Article

Measuring and Promoting Self-Regulation for Equity and Quality of Online Learning: New Evidence from a Multi-Institutional Survey during COVID-19

Jiao Guo 1,*, Ronnel B. King 2, Qinnan Ding 3 and Miaomiao Fan 1

1 Institute of Higher Education, East China Normal University, Shanghai 200062, China; 51204105006@stu.ecnu.edu.cn
2 Centre for the Enhancement of Teaching and Learning, Faculty of Education, University of Hong Kong, Hong Kong, China; rking@hku.hk
3 National Institute of Education Policy Research, East China Normal University, Shanghai 200241, China; qnding@admin.ecnu.edu.cn
* Correspondence: jguo@ses.ecnu.edu.cn

Abstract: Self-regulation is a core concept to understand the metacognitive, motivational, and emotional aspects of learning. The outbreak of COVID-19 resulted in large numbers of courses being shifted online, thus providing a large-scale setting to collect new empirical evidence to shed light on the specific challenges that different learner subgroups struggle with in the authentic online learning environment and then to provide practical implications to improve the learning outcomes by promoting learners’ online self-regulation. Based on a survey of 64,949 participants enrolled at 39 universities in a metropolitan city of China during the pandemic, we developed the Undergraduate Online Self-regulated Learning Questionnaire (UOSL), tested the reliability and construct validity of the UOSL items, and then built regression models to estimate the associations between online self-regulation and mastery of eight skills across different subgroups in a diverse student body. Disadvantaged subgroups such as rural, first-generation college students reported significantly lower UOSL scores as well as lower skill mastery in online learning. After controlling self-regulation in the regression model, these gaps related to student demographics have shrunk substantially and some become statistically insignificant (e.g., the gender gap in online skill mastery). The findings highlight the critical role of the targeted interventions of self-regulation to promote equity and enhance quality in the online teaching design and learning support.

Keywords: self-regulation; skill mastery; equity; quality; online teaching and learning

1. Introduction

The outbreak of COVID-19 triggered a large-scale learning environment switch to online. Such an unexpected massive switch provided a quasi-experimental opportunity to compare the two different learning settings, specifically, the online learning during COVID-19 (i.e., the treatment) and the traditional face-to-face experience right before the switch (i.e., the control). Our study aims to leverage this opportunity to investigate whether a substantial gap in learning outcomes associated with self-regulation across different learner subgroups. The investigation has policy implications to estimate the long-term effect of the pandemic on learning outcomes through the massive switch to online, in particular to address the concern for equity and quality relevant to the disadvantaged subgroups who face more challenges in the online learning settings during the pandemic.

To focus on non-academic outcomes beyond standardized testing in basic education, the participants of our study were college students in a metropolitan city of China. Despite that the city is classified as high-income group by income level 2020–2021, the city serves the population similar as a medium-size developing countries, e.g., Niger or Śri Lanka.
Furthermore, the lockdown policy during the pandemic caused campus closure. The participants of our survey from rural, low-income families face similar challenges (e.g., lack of computers or laptops, unaffordable bandwidth, heavy housework and other learning distractions) that trouble online learners in underdeveloped countries. Our survey results, thus, contribute to the specific theme of this special issue (i.e., to promote equity and enhance quality of online teaching and learning) by providing a case study of an emerging economy that is still undergoing social economic transition and is in the urgent need for equity and sustainability.

Thanks to the large sample size (N = 64,949) and a diversity of student body (e.g., 18.3% participants from rural families and 48.6% were first-generation college students, see the sample distribution for more details), we designed a multi-institutional survey including an instrument to measure online self-regulation across student subgroups and shed light on the inequity in self-regulation associated with learners’ background. A hypothetic link between self-regulation and learning outcomes was then tested to search for the practical implications to promote self-regulation for enhancing the quality of online teaching and learning.

2. Literature Review

Self-regulation is a core concept to approximate the nature of learning. Since the first attempts to distinguish self-regulation from other psychological drivers in the learning process three decades ago [1–3], self-regulated learning (SRL) has developed a sophisticated and evolving conceptual framework with theoretical models and empirical evidence. Three handbooks have been published to disseminate the comprehensive progress of SRL [4–6]. The systematic reviews of literature and meta-analysis studies on different aspects of SRL (e.g., in primary education [7], for work-related training [8], on medical simulation-based learning [9]) have also been presented to push forward the field.

The following literature review is prepared to justify our research design from two perspectives: one is a theoretical perspective on the conceptual framework and the empirical measures (i.e., surveys and other self-reporting instruments) of SRL; the other is a practical lens with a focus on the prior empirical studies to reveal the inequity in self-reported SRL cross different learner subgroups, to link SRL with the measures for quality of teaching and learning (e.g., GPA for academic outcomes and student engagement for non-academic outcomes), and to review the innovation for teaching design and learning support for promoting learners’ online self-regulation.

2.1. Theoretical Models and Empirical Instruments of SRL

A fundamental three-phase construct of SRL has been identified by prior theoretical reviews [10]. The three phases (i.e., preparatory, performance, and appraisal) form a feedback loop to map learners’ observable strategic efforts such as setting goals, monitoring progress, and reflection. Within each phase, there are subtle and sophisticated subphases that reveal the intellectual curiosity and theoretical background of different SRL scholars. Among the four conceptual models of SRL with thousands of citations for each [11], Zimmerman (2000) built a three-phase model of SRL, i.e., forethought, performance, and self-reflection [12]. Pintrich (2000)’s four-phase model is essentially similar to the fundamental three-phase construct, except that he split the second phase into two subphases as monitoring and control [13]. Winne and Hadwin (1998) also designed a four-phase SRL model by proposing two subphases (i.e., reflection and adapation) in the third phase [14]. Boekaerts and Corno (2005) adopted the core of three-phase construct (i.e., goal setting, goal striving, and performance feedback) within a dual processing model to emphasize the moderating role of emotion [15]. The positive emotion inspires a skill-expansion model of SRL for growth while the negative emotion triggers a well-being model of SRL for prevention or safety concerns. The three-phase construct of SRL, therefore, is the key concept underpinning our survey design of this paper.
The growing interest in the field of SRL has led to the development of diverse instruments to measure the cognitive, metacognitive, motivational, behavioral, emotional, and other aspects of SRL. Based on a systematic review of literature on the instruments measuring SRL in higher education (i.e., the target educational level of this paper), self-reporting instruments are most frequently used [16]. In comparison with interviews, think-aloud technique and writing diaries, questionnaires are the most common self-reporting method (i.e., 87.1% of the 225 instruments in the systematic review) to assess the different theoretical components of SRL. Near half of the questionnaires reviewed used the complete version or certain parts of an established instrument, i.e., Motivated Strategies for Learning Questionnaire (MSLQ) [17]. MSLQ has been translated into Chinese and has inspired the development of new instruments that consist of a compilation of MSLQ items and self-developed items. One of the MSLQ-inspired questionnaires developed by two Chinese scholars, Huang and Zhang, the Student Survey on Online Learning Experience (SSOLE), is a major source of survey design for our study [18]. Another major source for our survey design is the Online Self-regulated Learning Questionnaire (OSLQ) developed by Lucy Barnard and her coauthors [19] that is appropriate to measure SRL in the online learning environment. More details about the two instruments and their inspirations (e.g., specific dimensions and individual items) for our survey will be elaborated in the part on measurement.

2.2. Practical Findings to Link SRL with Equity and Quality of Online Teaching and Learning

This evidence-informed model building efforts of scholars to link SRL with learning environment and student outcomes empower education practitioners to unpack some myths (e.g., goal-driven, over-achievers from rural, low-income families attending non-elite colleges who leverage the online learning to improve their skills substantially) by the inclusion of learner background into analysis [20]. The literature of online SRL identified that demographics (e.g., age, gender, country of origin), personal background (e.g., educational attainment, social economic status), and other person-level differences (e.g., prior online learning experience) are linked with learners’ self-regulation and learning outcomes [21–23], but most findings were limited to the Massive Open Online Courses (MOOCs) whose target population (e.g., lifelong learners) and outcome measures (e.g., watching course lectures) are different from college students and their learning goals. Our study aims to extend the prior work to more personal-level differences (e.g., the subgroup of first-generation college students, the students attending elite universities, the students studying science, engineering, agriculture, or medicine programs) relevant to the theme of equity and quality in higher education.

The critical role of SRL plays in the learning process is proven for both academic performance such as course grades [24] and non-academic outcomes such as student engagement [25] that matter more for long-term knowledge retention and transferrable skills. The associations between SRL and learning outcomes have supportive evidence across learning environments, but the estimated effect sizes of SRL vary substantially. With an aim to enhance the quality of online teaching and learning by promoting learners’ self-regulation, it urges researchers to further identify which SRL strategy is most effective where, when, and for whom. The effect of SRL is assumed to be more important in the online learning environment with more autonomy for learners and less guidance from instructors than that in the traditional face-to-face learning [11,19]. This assumption about SRL and online learning, however, has been challenged by empirical findings. A meta-analysis of 12 studies on academic outcomes (e.g., GPA) found that the effect of SRL in the online setting is weaker than that in the face-to-face classroom [26]. Our study contributes to the latter thread of investigation on the associations between SRL and non-academic outcomes (i.e., skill mastery) through a comparative lens of online vs. face-to-face learning. Our findings built on the comparison, then, provide practical implications on future design of SRL-related interventions (with a focus to support rural, low-income, first-generation...
college students or other vulnerable subgroups) to promote equity and enhance quality of online teaching and learning.

3. Context

The metropolitan city in the current study is located in East China. In 2020, its population was near 25 million. In comparison with developing countries, the population size of the current city is smaller than that of Korea, DPA., but larger than that of Niger or Sri Lanka [27]. The statistics of income level demonstrated that the GDP per capita of the city was around 25,000 (current US$) in 2020, which can be classified into the high-income group and is similar to that of Slovenia, Bahamas, or Kuwait by income level 2020–2021 [28].

The higher education system in the city consists of four types of higher education institutions: 40 four-year universities, 23 three-year vocational colleges, 14 institutions for adults, and 227 private institutions [29]. The higher education enrollment statistics from the website of China Ministry of Education excludes the institutions for adults and those for non-degree programs. The total enrollment in 2020 of the city was 399,986 students in the four-year universities pursuing a bachelor’s degree and 140,709 students in the three-year vocational colleges for an associate degree as well as a vocational certificate. The gender ratio in the total enrollment for degree programs in the city was around 107 females per 100 males [30]. More detailed data on student social economic status such as region of birth and family background are unavailable. Filling this gap in the statistics data by a multi-institutional survey or other instruments is a critical step to shed light on any hidden inequity of the higher education system.

The teaching and learning in higher education were switched to online in 2020 spring semester because of the outbreak of COVID-19. The facts and figures released by the Division of Higher Education in China Ministry of Education indicated that it was an unprecedented large-scale evolution in online learning. About 1.08 million instructors and 22.59 million students in 1270 four-year universities all over the country participated in the online teaching and learning from March to August in 2020. The overall student satisfaction rate for online teaching and learning during the pandemic reached 85% [31]. These numbers exclude vocational colleges, no-degree programs, and institutions for adults.

The statistics about online learning in the city’s formal higher education system were unavailable, but case studies from individual institutions reveal the special learning needs of different student subgroups during the pandemic. For example, a public university in the city under study delivered laptops for ethnic minority students who were locked down in the dormitories due to the outbreak of COVID-19. Furthermore, the university provided a total amount of 163,300 RMB (around 23,667 US$ in 2020) bandwidth vouchers for all the students from low-income families who struggled with Internet access for the online learning [32]. Another case is a private university in the city that supported the disadvantaged students not only with the cash aid for those whose parents or who themselves severely affected by COVID-19, but also with a special aid to cover 30 GB data transfer for 30 days per student. For the students whose hometown were in Wuhan, the epicenter of COVID-19, they were encouraged by the university to choose and buy e-books when learning at home. The book purchase was reimbursed by the university with a maximum of 200 RMB (around 30 US$ in 2020) per student [33]. The third case is a public elite university of the city that conducted a survey for over 2000 instructors and near 30,000 students on the Internet usage. The survey result demonstrated that only 10 students were unable to get access to Internet while 2500 students suffered from slow data transfer and 12,400 students self-reported struggling with online learning. The survey result demonstrated that only 10 students were unable to get access to Internet while 2500 students suffered from slow data transfer and 12,400 students self-reported struggling with online learning. The university then conducted one-to-one in-depth student interviews to figure out the barriers in online learning during the pandemic. Instructors were encouraged to hold online office hours by e-learning and other platforms suitable for text or voice messages. Furthermore, near 100 students volunteered to learn the platform user handbooks, to pass an online evaluation, and then to play the role of virtual teaching assistants for more than 1000 instructors in online teaching [34].
Despite of its high-income economic status and female advantage in enrollment, the city in our study provides the following unique insights to promote equity and enhance quality of online teaching and learning during the pandemic. First, the higher education system of the city serves a diverse student body including those from rural, low-income families who are missing from previous national or regional statistics and who face essentially the same challenges as students in underdeveloped countries when studying from home online. Second, case studies from individual universities during the pandemic highlight diverse needs (from laptop donation to personalized e-learning solutions) of disadvantaged subgroups for the online learning ranging. Third, the needs mirror the emerging literature on a pivot of digital divide from the Internet access to the inequity of learner usage, skills, and reflection [35,36]. This pivot is relevant to the focus of our study on the hidden divide of SRL across learner subgroups. Last but not least, the empirical results of prior survey-based SRL studies in China support the positive associations between online self-regulation and learning outcomes (e.g., a link between the use of reflection strategy and the growth of critical thinking skills), but their sample sizes are smaller than 1000, their learning environments are blended [37] or in a MOOC platform before the pandemic [18], and their research priorities are data-analytic methods (e.g., process mining) instead of the purpose to promote equity and enhance quality. Thus, our study contributes new empirical evidence based on a large-scale survey during the outbreak of COVID-19 to capture the importance of SRL on skill mastery in the online learning environment.

4. Materials and Methods

4.1. Participants

Data were collected by a multi-institutional survey during the outbreak of COVID-19 in 2020 summer. The survey is exempt from IRB review based on the following guidelines after the consultation with our IRB office: one is that the survey, interview, or observation is used only for the purpose of teaching and learning and that students are unable to be directly or indirectly identified by the data collection process, the other is that the research project is contracted, sponsored, or approved by government. Our survey was a project sponsored by Ministry of Education to improve online teaching and learning effectiveness. The individual information is confidential due to a standard non-disclosure agreement and a strict process of data collection, storage, and analysis. Furthermore, a brief consent form was included in the survey introduction for participants to permit their submitted information to be used for the purpose of academic research and policy evaluation.

An invitation to participate the online survey was sent to all the 40 four-year universities in a metropolitan city in China. About 39 universities participated and only one university submitted zero response. A total of 84,720 students responded and completed the survey. The sample size is slightly over 21% of the undergraduate population in the city. The reason is that once the survey response rate reached 20%, i.e., the expected percentage set by the survey team, the online survey system monitored this progress and then stopped collecting data.

A pilot study had been conducted to estimate the time that was required to complete the survey. Over 30 participants of the pilot study spent about 30–45 min to complete and submit the survey. A follow-up research team meeting then designated that response time less than 5 min is too short and that more than an hour is too long. After deleting the participants who submitted their surveys within too short or too long time, the final valid sample for the study includes 64,949 students.

Given that the focus of our investigation is to compare between subgroups and identify the heterogeneity of association between online SRL on skill mastery if it exists, the demographic and educational background diversity of sample composition is a major concern to address the data-collection process. In comparison with the undergraduate population of the city, the percentage of female students in our sample (61.2%) is higher than that of the population (51.7%), but comparable to a prior study in China on SRL in blended learning environment (for reference, their sample percentage of females is 62%) [18].
statistics of the population, however, are unavailable for us to check the representativeness of our sample. See Table 1 for the sample distribution of selected variables.

Table 1. Sample distribution of selected variables (N = 64,949).

<table>
<thead>
<tr>
<th>Selected Variables</th>
<th>Categories</th>
<th>Sample Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female</td>
<td>61.2</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>38.8</td>
</tr>
<tr>
<td>Household residency</td>
<td>Rural</td>
<td>18.3</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>81.7</td>
</tr>
<tr>
<td>Family social-economic status</td>
<td>First-generation college students</td>
<td>48.6</td>
</tr>
<tr>
<td></td>
<td>Second-generation college students</td>
<td>51.4</td>
</tr>
<tr>
<td>Grade year</td>
<td>Freshman</td>
<td>41.0</td>
</tr>
<tr>
<td></td>
<td>Sophomore</td>
<td>32.1</td>
</tr>
<tr>
<td></td>
<td>Junior</td>
<td>25.3</td>
</tr>
<tr>
<td></td>
<td>Senior</td>
<td>1.6</td>
</tr>
<tr>
<td>Subject fields</td>
<td>Humanities and social sciences</td>
<td>60.4</td>
</tr>
<tr>
<td></td>
<td>Science, engineering, agriculture, and medicine</td>
<td>39.6</td>
</tr>
<tr>
<td>University types</td>
<td>Elite universities</td>
<td>24.2</td>
</tr>
<tr>
<td></td>
<td>Other universities</td>
<td>75.8</td>
</tr>
</tbody>
</table>

1 Family social-economic status is measured by educational attainments of parents. First-generation college students in our study refer to students whose parents never went to college. 2 The response rate was low for college seniors (i.e., class’20) because the survey was conducted in June 2020 when most seniors already graduated and few of them took online courses during COVID-19. 3 Elite universities in our research design refer to 42 universities selected by China’s Ministry of Education for the World-class University Construction and other 95 universities selected for the World-class Discipline Construction in 2017. Both constructions are known in China as the Double World-class Construction (DWC) project in higher education.

4.2. Measures

The outcome variable to measure the non-academic outcomes in our survey is the participant’s self-rated skill mastery. Eight skills are rated, including three subject-specific skills (basic knowledge, frontier research, and practical skills) and five transferrable skills (effective communication, information process, critical thinking, problem solving, and planning). Each of the eight skills is rated by participants with a Likert scale ranging from one (“no improvement”) to five (“improve a lot”). The survey platform sent a reminder to the participants for rating their skill mastery from the online learning during COVID-19 through a comparison lens against the face-to-face learning before the pandemic.

The question variable in the study is measured by the Undergraduate Online Self-Regulated Learning Scale (UOSL). The UOSL has been built on two instruments introduced in the part of literature review: the OSLQ and the MSLQ-inspired SSOLE [18,19]. The OSLQ instrument is more established and includes 24 items of six subscales, i.e., goal setting, environment structuring, task strategies, time management, help seeking, and self-evaluation. The SSOLE (in Chinese only) consists of 11 items of three phases, i.e., goal setting and planning, execution and behavior adjustment, and reflection and cognitive monitoring. The maximum number of UOSL items is set by our survey team at six. Two items for each of the three fundamental phases identified by prior literature review on SRL, i.e., preparatory, performance, and appraisal, have then been selected from the OSLQ and the SSOLQ. The items selected from the OSLQ has been translated from English to Chinese and then revised for comparison, for example, the original OSLQ item is I set short-term (daily or weekly) goals as well as long-term goals (monthly or for the semester) while the revised UOSL item, I plan for short-term and long-term goals more frequently, emphasizes the comparison between online and face-to-face learning settings.

The Cronbach’s alpha value of UOSL items is 0.942, which proves the internal consistency of the scale. In order to investigate the latent construct of the UOSL items, an exploratory factor analysis (EFA) was conducted to test the construct validity. The EFA
results demonstrated that one factor is sufficient for the six items to measure the online self-regulated learning. All standardized factor loadings of the six items of UOSL scores range from 0.79 to 0.89.

4.3. Hypothesized Associations

Our investigation aims to test the two hypotheses by the following equation:

$$\text{Skill} = \beta_0 + \beta_1 \text{UOSL} + \sum_{j=2}^{n} \beta_j X_j$$

(1)

where $\beta_0$ is the constant term.

1. SRL is positively associated with skill mastery of online learning. The hypothesized association can be proven by the positive and statistically significant coefficient, $\beta_1$, of participants' self-rated UOSL scores.

2. The background variables, $X_j$, in the hypothesized regression model include gender, household residency, educational attainments of parents, subject fields and institutional types (i.e., elite universities versus other four-year universities). Males, rural residents, first-generation college students, SEAM (science, engineering, agriculture, or medical) majors, and undergraduates at non-elite universities are hypothesized as disadvantaged subgroups who face more challenges of online learning than other subgroups. The corresponding set of coefficients, $\beta_j$, are hypothesized negative and statistically significant.

5. Results

5.1. Descriptive Statistics

We began with descriptive statistics of the survey data before conducting further analysis. As a sum of total six items of the UOSL scores, participants reported an average value of 19.64 (Median = 18, SD = 4.246) with a min of 6 and a max of 30. Due to the large sample size ($N = 64,949$), the self-rated UOSL scores approximate a normal distribution. Specifically (see Table 2), the highest SRL item rated by participants is that they are better at adjustment for course pace in online setting than the traditional face-to-face learning environment (Mean $= 3.347$), which supports the expected high autonomy of online learning; the lowest SRL item is that participants set short- or long-term goals more frequently online than the face-to-face learning (Mean $= 3.197$).

Table 2. Descriptive statistics of the self-rated UOSL scores ($N = 64,949$).

<table>
<thead>
<tr>
<th>UOSL Items</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I plan for short-term and long-term goals more frequently.</td>
<td>3.197</td>
<td>0.803</td>
</tr>
<tr>
<td>2. I am better at completing weekly assignments $^1$.</td>
<td>3.335</td>
<td>0.835</td>
</tr>
<tr>
<td>3. I am better at adjusting reading methods for challenge materials.</td>
<td>3.251</td>
<td>0.801</td>
</tr>
<tr>
<td>4. I am better at adjusting learning methods to cope with the course pace.</td>
<td>3.347</td>
<td>0.801</td>
</tr>
<tr>
<td>5. I am better at summarizing what I learned.</td>
<td>3.245</td>
<td>0.791</td>
</tr>
<tr>
<td>6. I ask myself if I understand the new content more frequently.</td>
<td>3.267</td>
<td>0.790</td>
</tr>
<tr>
<td>TOTAL</td>
<td>19.64</td>
<td>4.246</td>
</tr>
</tbody>
</table>

$^1$ There is an introduction sentence of the UOSL items to remind the survey participants to compare their online learning during COVID-19 to the face-to-face learning before the pandemic.

Based on the t-test to compare the mean scores between different subgroups, we identified that males, rural students, first-generation college students, SEAM majors and undergraduates at elite universities reported an average lower UOSL score in online vs. face-to-face learning comparison, which implies that they face more challenges in self-regulation than other subgroups for the large-scale switch to online learning during the pandemic. All the self-reported gaps between subgroups are statistically significant (see Table 3) and all
the results (except that relevant to elite universities) support the hypothesized vulnerable subgroups in online SRL. The $t$-test comparison between subgroups provides a similar pattern in skill mastery. The outcome measure is the total sum of participants’ self-rated mastery of eight skills (Mean = 25.73, Median = 24, SD = 5.475). The subgroups who reported lower level of online SRL also reported the lower average sum of skill mastery from online learning during COVID-19.

Table 3. $t$-test to compare the mean total UOSL scores and skill mastery of online learning between different subgroups (N = 64,949).

<table>
<thead>
<tr>
<th></th>
<th>USOL</th>
<th>Skill Mastery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Female</td>
<td>19.867</td>
<td>3.964</td>
</tr>
<tr>
<td>Urban residents</td>
<td>19.748</td>
<td>4.287</td>
</tr>
<tr>
<td>Second-generation college students</td>
<td>19.873</td>
<td>4.411</td>
</tr>
<tr>
<td>Non-SEAM majors</td>
<td>19.801</td>
<td>4.211</td>
</tr>
<tr>
<td>Students at elite universities</td>
<td>19.544</td>
<td>4.179</td>
</tr>
<tr>
<td>Students at non-elite universities</td>
<td>19.673</td>
<td>4.267</td>
</tr>
</tbody>
</table>

$^* p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$.

5.2. Correlation Analysis

We conducted the correlation analysis to confirm the associations between learners’ background, online self-regulation, and skill mastery (see the corresponding columns in Table 4). The pattern remains essentially the same. We identified negative correlations between online SRL and four background variables (i.e., males, rural residency, first-generation college students, and SEAM majors) as the Hypothesis 2 (see Section 4.3) predicts. Only the correlation between online SRL and non-elite university is positive and opposite to the hypothesis. All the correlation coefficients are statistically significant, but weak, ranging from 0.013 to 0.071.

Table 4. Correlations between learners’ background, self-regulation, and skill mastery of online learning (N = 64,949).

<table>
<thead>
<tr>
<th></th>
<th>Self-Regulation (USOL Scores)</th>
<th>Skill Mastery</th>
</tr>
</thead>
<tbody>
<tr>
<td>males</td>
<td>−0.067 ***</td>
<td>−0.055 ***</td>
</tr>
<tr>
<td>Rural residents</td>
<td>−0.053 ***</td>
<td>−0.053 ***</td>
</tr>
<tr>
<td>First-generation college students</td>
<td>−0.056 ***</td>
<td>−0.065 ***</td>
</tr>
<tr>
<td>SEAM majors</td>
<td>−0.045 ***</td>
<td></td>
</tr>
<tr>
<td>Students of non-elite universities</td>
<td>0.016 ***</td>
<td>0.013 ***</td>
</tr>
</tbody>
</table>

$^* p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$.

5.3. Regression Analysis

5.3.1. Baseline Model with Learners’ Background

The baseline regression model includes a set of learners’ demographic and educational background variables associated with learning outcome. All regression coefficients are
statistically significant with the same direction as Hypothesis 2 predicted (see Section 4.3, hypotheses, bullet 2), except the coefficient of non-elite universities. See the detailed results of Model 1 in Table 5. The findings relevant to demographics are consistent. Males, rural residents, and first-generation college students face more challenges of self-regulation when studying from home online during the pandemic. Their self-reported skill mastery from online learning is lower than other subgroups and all the gaps are statistically significant ($p < 0.001$). The findings relevant to educational background are mixed. Lower division of undergraduates (e.g., freshmen or sophomore) and SEAM majors reported lower skill mastery of online learning than other subgroups. The regression coefficient of non-elite universities, however, is positive and opposite to the hypothesis. This perplexing pattern will be interpreted further in the discussion section.

Table 5. Regression analysis of learners’ background, self-regulation and skill mastery of online learning ($N = 64,949$).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1 Baseline</th>
<th>Model 2 +Self-Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>$-0.041^{***}$</td>
<td>$-0.004$</td>
</tr>
<tr>
<td>Rural residency</td>
<td>$-0.029^{***}$</td>
<td>$-0.008^*$</td>
</tr>
<tr>
<td>First-generation</td>
<td>$-0.050^{***}$</td>
<td>$-0.024^{***}$</td>
</tr>
<tr>
<td>Grade year</td>
<td>$0.031^{***}$</td>
<td>$0.017^{***}$</td>
</tr>
<tr>
<td>SEAM majors</td>
<td>$-0.044^{***}$</td>
<td>$-0.028^{***}$</td>
</tr>
<tr>
<td>Non-elite universities</td>
<td>$0.012^{**}$</td>
<td>$0.004$</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>$0.637^{***}$</td>
<td></td>
</tr>
</tbody>
</table>

$df$ 64942 64941  
adjusted R$^2$ 0.011 0.413  
F-value 119 6521

$^1$ $p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$, the standardized regression coefficients are reported.

5.3.2. Self-Regulation and Skill Mastery of Online Learning

After adding the question predictor, self-regulation, to the regression model (see Model 2 in Table 5), the variation of self-rated skill mastery explained by the model increased substantially. The adjusted $R^2$ value increased from 0.011 in the baseline model to 0.413 in the model 2 with UOSL scores. The standardized regression coefficient of UOSL scores, $\beta_1$, is positive, statistically significant, and the effect size is 0.637, which supports our hypothesis on the strong association between online SRL and learning outcomes. Furthermore, after controlling the individual-level difference in self-regulation, the gaps in skill mastery of online learning between different demographical subgroups shrunk substantially. Specifically, the self-reported skill mastery gaps related to rural residency and that to first-generation college students decreased by 72% and 52% in effect size, respectively. The gender gap of skill mastery shrunk by 90% and became statistically insignificant. The similar pattern was illustrated with educational background. The gap in skill mastery of online learning associated with a linear variable of the grade year decreased by 45% and that associated with a dummy variable of SEAM majors reduced by 36% in effect size. The predicted advantage associated with non-elite universities in online learning effect became statistically insignificant.

5.3.3. Robustness Check to Identify Specific Challenges for Each Subgroup

To test the robustness of regression results, each of the eight self-reported skills instead of their sum were taken as learning outcome measures of the regression model (see Table 6). The coefficients of the self-regulation are consistent across the eight model specifications. All the coefficients of self-regulation are positive, statistically significant, and with an effect size ranging from 0.462 to 0.584. These results confirm the important role of self-regulation...
in online learning context for all eight skills in general and for some transferable skills (e.g., critical thinking, problem solving, and planning) in particular.

Table 6. Regression analysis of eight self-rated skills of online learning by undergraduate subgroups on self-regulation (N = 64,949).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Basic Knowledge</th>
<th>Frontier Research</th>
<th>Practical Skills</th>
<th>Communication</th>
<th>Critical Thinking</th>
<th>IT Competency</th>
<th>Problem-Solving</th>
<th>Planning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.003</td>
<td>0.022***</td>
<td>0.028***</td>
<td>-0.012***</td>
<td>0.010***</td>
<td>-0.016**</td>
<td>-0.031***</td>
<td>-0.007*</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.011**</td>
<td>-0.003</td>
<td>0.000</td>
<td>-0.019***</td>
<td>-0.003</td>
<td>-0.016**</td>
<td>-0.016*</td>
<td>0.006</td>
</tr>
<tr>
<td>First-generation</td>
<td>-0.027***</td>
<td>-0.025***</td>
<td>-0.015***</td>
<td>-0.025***</td>
<td>-0.022***</td>
<td>-0.019***</td>
<td>-0.021***</td>
<td>-0.014***</td>
</tr>
<tr>
<td>Grade year</td>
<td>0.027***</td>
<td>0.023***</td>
<td>0.015***</td>
<td>0.003</td>
<td>0.007*</td>
<td>0.018***</td>
<td>0.015***</td>
<td>0.011***</td>
</tr>
<tr>
<td>SEAM majors</td>
<td>-0.025***</td>
<td>-0.022***</td>
<td>-0.099***</td>
<td>-0.043***</td>
<td>-0.026***</td>
<td>-0.017***</td>
<td>-0.011***</td>
<td>-0.016***</td>
</tr>
<tr>
<td>Non-elite universities</td>
<td>-0.007*</td>
<td>0.016***</td>
<td>0.044***</td>
<td>0.003</td>
<td>0.005</td>
<td>-0.026***</td>
<td>-0.019***</td>
<td>0.008*</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>0.584***</td>
<td>0.544***</td>
<td>0.462***</td>
<td>0.530***</td>
<td>0.571***</td>
<td>0.568***</td>
<td>0.578***</td>
<td>0.569***</td>
</tr>
</tbody>
</table>

$df$ 64941 64941 64941 64941 64941 64941 64941 64941
Adjusted $R^2$ 0.347 0.300 0.230 0.290 0.329 0.329 0.342 0.327
F-value 4941 3984 2775 3793 4556 4542 4815 4506

$^1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, the standardized regression coefficients are reported.

Another purpose of the robustness check in Table 6 is to identify the specific challenge facing each subgroup, which is the key to design and deliver interventions tailored to the special needs of different subgroups. For example, the gender gap of online learning identified in Table 6 includes three skills for females (i.e., frontier research, practical skills, and critical thinking) and four skills for males (i.e., communication, IT competency, problem solving, and planning). Those from rural families mainly need support in transferrable skills learned online such as communication, IT competency, and problem solving. First-generation college students face challenges in all eight skills, especially in basic knowledge, frontier research, and communication. SEAM majors face the similar challenge in all eight skills and their top need is the support for practical skills training in the online learning environment. Students at elite universities self-reported more skill mastery in basic knowledge, IT competency, and problem solving while those at non-elite universities rated more skill mastery of online learning in frontier research, practical skills, and planning.

6. Discussion

The outbreak of COVID-19 resulted in large numbers of courses being shifted online, thus providing a large-scale setting to shed light on the specific challenges that different learner subgroups struggle with in the authentic online learning environment and then to provide practical implications to improve the quality of online teaching and learning by promoting learners’ self-regulation. Based on the survey of 64,949 participants at 39 universities in a metropolitan city of China, we developed the Undergraduate Online Self-regulated Learning Questionnaire (UOSL) and then built regression models to estimate the associations between learner background, self-regulation, and skill mastery in the online learning setting across different subgroups in a heterogenous student population.

The following discussion about our findings includes theoretical contributions and practical implications, which are the same two perspectives to frame our literature review and to extend our investigation. In addition to self-developed instrument to measure online self-regulation, one of our theoretical contributions is to highlight the gaps in self-regulation associated with different human factors (e.g., gender, grade year, family background), i.e., where the inequity in online learning lies. Practical implications built on the gaps identified, then, shed light on the design principles and learning supports (e.g., a pre-course reminder of goal-setting sent to students) to improve the quality of online teaching and learning by promoting learners’ online self-regulation. Limitations and future research directions are also addressed in the discussion.
6.1. Theoretical Contributions on Measuring SRL and Highlighting the Importance of Human Factor

Our investigation made two contributions to the literature of SRL. First, we developed Undergraduate Online Self-regulated Learning Questionnaire (UOSL). UOSL is based on the parts of two established instruments (i.e., OSLQ and MSLQ-inspired SSOLE [18,19]) to measure online SRL from a fundamental three-phase construct (i.e., preparatory, performance, and appraisal) [1,11]. Both reliability and construct validity of UOSL were proven with our survey data. With six items only, UOSL is easy to be applied to a combination of measurement in a large-scale survey. The combination enables researchers to use one survey to simultaneously collect data from other measurement such as social and emotional scales that the literature on SRL lacks [16]. For example, our survey results demonstrate the positive, statistically significant, and strong association between SRL and skill mastery during COVID-19, which supports the skill growth model in Boekaerts and Corno (2005)’s dual processing framework of SRL [15]. This finding implies the positive moderating role of emotion that dominates the online learning process of our survey participants despite of the disruptions and other emotional anxiety triggered by the pandemic. The dual processing framework, both cognitive and emotional self-regulations, needs further investigation.

Second, our results highlight the importance of human factors (e.g., gender, social economic status) to identify if the inequity in learners’ self-regulation exits in the authentic online learning environment. The concern for human factors can be traced back to Zimmerman’s early publications that unpack the concept of SRL from a social cognitive perspective to emphasize the biological, developmental, contextual, and other individual constraints on learners’ self-regulation [1,20–23,38]. Age, gender, experience, and other individual-level differences associated with SRL have been summarized in the systematic reviews of empirical studies about online learning in general and on MOOC platforms in particular [39–41]. Our survey-based findings extend this thread of investigation by adding new evidence. For example, prior studies identified that female and older learners are better at time management and other SRL strategies in online learning [42]. Our findings confirmed the similar advantages of females and college juniors or seniors in online SRL as well as skill mastery. Moreover, our regression results demonstrated that the gaps cross subgroups on skill mastery shrunk substantially after controlling SRL, which indicates that disadvantaged students have the potentials to catch up if they are trained and supported appropriately on SRL strategies. This potential scenario has been verified in prior experimental studies on SRL-related training and prompts [43], which leads to the discussion on practical implications to promote equity and enhance quality of online teaching and learning in higher education.

6.2. Practical Implications on Promoting SRL for Equity and Quality of Online Teaching and Learning

The literature on SRL-related support for online teaching and learning (e.g., to support MOOC leaners) is growing, which indicates a transition in the field from measuring SRL to promoting SRL [41]. The assumption for this transition is that SRL is a key element for online learning success and that learners struggle if they do not use the critical SRL strategies [38]. Our findings reveal the assumed SRL gaps cross different subgroups and highlight the specific struggling identified for each subgroup in online learning. Another assumption that learner’s self-regulation is neither fixed nor spontaneous has also been proven by prior empirical studies [43], which has justified the increasing practices that train and support learners to use the SRL strategies online. The training or support is more important for disadvantaged subgroups such as the students from rural, low-income families who reported low SRL scores and low skill mastery in our study. We encourage future interventions targeting at the specific challenges of those subgroups in higher education, e.g., an effective communication workshop tailored to males, rural residents, first-generation college students, and SEAM majors in the context of online learning. The interventions reflect the importance of equity that recognizes the special needs of each
subgroup who faces different constraints and demands different support for self-regulation to reach an equal outcome in online learning. This demand for equity is demonstrated, by our empirical findings, more urgent than the demand for equality that means providing the same Internet access, platform features, or other resource for each subgroup of a diverse student body in the online learning environment.

Prior systematic reviews of the literature summarized the design principles to promote online SRL in MOOC or other learning platforms for instructors and students of higher education. A review of 21 empirical studies identified different MOOC system features for each of the three fundamental phases of self-regulation [44]. For example, in the planning phase, MOOC platforms should offer the possibility for learners to identify the tasks (e.g., specific courses to take and external resources to leverage), to predict the time commitment, and to set the learning schedule [45]. This design principle is actionable for other instructors of online teaching in higher education. A pre-course reminder should be sent to learners to support their self-regulation on planning, e.g., to write down their short-term and long-term goals for taking this online course, to estimate the learning workload and to make the weekly study plan. Some learners struggle with self-regulated planning because their course schedule is already too tight (e.g., SEAM majors struggling with planning in our study who often need submit more assignments and complete more credit hours to graduate) while other subgroups (e.g., first-generation college freshmen in our sample who are often unfamiliar with learning support and other academic resource on campus) may need the navigation to use the additional social cognitive resource for planning. This practical implication is relevant to academic planning, curriculum reform, freshman orientation, online advising or tutoring, platform improvement, and other related changes that need to be made in higher education.

The findings of our study provide similar practical implications for other two phases of SRL (i.e., performance and appraisal) to promote SRL online. Take the SRL-related prompts as an example, guiding questions (e.g., What are the key points?), problem-solving suggestions (e.g., Take notes and highlight the main disagreement of the three parties), or reflection prompts (e.g., Rate the confidence level of your understanding about this topic) should be integrated into the online learning process to promote learners’ self-regulation with a focus on addressing the need of disadvantaged subgroups, for example, in similar practices as a workshop for the rural students to adjust their reading strategies for difficult course materials (i.e., a specific SRL-related challenge identified by our study).

6.3. Limitations

Two limitations of our research design should be noted. One limitation is external validity of our sample. Participants of the study are college students in a metropolitan city of China during COVID-19. Due to the campus closure during the pandemic, disadvantaged subgroups (e.g., rural, first-generation college students) had to study from home online. They faced similar online learning challenges (e.g., poor bandwidth coverage in rural villages) as the learners in low-income neighborhoods of developing countries. We are fully aware of the potential impact of Chinese culture on the findings. For example, prior work identified a maladaptive-but-engaged student subgroup in China who were encouraged to transform inadequacy, failure, or other obstacles as an opportunity for improvement [46]. This cultural context is likely to be intertwined with self-regulation and other metacognitive or psychological mechanisms associated with skill mastery of online learning. Our conclusions need further investigation with future empirical evidence collected from other countries.

The other limitation is internal validity of our instrument. Both the measurement of online SRL and skill mastery were self-rated by participants of our survey sample. For example, the participants attending elite universities in our sample report on average lower SRL scores and skill mastery of online learning than those at non-elite universities. A threat to validity for this finding is that the students of elite universities are likely to have a higher expectation of themselves about their online learning performance. Further investigation
needs to address this self-reporting bias, and more discussion to interpret this finding is provided below (including the insider view of the subgroup of elite-university participants who shared with our survey team). Although learning behaviors can be observed and recorded in an authentic online learning environment, we argue that metacognitive, psychological, or other hidden mechanisms should still be measured by self-reporting instruments [16] that are tailored to the population of interest and the context of culture. A combination of multiple sources consisted of observed behaviors, self-reported perceptions, platform logs, and experimental results will be discussed further as a direction for future investigation.

6.4. Future Research Directions

Our study yielded some inspiring findings for future research. One perplexing pattern in the results is the associations between university type, self-regulation, and skill mastery of online learning. After controlling learners’ background, students at elite universities reported lower level of SRL and lower-rated skill mastery. This pattern was opposite to our hypothesis, but revealed the essential context-dependent nature of SRL, apart from the concern about the self-reporting bias that we already discussed. SRL is highly context-dependent, which refers to that learning environment design is the key to promote SRL strategies, to shape learners’ behaviors, and then to make these behaviors into habits. For example, students at elite universities were separated not only from labs, libraries, and other study space designed for facilitating advanced learning and scientific inquiry, but also separated from dorms, cafeterias, gyms, and other social space to interact with other students. Formal or informal interactions were the most missed activities reported by some participants of our survey in follow-up interviews. The interviewed students at elite universities were eager to receive feedbacks on hands-on lab skills from the face-to-face instruction or to play games while hang out with their friends in a basketball court or a soccer field. The social, emotional, interactive, or collaborative perspective is one of the future directions to extend our understanding about self-regulation in the online learning [47,48].

Another direction for further investigation is the improvement of research design about SRL. Specifically, we encourage three directions in measures and methods. First, in addition to questionnaires, other self-report instruments (such as in-depth interviews, think-aloud technique, learning diaries [16]) should be combined in the research design to leverage both quantitative and qualitative methods. Second, the current development of learning analytics enables the investigation of SRL from the traditional personal level to a finer-grained level of learning tasks [11]. Platform logs record the trace of each learner in the authentic online learning environment. Trace-based learning process mining and other log analytic instruments have been used to understand the association between SRL and motivation [49] or to predict the mastery of writing skill [50]. Third, to address the skeptical concern about the interventions consisted of scaffolding prompts (e.g., an alert tool to remind time allocation [51]), we encourage future intervention design to target at the heterogeneity of self-regulation associated with learners’ background. For example, writing down the personal intention to take an online course (i.e., related to the goal setting strategy in our study) was proven to be an effective intervention for MOOC learners from developing countries [52]. Hispanic students, males, low-achievers were identified facing larger challenges in online learning [53]. Randomized control experiments were conducted to make causal inference about the unbiased effect of those interventions across different learning environments [54,55]. More rigorous empirical evidence combining mixed method, learning analytics, and experimental design needs to be collected in large scale to promote self-regulation and skill mastery of online learners from disadvantaged background.

Author Contributions: Conceptualization, methodology, formal analysis, writing—original draft preparation, funding acquisition, J.G.; writing—review and editing, J.G. and R.B.K.; survey design, Q.D.; data curation, Q.D. and M.F.; literature review, project administration, M.F.; All authors have read and agreed to the published version of the manuscript.
Funding: This research was funded by the Peak Project of Faculty of Education at East China Normal University and also funded by the Undergraduate Online Learning Behavior Analysis Based on Trace-SRL project (grant number 21YJ A880014) of the Humanities and Social Science Research Planning Fund at China’s Ministry of Education.

Institutional Review Board Statement: The survey is exempt from IRB review based on the following guidelines after the consultation with the IRB office at East China Normal University: one is that the survey, interview, or observation is used only for the purpose of teaching and learning and that participants are unable to be directly or indirectly identified by the data collection process, the other is that the research project is contracted, sponsored, or approved by government. The survey in the current study was a project sponsored by China Ministry of Education to improve online teaching and learning effectiveness for the IRB exemption.

Informed Consent Statement: Written informed consent has been obtained from the participants to publish this paper.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the restrictions of participant privacy.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

22. Hansen, J.D.; Reich, J. Democratizing education? Examining access and usage patterns in massive open online courses. *Science* 2015, 350, 1245–1248. [CrossRef]


32. Education with Care: Shanghai Normal University Set COVID-19 Funds to Help Struggling Students. Available online: https://news.gmw.cn/2020-03/10/content_33636516.htm (accessed on 10 May 2022).


34. Information Office of Fudan University. Fudan University: 1+N Practice of Online Teaching Model. Available online: https://www.fudan.edu.cn/202108/t20210831_556506.html (accessed on 10 May 2022).


