ABSTRACT
In both flipped classroom settings and distance learning, educational content is typically delivered via video lectures that students watch alone. While flipped classrooms typically provide students with opportunities for social interaction that feature active learning, online learners are not typically afforded these opportunities. Cooperative learning techniques like Lightweight Teams [15] provide social, collaborative learning opportunities to students in flipped classrooms but extending these techniques to distance learning settings is not straightforward. In this paper, we present our experiences with online, distributed Lightweight Teams. We present an in-the-wild study that compares learning outcomes and student preferences between co-located and distributed Lightweight Teams against the base case of individual learning. Our results show that while there are no significant learning differences between the two team conditions and the individual condition, students significantly prefer the team conditions.

Keywords
Distance learning; Team-based Learning; Lightweight Teams; Distributed Teams; Synchronous Learning

Categories and Subject Descriptors
K.3.2 [Computer Science Education]: Computer and Information Science Education

1. INTRODUCTION
Flipped classroom learning environments have been receiving much attention over the past few years; partly because they help to provide more opportunities for active and engaged learning. Flipped classroom setups tend to foster collaborative peer learning, which has been shown to improve learning outcomes [14, 8, 1, 17]. Researchers are investigating how instructional design techniques that leverage collaboration in face-to-face classrooms can be extended or applied in distance learning. This research has become even more relevant with the recent surge of Massively Open Online Courses (MOOCs). Traditional distance learning classes and MOOCs both share similarities with flipped classrooms in that students typically consume educational content by watching lecture videos alone. However, with flipped classrooms, these lecture videos are then supplemented with face-to-face active learning activities in the classroom that are often collaborative in nature. In distance learning, this type of interaction either doesn’t happen at all or has to be mediated by technology. Keeping distributed students engaged and encouraging collaborative peer learning in distance environments is challenging. Extending effective face-to-face teaching techniques to an online context requires an understanding of the challenges that distance creates as well as strategies for mitigating those challenges [21].

We build on our previous work with Lightweight Teams [15], which were used in a face-to-face flipped classroom environment, by creating an online version. Lightweight Teams (LWT) are a cooperative learning approach in which teams of students are assigned to work together throughout a course, but where the activities performed by the team are lightweight and have little or no direct impact on each individual student’s grade. This approach leverages the benefits of peer learning while removing the stress typically associated with team-based learning.

Our contributions in this paper are: (1) to present an online, distributed version of LWT; (2) to provide a comparative study to investigate learning in co-located LWT, distributed LWT, and students learning individually; (3) to elicit student preferences to determine whether the latent, non-performance-specific benefits of social learning - such as engagement, enjoyment and relationship building - translate to an online environment; (4) and finally, to provide a discussion that highlights the challenges of moving LWT online and describes some unexpected benefits.

2. RELATED WORK
2.1 Lightweight Work
We previously define Lightweight Teams (LWT) as small teams in which students collaborate throughout the semester on tasks that do not represent a significant portion of their final grade [15]. In our initial implementation, teams en-
gaged in various activities such as peer-instruction clicker quizzes, structured active-learning challenges, and sharing course notes. Each of these were specifically designed to promote team cohesion and encourage collaborative peer learning. Students sit together in class to develop cohesion and eliminate any potential anxiety experienced by freshmen students entering large classes, who might wonder who they should sit with, whether anyone will want to sit with them or talk to them, etc.

An important activity for LWT was the clicker quizzes, which followed a modified peer instruction method [6]. After each question was projected on screen, students discussed the question with their teammates and then answered individually, without necessarily forming a consensus. In our initial implementation and in this study, clicker quizzes accounted for only 10% of each student’s grade, which was enough to ensure participation, but not enough to cause stress associated with having unprepared teammates.

Previously, we observed positive qualitative feedback from the majority of students that used LWT, with students reporting that they liked being on their teams. In addition, 62% of student respondents noted that they had made 5 or more new friends as a result of taking the class. Students also performed better in a follow-on programming course than students who took a traditional lecture-based course instead. These results demonstrate that LWT create a positive and engaging learning environment. Furthermore, this approach mimics real world environments where graduates must be able to work with others.

Approaches like these provide scaffolding for team-work and for self-regulated learning where students can observe from their peers’ study habits and correct their own as necessary. These opportunities for observation and imitation lead to improvements in self-regulation [22].

2.2 Online Learning

Increasingly, employers expect employees to adjust their skills to meet changing markets and demands. As a result, employees must be lifelong learners who are capable of self-guided learning. For many working adults, the need for continuing education is met through online learning environments, though this is often a frustrating and lonely experience. In addition, a large population of the developing world, connected via the Internet, are reaching out for online educational opportunities that are not readily available locally [7, 3].

Despite surging popularity and the promise of continuing and global education, online learners face many unique challenges. Many of the original challenges described in the canonical work ‘Distance Matters’ [21] - such as establishing common ground, coupling in work, collaboration readiness, and technology readiness - are still issues today. Some challenges associated with working remotely are not limited to remote contexts. For example, establishing trust is challenging in both a remote [2] and face-to-face environment [19].

Some challenges, such as cultural differences, are more unique to remote collaboration. These cultural challenges remind us that distance is both physical and cultural [23]. Some of these cultural aspects of distance are both positive and negative - distance provides more perspectives; however, managing conflicts that naturally arise from multiple perspectives can be challenging to resolve [9, 12].

There are also some unique benefits afforded by distance and computer-mediated communication (CMC) in educational settings. For example, students’ inhibitions about participating in discussions can be alleviated in online contexts [10]. Similarly, online worlds like Second Life provide contextualization, recreating historically accurate environments that no longer exist in the real world [25].

Finally, distance implies geographic diversity which can naturally deter homophily, the tendency people have to associate with people similar to themselves [18]. And peer learning, which is an effective way of developing positive race relationships [24], is well situated to take advantage of this geographic diversity; especially for MOOCs, where many cultures are potentially interacting [7, 3].

Recent research has aimed to provide a social learning environment online via synchronous and asynchronous discussions. Synchronous techniques include textual chat and video conferencing. Some recent examples are a series of studies that investigate learning in discussion forums [4, 5], how new students benefit from interactions with alumni that continue to participate in an IRC chatroom for a course [20], and the online video-chat discussion system, Talkabout [13].

Previously, we investigated an online, modified version of LWT. This modified version, “Low-Stakes Teams” [16], used the peer-instruction and low-stress aspects of LWT but ignored the aspects of temporality and team cohesion. This one-off controlled experiment provided insights about instantaneous learning gains, but didn’t address latent benefits - such as engagement, enjoyment, preference - which require longitudinal studies where teams develop strong cohesion.

<table>
<thead>
<tr>
<th>Team</th>
<th>Week1</th>
<th>Week2</th>
<th>Week3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teams 1-3</td>
<td>Face-to-Face</td>
<td>Distributed</td>
<td>Individual</td>
</tr>
<tr>
<td>Teams 4-6</td>
<td>Distributed</td>
<td>Individual</td>
<td>Face-to-Face</td>
</tr>
<tr>
<td>Teams 7-9</td>
<td>Individual</td>
<td>Face-to-Face</td>
<td>Distributed</td>
</tr>
</tbody>
</table>

3. COMPARATIVE STUDY

For two years, we have been using LWT in our CS1/2 courses. Students in teams of 5 sit together each week and work on clicker quizzes and other activities together during class time. We were interested in extending this approach to support students socially when learning online.

To investigate the effect of adding distance to LWT, we conducted a three week experiment that allowed students to learn in three environments: individually in a traditional lecture format, as co-located LWT, and finally as distributed LWT collaborating remotely. Each participant rotated conditions each week so that they experienced every condition (see Table 1). This resulted in a partially counter-balanced design which accounted for some of the ordering effects and variations in the difficulty of the topics from week to week.

As an in-the-wild study, we designed our experiment with as few changes as possible to the normal course flow.

By using the same teams that had been working together co-located in class through the semester, we were able to mimic a long-term team dynamic in our study, preserving team cohesion. Role negotiation was informal and students
were only guided by the clicker questions. Each week of the experiment consisted of a quiz where students had an opportunity to engage with material and a post test to gauge how their interaction with the material transferred to isomorphic questions (questions that have the same structure or form, but differ in actual content). Each week we booked two classrooms: (1) the normal classroom where students engaged in co-located LWT, (2) a classroom for the "lecture" where students worked alone. In addition to these two rooms, some students worked remotely using Google Hangouts. These differences are described in more detail in the next section. Generally, the quizzes took 45 minutes and the distributed team typically took at least 15 minutes more than other conditions. It is also important to note that in our comparative study, discussion with the instructor is replaced by an explanation slide after each question to control for multiple instructors engaging differently with the groups.

We provided a few tutoring sessions on Google Hangouts leading up to the experiment weeks to get students familiar with using the platform. The tutoring sessions were in the form of review sessions before each test and provided opportunities to promote technology readiness [21]. Approximately 10 students attended each session which served as an opportunity for working out the kinks of using this kind of communication. The week before our first experiment all students were asked to join a Google Hangout as a team and invite the TA to their session to collect attendance. This ensured that all students tested Google Hangouts.

We used a counter-balanced design with three conditions:

**Individual (I):** Students were in a lecture hall where they took a clicker quiz individually. Students were not able to discuss the questions. This approach mimics the traditional lecture. After each question, the correct answer was shown on the screen, along with an explanation. The students took the post-test individually, online or on paper.

**Group Co-located (GC):** Students were in their normal classroom engaging in LWT to answer each clicker question. After discussing each question, they answered individually, without necessarily having reached consensus. Unlike LWT, the instructor didn’t discuss the answers with the class. Instead, the answer was shown after each question and students could discuss the answer. Students took the post-test online, individually.

**Group Distributed (GD):** Students in this condition took the clicker quiz online using Socrative instead of the clicker software used in the other two conditions. Team members discussed clicker questions using audio and video on Google Hangouts. In each Hangout, a TA was present to observe team behavior. The TA was also responsible for pacing the clicker quiz. After all team members answered the question, the TA displayed the explanation followed by the next question. Students in this group left the Google Hangout session and answered the post-test questions online, individually.

Our general research question is whether LWT can provide educational benefit, in terms of learning outcomes and social engagement, to students who are learning at a distance. We had the following hypotheses:

1. Learning gains will be highest for the students in the co-located group.
2. Learning gains will be higher for students in the distributed group than for students working alone.
3. Team conditions will be perceived as being more enjoyable, more rewarding, and requiring less effort than learning alone.
4. Team conditions will be preferred to learning alone.

### 3.2 Participants

Our experiment consisted of 58 student participants in their normal scheduled class period for three weeks. After removing incomplete cases, we analyzed the results for 52 students. Those 52 students contributed 156 observations for each condition over the course of three weeks, where each observation represents a normalized test score. These observations were counterbalanced so that each participant contributed only one observation each week.

### 3.3 Statistical Analysis

Our experiment investigates students’ performance as measured at three points in time. These points correspond to a different week covering a specific topic and a different condition as each participant rotates through all of the conditions. For example, in the first week students learned about enumerated types and in the second week they learned about threads. It is unclear whether these topics are equivalent in terms of difficulty. As a result, our analysis needed to model participant, team and content-related error terms.

Given these considerations, we analyzed the data using multilevel mixed effects models; fit by restricted maximum likelihood estimation. Condition was a fixed effect and both time and students nested within their groups were modeled as random effects. This and the counterbalanced design mitigate some effects of multiple instruments measuring the same construct. When significant, comparisons between groups were made using Tukey contrasts applied to the model. These results are reported using normally distributed Z values. To control family-wise error rates, p-values were adjusted using the single-step method. We checked assumptions of normality. Homoscedasticity of standardized residuals and model fitness were checked using residual-to-predicted plots and Q-Q plots (quantile-quantile).

| Comparison          | Difference | P (> |z|)  |
|---------------------|------------|-------|
| Co-located - Distributed | -0.312     | 0.722 |
| Co-located - Individual | 0.771      | 0.160 |
| Distributed - Individual | 1.084      | 0.018 |

### 3.4 Quantitative Results

We modeled both the clicker quiz and the post-test scores using a multilevel mixed effects model. Assumptions were upheld, but we did observe a slight bias for both models when plotting fitted against residuals. Only the model for clicker quizzes showed significant differences and those differences were explored using Tukey contrasts, shown in Table 2. Generally, teams did better than individuals on the
clicker quizzes; although, this was only significant for distributed LWTs. This difference didn’t appear to translate to post-test scores. High performance during collaboration which doesn’t translate to post-test scores partially supports our previous work with Low-Stakes teams [16].

There are a number of explanations for this: (1) explanations were provided after each question (2) the post-test was too short to delineate differences in understanding (3) isomorphic questions may not have been challenging enough to distinguish between partial and holistic understanding. It is unclear whether LWT provide benefits that apply immediately to transfer questions. Our results also may not reflect the learning gains that could be achieved through long-term sustained engagement and effort.

In our analysis, Week 1 clicker data was not used. In Week 1, both team conditions outperformed individuals; however, distributed teams were paced by all teams instead of being individually paced. Consequently, they took nearly twice as long to finish. This was fixed in Weeks 2 and 3.

4. SUMMATIVE ATTITUDINAL SURVEY

To elicit student preferences about each condition we created a summative attitudinal survey with 11 questions. The first 9 questions were Likert scale questions that asked students about the perceived effort, rewards, and enjoyment associated with each condition. Following these 9 questions, we asked them to rank the conditions by overall preference and then finally we provided them with an open-ended text box where they could comment on any aspect of the study. The preference survey was provided after finals to ensure that students didn’t feel obligated to complete it and to make sure it didn’t interfere with their preparation for their finals. As a result, only 16 students completed this survey.

4.1 Statistical Analysis

We analyze both the rankings of conditions and attitudes about each condition based on a 5-point Likert scale. Ranks and Likert data are ordinal and consequently we used non-parametric tests to analyze them. We used the Friedman test because each participant provided ranks for all conditions. To analyze significant differences, we performed a post-hoc analysis using Fisher’s LSD test for multiple comparisons. We adjusted p-values using the Bonferroni method to control family-wise error rates. The Bonferroni method is typically overly conservative in its adjustments but this adjustment is safest given our small sample size.

Table 3: Rankings between conditions. A low rank indicates that it was selected as best condition more.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Rank-Sum</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-located - Individual</td>
<td>-25</td>
<td>0.002</td>
</tr>
<tr>
<td>Co-located - Distributed</td>
<td>-14</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Distributed - Individual</td>
<td>-11</td>
<td>0.015</td>
</tr>
</tbody>
</table>

4.2 Quantitative Results

We investigated two aspects of our survey data: (1) overall rankings for each condition (2) Likert data related to perceived effort, enjoyment, and rewards.

<table>
<thead>
<tr>
<th>Preference</th>
<th>Comparison</th>
<th>Rank-Sum Difference</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort</td>
<td>Col - Ind</td>
<td>-9.0</td>
<td>&gt; 0.99</td>
</tr>
<tr>
<td></td>
<td>Col - Distr</td>
<td>-22.0</td>
<td>&gt; 0.99</td>
</tr>
<tr>
<td></td>
<td>Distr - Ind</td>
<td>35.5</td>
<td>&gt; 0.99</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>Col - Ind</td>
<td>86.0</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Col - Distr</td>
<td>39.5</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>Distr - Ind</td>
<td>46.5</td>
<td>0.003</td>
</tr>
<tr>
<td>Rewards</td>
<td>Col - Ind</td>
<td>56.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Col - Distr</td>
<td>27.5</td>
<td>0.655</td>
</tr>
<tr>
<td></td>
<td>Distr - Ind</td>
<td>29.0</td>
<td>0.464</td>
</tr>
</tbody>
</table>

The Friedman test found a significant difference between conditions in their rankings of overall preference between conditions (\(\chi^2 = 19.63, p < 0.001\)). Post-hoc analyses also revealed significant differences for all pairwise comparisons. Participants ranked co-located teams as their most preferred condition and working individually as their least favorite condition (see Table 3).

To investigate people’s Likert responses we conducted a Friedman test and found a significant difference in terms of preference (\(\chi^2 = 53.72, p < 0.001\)). Post-hoc analyses also revealed significant differences for some pairwise comparisons. All the comparisons of interest are shown in Table 4.

With a limited sample size and a conservative p-value adjustment, we still obtained significant findings. Specifically, people strongly prefer being face-to-face and are ambivalent about working in distributed teams; however, they generally don’t like to work alone. This evidence supports our fourth hypothesis and is compelling because it indicates that students may be more willing to engage in these team-based alternatives than traditional lectures which supports previous findings [5, 13].

Revisiting our other hypotheses, team conditions were significantly more enjoyable than working alone, shown in Table 4. Our hypothesis that the team conditions would be more rewarding was only significant for face-to-face teams and not for the distributed version, as shown in Table 4. This may relate to the fact that distributed teams typically took at least 15 minutes longer than either of the other conditions. Finally, we did not observe any significant differences between conditions for perceived effort. This was surprising; however, it may be that the reduced cognitive load due to multiple people collaboratively answering a question is offset by the effort of communicating.

5. QUALITATIVE RESULTS

The TAs used a rubric to code behaviors, team dynamics and technical difficulties. These rubrics guided observations and served as common ground for analysis. Many students mentioned that they liked Google Hangouts because they could work from home. Students found new common ground; for example, students discussed hobbies based on cues they observed in each other’s environments, such as a motorcycle poster in one student’s room and a guitar in another student’s room. Figure 1 shows two examples where
students met up for food based on a discussion that started in Google Hangouts and continued in Google Chat. Students continued to use Google Chat throughout the semester to remind each other about deadlines and to meet up to study.

In addition to these advantages, there were some challenges. In one distributed team, a student was watching TV and continued to watch TV even after his teammates mentioned it. He only lowered the volume. Online it is more difficult to hold teammates accountable for their actions since many actions are not as overt as this one. A second challenge is related to time. Students experienced technical difficulties or arrived late and the teams were disrupted as a result. In one case, a student’s laptop battery died and her team waited for her to charge it. One student even participated in the Hangout while driving to campus.

It is important to consider the implications of these observations for future implementations of distributed LWT. Features like supporting teams on mobile devices may or may not be preferable. Similarly, social accountability, which is a key component to successful collaborative work [11], becomes challenging when students can multitask discreetly.

6. DISCUSSION

Our findings suggest that the successes associated with LWT may be primarily associated with the more frequent social interactions occurring between students and increased engagement over the course of the semester. In our study, specific short-term performance benefits were not seen when students were asked to answer isomorphic transfer questions. The team conditions generally didn’t do any better than the students working alone; however, all conditions received the same explanations for each question. This suggests that students working alone with explanations and supplemental material may be able to perform similarly to teams on similar questions. On the other hand, students drastically preferred working in distributed LWT than working alone in a traditional lecture format. As a result, students may be more willing to interact with the material for longer duration. Although distributed teams took the longest of the three conditions and despite technical glitches, students still preferred learning in distributed teams to working alone.

We encountered numerous challenges in moving LWT online. Even though we provided tutoring sessions using CMC to ensure technology readiness, students still experienced moderate technical difficulties, such as battery life, lack of headphones in public places, latency, and microphone problems. Despite an effort to ensure common ground, students weren’t always respectful of each other’s time: arriving late, and multi-tasking while participating.

We think that these study results make an extremely strong case for the importance of studying face-to-face educational methods in an online context before actually deploying them online at large scale. Even with design considerations that specifically accounted for distance, we still experienced many challenges associated with distance. Furthermore, design decisions like using small groups to increase accountability may be problematic if students don’t respect each other’s time and skip a scheduled discussion.

Finally, we think there are some specific changes that we could investigate in the future to improve the efficacy of distributed LWT. First, we think that the group size is very important for online discussions. Specifically, it is much easier for students to become distracted online and there are less overt signs that they have become distracted. We think that by limiting the size of the groups students might be accountable to each other and less able to discreetly multitask. We also found that question pacing is a double edged sword. In the first week all distributed teams were paced by one quiz and as a result some teams had significant down-time between questions. Different teams experienced down-time but generally we noticed that this “positive down-time” may have been beneficial. Students discussed weekend plans and asked about each other’s interests. That week, distributed teams took nearly twice as long as the other conditions.

7. LIMITATIONS

Our study had some limitations, mainly due to the in-the-wild nature of the setup. We integrated this study into a class where we were teaching real students. We considered their learning goals most important and our study secondary. As a result, we were forced to make some concessions in our study, namely: (1) groups were not modified in any way, (2) explanations were provided after each question to ensure students could learn the material, (3) post-tests were short so students weren’t significantly taxed by additional work, (4) post-tests were optional and ungraded, (5) the instructor didn’t interact with students during the clicker quiz (as they traditionally would), to eliminate differences in teaching style and teaching ability, and (6) students in this course typically learn in face-to-face LWT so they may have been more comfortable with this modality.

8. CONCLUSIONS

We presented an in-the-wild comparative study of LWT where students learned individually, face-to-face and online. Our distributed version of LWT was similar in terms of performance to the original implementation of LWT; however, in our study students were unable to interact with the instructor. Our results provide insight about the potential for distributed LWT. In terms of performance, we only saw a significant difference on the clicker quizzes, which didn’t translate to the post-test. However, we did show strong evidence that students prefer working in LWT and enjoyed working in them. Our main qualitative findings convey the
difficulties of learning remotely. As we show in this paper, the challenges that distance creates are difficult to address. Given that distributed and co-located LWT performed similarly and that students strongly preferred these environments to learning individually, distributed LWT show promise as a social, collaborative learning approach that may alleviate some of the loneliness associated with distributed learning and may help to ameliorate low retention rates in online classes.

9. ACKNOWLEDGEMENTS

This work was partially supported by the U.S. Department of Education under grant P200A130088. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsor.

10. REFERENCES


