Student Stress and Learning Outcomes: Effect of Deploying Single-Item Student Check-in on Student Learning Outcomes

Sage Kittelman, Ph.D., Montana State University
Ganapathy Natarajan, Ph.D., University of Wisconsin-Platteville
Ean H. Ng, Ph.D., Oregon State University
Javier Calvo-Amodio, Ph.D., Oregon State University

Abstract
Research has shown that students' mental wellness and emotional state affect student learning. In a prior study, a Check-In Tool$^©$ was developed for the industry shop floor to assess the mental wellness of the shop floor employees before the start of their shift. The study provided evidence that the tool improved communication, and as a result, workers were able to address performance issues at the beginning of their shifts more effectively. The objective of this study is to evaluate the effect of using a single-item questionnaire to check-in with students on their emotional well-being at the beginning of every class and on the student learning outcome.

The Check-In Tool$^©$ was adapted from the industry shop floor purpose to that of academia to check in with the students at the beginning of every class meeting. To assess their semester-long learning outcomes, the learning outcomes of the semester where check-ins were conducted were compared with the learning outcomes from the previous year's cohort. The result shows a significant positive difference in student learning outcomes for the cohort of students where the Check-In Tool$^©$ was deployed. Further study is needed to assess the validity of the single-item questionnaire and to assess whether the significant improvement found in this study was due to the deployment of the Check-In Tool$^©$ or due to a Hawthorne effect.

Background

Research has shown that some level of stress is necessary for effective learning (a.k.a. eustress), however, too much stress leads to distress and can negatively affect student learning [1]. In addition to seeking help from counseling for students to manage stress, research has shown that short and simple interventions conducted by instructors during class time can improve student's mental wellness which in turn can lead to better learning. Examples of simple interventions during class time include deep breathing exercise [2] and short physical activity [3].

This study adapted a Check-In Tool$^©$ previously developed for an industry partner to assess workers' mental state at the beginning of each shift. As the work for the industry partner was proprietary, we are not able to share the results from the study. The premise behind the shop-floor Check-In Tool$^©$ was that if employees and their supervisors who work closely (in close vicinity or closely related job function) are aware of the mental state of their fellow co-worker, it can improve the overall job outcomes for all employees. From the industry study, a correlation between workers knowing the mental state of co-workers and the improvement in the job outcomes was observed.

For this study, the Check-In Tool$^©$ was deployed at the start of each class session over the course of the semester and data collected to assess student engagement and wellbeing. With the data collected, our objective is to identify if the change in their self-reported mental wellness is correlated with the change in student learning outcome. Our premise is that if individual students began to show a decline in their satisfaction with respect to school, it would provide sufficient information for an instructor to intervene and provide the students with additional support or refer them to other resources early on, rather than wait until the student failed course assignments or exams. Additionally, if the majority of cohort of students began to show a decline in their satisfaction with respect to school, it might provide sufficient information for the instructor to evaluate their own teaching pedagogy and assess, for instance, whether more time needed to be spent on specific content before the students were prepared to move on to the next topic. The early detection and evidence for intervention would provide fast and frequent feedback for assessing the current state of the system and adjusting it quickly, rather than after divergent behavior has progressed to a point where significant intervention or student failure was the likely result (see Figure 1). The theoretical support for the framework presented in Figure 1 is detailed in [4].
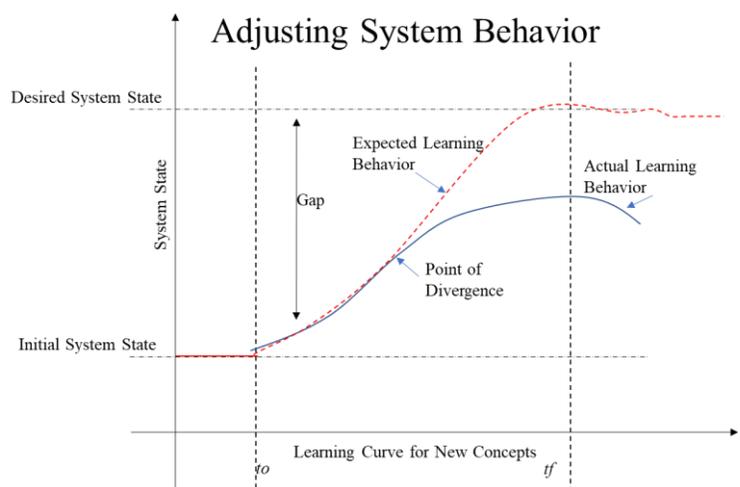

Figure 1. Early Detection of Divergent Behavior

Kittelman [4] conducted an exploratory study at Montana State University during Spring semester 2023 in a senior engineering capstone design course to determine if the students would check-in on a consistent basis, whether they would answer honestly, and whether the tool would provide useful information for the general trends of student wellbeing. The exploratory study results show that when provided, students will consistently use the Check-In Tool$^©$, the integration of this Check-In Tool$^©$ with classroom teaching was seamless, and that useful information regarding group and individual stresses could be collected and measured [4]. Kittelman [4] study did not evaluate the relationship between the frequent student check-in and the student course learning outcomes.

Methodology
This study tracks the same participants in a single senior level course over the a single 15-week semester with the goal of observing changes in student self-assessments on a class-to-class basis and evaluate their learning outcomes. The same course offered in the Spring 2022 semester with different cohort of students was used as a control group, where the Check-In Tool© was not deployed in the classroom during the Spring 2022 offering.

Check-In Tool$^©$ Design
The Check-In Tool$^©$ was employed using the online platform Mentimeter.com, an interactive application designed for collecting audience feedback. Mentimeter.com is an application commonly used in classrooms, focus groups, and professional presentations where collecting feedback or engaging participants is the goal. However, unlike the typical application of Mentimeter.com, where participants can see the result of the group responses, a QR code was displayed for students to access the check-in tool and the results remained anonymous after they submitted their responses to avoid group bias or shy responders. During the capstone course in Spring 2023, this QR code was displayed via the projector at the start of every class session to develop the habit in students of answering the survey first thing when arriving to class.

The QR code would take students to the Mentimeter.com poll with two prompts, 1) asking the students to type their first name and first initial of last name for attendance records and 2) display the question "How are you doing today?" Students were informed at the start of the semester that the first prompt was mandatory for attendance records, however, answering the second prompt was completely voluntary, and by answering the second prompt, student is consenting to participating in the research study. While attendance in capstone was mandatory for the success of the student teams, this attendance requirement was designed so that reporting may be increased by students already being logged on to Mentimeter.com to answer the second question versus not taking the effort to log on at all if attendance was not required or tracked via another method.

Figure 2 shows both the student (audience) view as well as the results (presenter/instructor) view of Mentimeter.com. Students answered the question "How are you doing today?" on a sliding scale from 0 to 10 for both quadrants "Could be better" or "Everything is great!" as well as "School related" or "Life related".
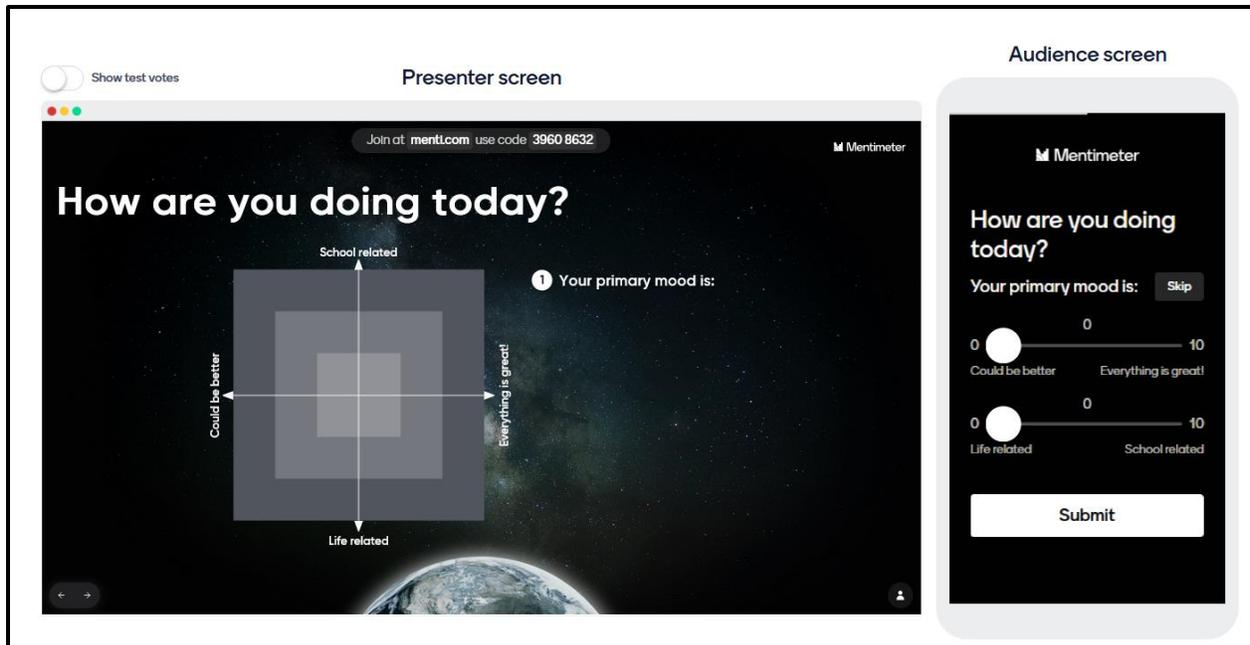
Figure 2. Mentimeter.com Survey Screens

This sliding scale design from 0 to 10, with a left justified 0 being the prompt initial default position (rather than a neutral center position of 5), encouraged more consideration from students to rate how they were actually feeling rather than students simply leaving the default prompt position for the sliding scale at a neutral center position. Additionally, the sub-prompt "your primary mood is:" focuses students on their *dominant* feeling versus two separate measurements of "Could be better" or "Everything is great!" for both domains "Life" and "School". The justification for this design is for the tool to mimic a rather typical engagement prompt between people greeting each other versus more of a survey with multiple prompts. Responders typically address the question verbally by stating they are "fine" or going into more detail on the dominant feeling they are experiencing and the cause of that feeling. From the moment students scanned the QR code to the completion of the Mentimeter.com prompts, the approximate time for students to check-in each class was 2-4 seconds.

Student Response to Check-In Tool©
The response from student will fall within a quadrant on the Check-In Tool©. Students were asked to indicate how they feel first, follow by their initial feeling is related to school or life. Student's response that falls within Quadrant II as shown in Figure 3 represents student feeling stress related to school work; whereas response within Quadrant IV represents students not feeling stress related to life. Responses falling within Quadrants I and II show that student is preoccupied with school-related activities, and it can be positively or negatively affecting them; whereas responses falling within Quadrants III and IV show that they are preoccupied with life-related activities. Life-related activities could include family commitment or job-search related concerns. The image in Figure 3 shows the plot of all students that check-in for one class. The average of all check-in responses, represented with "1" in the image shows the class average. In this case, it shows that students are mainly preoccupied with life-related stress, and the stress level is average for the cohort, and that, on average, the students are generally positive that day.
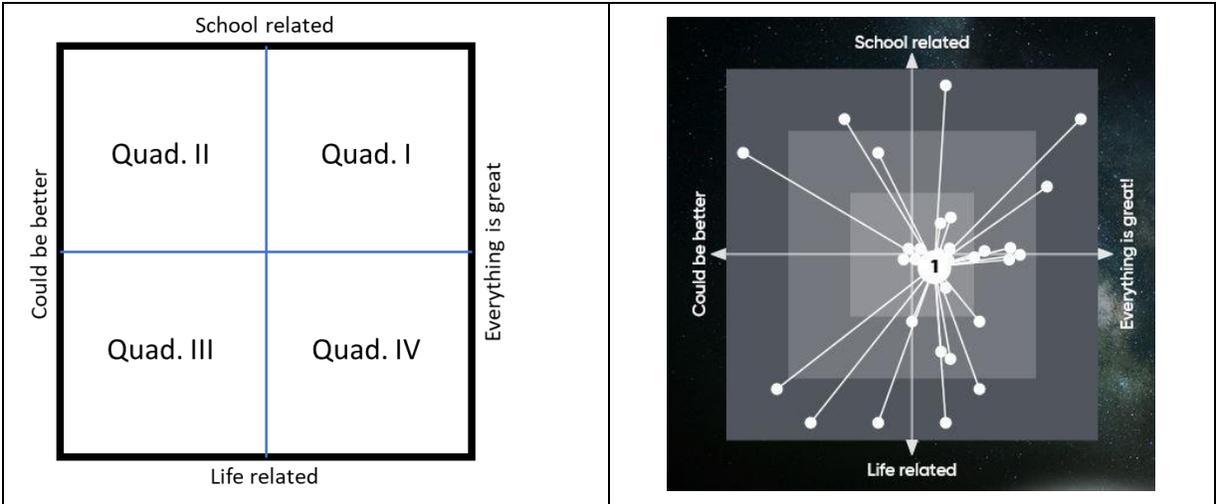
Figure 3. Plot of students' responses to the Check-In Tool©

Course Setup and Learning Outcomes Assessment
This study was conducted in a senior-level engineering capstone design course, which is a required course to be completed during the semester when they graduate. The course spanned 15 weeks, students are divided into teams of three to four students. Each team is assigned a real-world industry project where students are required to interact with the stakeholders (industry project sponsor) to define the project scope, design a solution, assess the feasibility of the solution, and in some instances, implement the solution. In each step of the process, students have to interact with the project sponsor as well as the course instructor to receive input and approval. The course learning outcomes assessment is conducted across seven sprints, where students have deliverables that are assessed for grades for each sprint. Each of the course learning outcome assessments bears a different weight proportion to the final grades for the course. The class met three times a week for 15-weeks from mid-January to mid-May.

Study Participants
There were 40 registered students in the capstone course, 28 Industrial Engineering and 12 Financial Engineering, all are seniors graduating after Spring or Fall semester. This demographic was chosen for two reasons, 1) their proximity to graduation and experiencing the additional stressors of applying for jobs, completing courses, and/or relocating, and 2) that capstone is a course that requires students to work in teams to solve real-world industry problem, adding additional stress through interpersonal cooperation of all team members and striving for a successful project completion.

Results
Control vs. Treatment Analysis: Spring 2022 was considered the control semester and Spring 2023, when the Mentimeter check-in was used was considered the treatment semester.

Data Processing
Following an agile management approach, the Capstone course was divided into seven sprints. Data were collected for the Spring 2022 and Spring 2023 semesters across the seven sprints. The total maximum score possible for each sprint varied. In which case, all the scores were normalized by dividing the raw student scores by the total possible scores. Upon further investigation, it was

noticed that the Sprint 1 scores were 1 for all students given that all students received full points if the work was completed on time. Since this would not provide any additional useful information, this Sprint was dropped from further analysis.

Descriptive Statistics
The descriptive statistics of the data collected is presented in Table 1.

Table 1. Descriptive Statistics of Control vs Treatment Group

Semester	Sprint	Mean	SD	Min	25%	50%	75%	Max
Spring 2022 (N = 25)	**Sprint 2**	0.8348	0.1042	0.5700	0.8200	0.8700	0.8880	0.9120
	Sprint 3	0.8499	0.0647	0.7900	0.8080	0.8100	0.9320	0.9500
	Sprint 4	0.8664	0.0486	0.7740	0.8675	0.8675	0.8885	0.9398
	Sprint 5	0.8803	0.0537	0.7910	0.8440	0.8860	0.9270	0.9460
	Sprint 6	0.9167	0.0519	0.8200	0.8920	0.9020	0.9670	0.9770
	Sprint 7	0.9133	0.0487	0.7967	0.9000	0.9333	0.9500	0.9500
Spring 2023 (N = 28)	**Sprint 2**	0.7961	0.1148	0.6400	0.7175	0.8000	0.8800	0.9600
	Sprint 3	0.8113	0.1506	0.5000	0.7350	0.8800	0.9050	0.9550
	Sprint 4	0.9132	0.0624	0.8100	0.8550	0.9300	0.9750	0.9900
	Sprint 5	0.8993	0.0913	0.7550	0.8250	0.9000	0.9913	1.0000
	Sprint 6	0.9168	0.1020	0.7400	0.8550	0.9800	1.0000	1.0000
	Sprint 7	0.9455	0.0609	0.8067	0.8950	0.9833	0.9933	1.0000

Visual Analysis
We expect the intervention of using daily check-ins to alleviate student stress and thus increase the learning outcomes in the treatment semester. Before performing a statistical analysis, a visual test was performed using boxplots, shown in Figure 4.
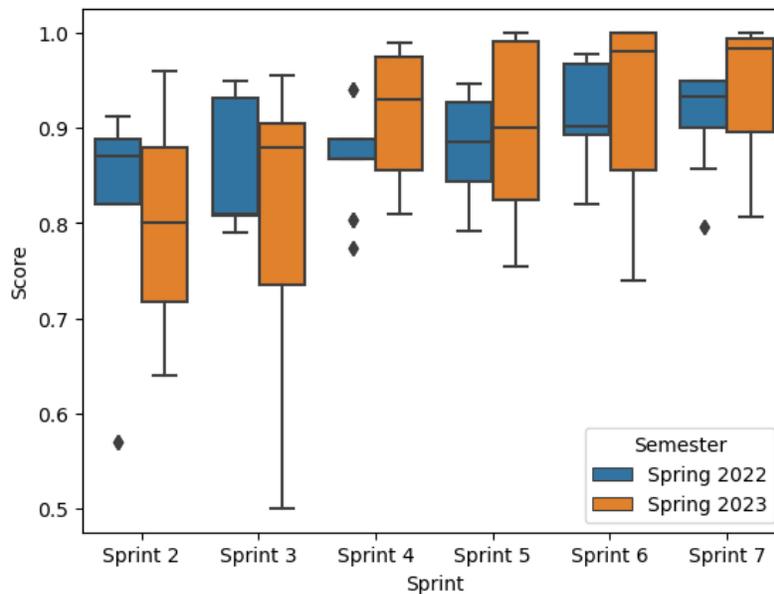

Figure 4: Boxplots of Student Learning Outcomes for Each Sprint between Control and Treatment Group. The learning outcomes score on the y-axis is the normalized score.

In the earlier sprints when students are still getting used to the daily check-in and the structure of the course, the scores in Spring 2023 (treatment) seems to be lower. However, in later sprints (usually the more strenuous or stress-inducing sprints), student scores are higher in the treatment semester than the control semester. Further statistical tests may show if the differences are significant.

Statistical Analysis
A two-way ANOVA was performed with Semester and Sprints as the levels, including the interaction effects of semester x sprint. The results are shown in Table 2.

Table 2. Effect of check-in on learning outcomes

	SS	DF	F	p
Semester	0.0009	1	0.1266	7.22E-01*
Sprint	0.5814	5	15.5476	1.24E-13**
Semester x Sprint	0.0860	5	2.2993	4.50E-02**
Residual	2.2885	306		

* Significant at the 0.1 level; **Significant at the 0.05 level

Semester was statistically significant at the $\alpha = 0.1$ level, but not at the $\alpha = 0.05$ level. Both Sprint and the interaction effects of Semester and Sprint were statistically significant at the $\alpha = 0.05$ level.

Post-hoc Analysis
Following the findings of the statistical analysis a post-hoc analysis was performed to explain the individual mean differences between the different Sprints. The Tukey test results are presented in Table 3.

Table 3. Tukey Test on the learning outcomes between sprints

Group 1	Group 2	Spring 2022 (Control)		Spring 2023 (Treatment)	
		Mean Diff	p	Mean Diff	p
Sprint 2	Sprint 3	0.0151	0.9634	0.0152	0.9935
Sprint 2	Sprint 4	0.0316	0.5224	0.1171	0.0004
Sprint 2	Sprint 5	0.0455	0.1384	0.1032	0.0028*
Sprint 2	Sprint 6	0.0819	0.0002*	0.1207	0.0002*
Sprint 2	Sprint 7	0.0785	0.0005*	0.1495	0.0000*
Sprint 3	Sprint 4	0.0165	0.9465	0.1020	0.0033*
Sprint 3	Sprint 5	0.0304	0.5635	0.0880	0.0182*
Sprint 3	Sprint 6	0.0668	0.0051*	0.1055	0.0021*
Sprint 3	Sprint 7	0.0635	0.0094*	0.1343	0.0000*
Sprint 4	Sprint 5	0.0139	0.9740	-0.0139	0.9957
Sprint 4	Sprint 6	0.0503	0.0746	0.0036	1.0000
Sprint 4	Sprint 7	0.0469	0.1160	0.0323	0.8422
Sprint 5	Sprint 6	0.0364	0.3605	0.0175	0.9875
Sprint 5	Sprint 7	0.0330	0.4720	0.0463	0.5342
Sprint 6	Sprint 7	-0.0033	1.0000	0.0288	0.8977

*Significant difference at the 0.05 level

The post-hoc analysis showed similar trends between the control and treatment groups. The significant differences seem to happen during the middle of the semester. With the treatment group the difference in mean scores between the sprints is very consistent during the middle portion of the semester. The gains realized from the Check-In Tool© seem to diminish as the semester progressed.

Within Treatment Analysis

Comparing the different semesters provided some insights into the usefulness of the Check-In Tool©. A within treatment group analysis based on the Spring 2023 data was also performed. The check-in asked about the general feeling and how the feeling related to life or school. Figure 5 illustrates the average rating per Sprint on both the prompts. The band shows the 95% confidence interval around the mean. Both the values change during the semester and the highest feeling seem to be during the middle of the semester. These two dimensions are not supposed to be looked at separately, so, the data was analyzed using Figure 6.
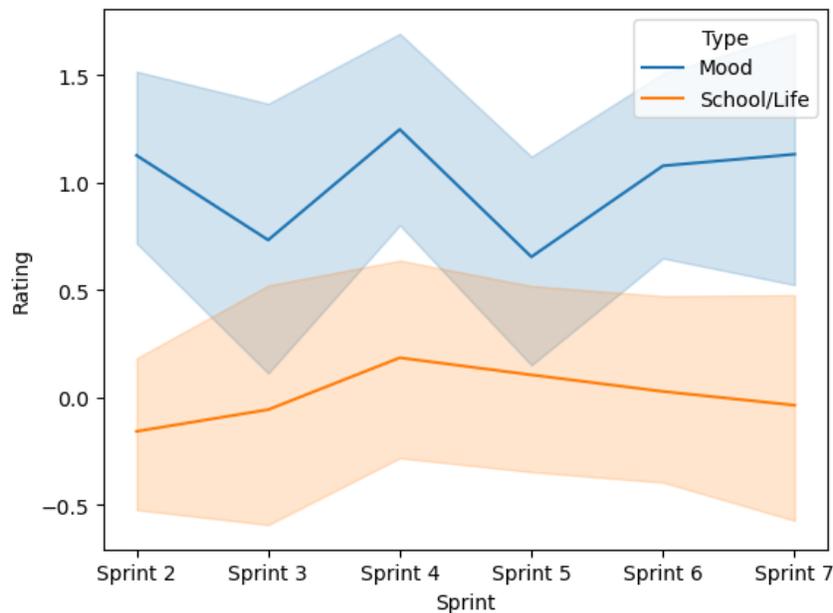

Figure 5. Average rating per sprint.

Figure 6 shows the scatter plot of mean values color coded by the Sprint. The means tend to generally fall on the two quadrants that relate to general positive outlook. Additionally, drilling deeper, it was important to determine if the feeling is life related (quadrant IV, negative y-values) or school related (quadrant I, positive y-values). The feeling is life-related in Sprints 2, 3, and 7, and school-related in Sprints 4, 5, and 6. Sprints 4, 5, and 6 are when the majority of the work related to the projects are completed (i.e. deliverables with higher weight are due), and the Check-In Tool© seem to capture the overall sense of the cohort in this regard. Interestingly, Sprint 7 consists of the final deliverable (report) for the course, which bears higher weight as well, yet, students are preoccupied with life-related stress.
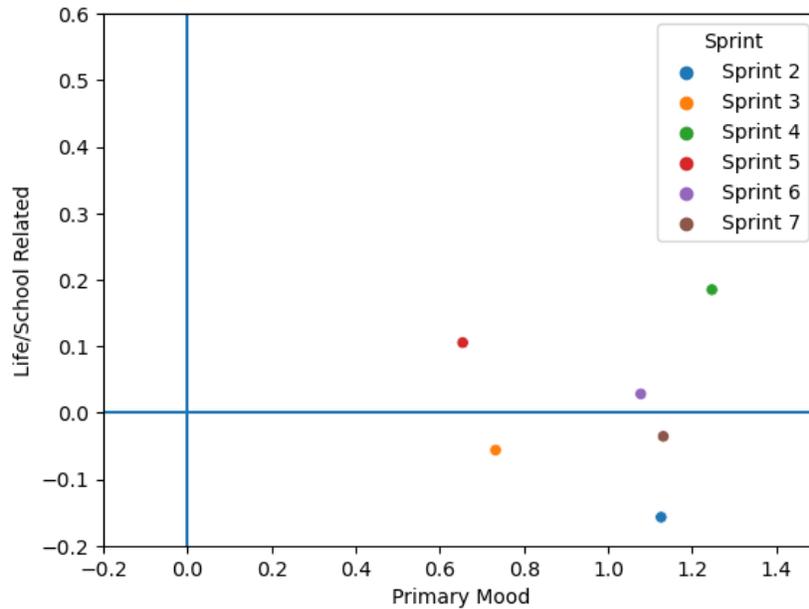
Figure 5. Scatter plot showing students' preoccupation with life versus school in different sprint.

Discussion

On student learning outcomes, the general trend shows that the treatment group's learning outcomes are significantly better than the control group during the sprints where more heavily weighted deliverables are due. The descriptive statistics show that the variability in student learning outcomes increases in the treatment group, along with the median. It can be inferred that the Check-In Tool© implementation does have an impact on the student learning outcomes.

On student's self-reported mental wellness, students are equally preoccupied with school and life-related activities. Depending on the sprint, if the sprint consists of heavily weight deliverables due, students are more preoccupied with school related stress, except for the final sprint. Students' stress level is generally within the range that shows they are slightly stressed, but not at the extreme that warrant intervention from the instructor when considering the average of the cohort results. This does not consider the single student analysis or evidence for interventions as discussed by Kittelman [4].

Another interesting finding is that when we track individual student's check-in response over the semester, a few students' responses over time show that they are experiencing higher than their cohort's stress level. This indicates that intervention from the instructor on an individual basis might be needed. As this is an exploratory study, we did not assess the check-in responses throughout the semester. If we were to deployed this Check-In Tool ©, we need to have a strategy on how to intervene and assist students who are exhibiting higher than cohort's stress level. The development of a strategy for individual intervention will have to consider student's right to privacy and confidentiality, and will require collaboration with the institution's student mental health services unit.

As this is an exploratory study, the validity and reliability of this single-item questionnaire has yet to be assessed. Even though the results show that there is an improvement in student learning

outcomes when comparing the treatment to the control group, one important consideration of this improvement is whether the improvement is due to the Hawthorne effect, where students perceived that the instructor is assessing their stress, and therefore, they unconsciously reacted to this check-in through a change in their performance. As for the reliability of this tool, a Cronbach alpha analysis was not conducted due to the limited sample.

Two other limitations of this exploratory study were the consideration of project types and team member influence on student responses. These questions were not asked directly though both variables are assumed to have an impact on how the students reported regarding the "Good/Bad" and "School" quadrants. For instance, regarding project-specific stressors, some teams experienced rough beginnings due to a lack of data available from the sponsors, other teams experienced stakeholders that changed their requirements mid-project, and still, other teams were met at the end of the course with either a successful project completion or just an implementation plan to handoff. Each experience may have contributed very directly to how individual students reported with the Check-In Tool©. However, this was not measured directly.

Similar to project experiences, team member influences were also expected to have a direct impact on individual student responses. Teams that experienced a lack of cohesion, effective communication, or clarity and responsibility of tasks between members suffered more in their ability to complete a successful and satisfying project. While these teams and team members were visibly stressed at times and often conveyed frustration, data was not specifically collected or analyzed to evaluate if and how much team members contribute to individual student stressors.

Neither the project stressors nor team dynamics were generally unique between the years before or during the collection of student feedback via the check-in. Both years experienced stressors of project and/or team member dynamics as is typical with the storming phase [5] of project management. However, future work would benefit from asking a follow-up question about why students rated themselves in the "Bad" and "School" quadrants and performing a qualitative analysis on the responses.

More fundamentally, the tool is not designed to influence the student(s) directly, but rather it gives the instructor information with which to modulate the education "system". Changes in the system, such as individual meetings, suggested reading material, or other modulations affect the student's performance. This tool gives the teacher increased ability to quickly adapt instruction to the needs of the students. Therefore, whatever the external variables bring, a more attentive and responsive instructor is far more likely to increase student performance.

In addition, variables from outside the educational confines (differences in the projects themselves, individual student families, and more) tend to normalize in the aggregate. In this case, there were 10 different capstone projects and 40 students for the treatment group. Net change in the aggregate performance is a good indicator of the effectiveness of the tool overall.

Future Work
To parse out if this change in performance is due to Hawthorne effect, a further study is planned that will require the course instructor to explicitly address the check-in results in class (1) for several sprints as treatment for some courses to assess the within class difference, and (2) for the

entire course as treatment, and another course with check-in but without addressing the check-in results for control.

For the questions, "How are you doing today?" and "Is this school or life-related", a future collaboration is planned with a researcher from counseling to assess if the questions asked and the language used is the most appropriate to capture the general feeling of stress among the students.

References
[1] J. R. Rudland, C. Golding and T. j. Wilkinson, "The stress paradox: how stress can be good for learning," Medical education, vol. 54, no. 1, pp. 40-45, 2020.
[2] A. Robertson, E. Mason, V. Placeres and H. Carter, "Investigating a deep breathing intervention to promote mental health in P-12 schools.," Psychology in the Schools, vol. 60, pp. 3056-3072, 2023.
[3] C. Müller, D. Dubiel, E. Kremeti, M. Lieb, E. Streicher, N. S. Oglou, C. Mickel and J. Karbach, "Effects of a Single Physical or Mindfulness Intervention on Mood, Attention, and Executive Functions: Results from two Randomized Controlled Studies in University Classes," Mindfulness, vol. 12, pp. 1282-1293, 2021.
[4] S. Kittelman, "Engineering Responsive Systems with the Fast and Frequent Feedback Methodology," in American Society for Engineering Management International Annual Conference Proceedings, Denver, Colorado, 2023.
[5] Tuckman, B. W., "Developmental sequence in small groups". Psychological bulletin, *63*(6), pp. 384, 1965.