Understanding Disruptions to Virtual Learning: Causes of and Variation in Lost Instructional Time

Xander Beberman, Sarah Novicoff, Ana Trindade Ribeiro, Carly Robinson, Susanna Loeb

Click here for latest version

April 2025

Abstract

Learning takes time. Tutoring is one of the most effective educational interventions for accelerating learning, granting students dedicated time to work with an educator on their specific needs. However, a disruptive environment can lead to significant instruction time losses and reduce the effectiveness of tutoring. Quantifying how much learning time is lost with disruptions within a classroom has previously presented a substantial challenge for research due to the laborious nature of classroom observations and human coding. In this paper, we use natural language processing methods to systematically identify and categorize disruptions at scale, and quantify instruction time loss for 21,937 virtual tutoring sessions. We estimate that the usable time for instruction during tutoring sessions was reduced by 18% on average due to disruptions. The most common disruptions identified in our sample were related to problems with the technology used for tutoring, followed by students being disruptive and unresponsive to the tutors. We find significant variation in the amount and types of disruption associated with student characteristics and the different campuses where tutoring was administered. We explore potentially disruptive elements of the school environment and discuss the implications for effective virtual tutoring implementation in schools.

1 Introduction

Virtual tutoring is an increasingly popular solution for school districts that wish to support struggling students with personalized instruction (**?**). However, in-school virtual tutoring programs tend to have smaller effect sizes compared to the in-person model or to out-of-school virtual tutoring. Although many factors may contribute to these differences, this study investigates the extent to which different types of disruption related to the student environment reduce instruction time in virtual tutoring and their relationship with academic achievement.

This discrepancy in the effectiveness of virtual tutoring in and outside school settings could be explained by competing hypotheses: first, out-of-school tutoring takes place among a group of students who are motivated or able to take up the program during those outside-of-school hours, and therefore, the effects of virtual tutoring may not generalize to virtual programming embedded within the school day for a more representative group of students with greater need and/or less motivation. Alternatively, or perhaps in addition, in-school tutoring is subject to the constraints of a school environment, where other students or teachers may interrupt a session directly or indirectly, intentionally or unintentionally.

In schools, external interruptions (Kraft and Monti-Nussbaum, 2021) and within-classroom disruptions have been shown to erode much of the learning time for students. Prior meta-analyses of the effect of time in school have shown that losses in instructional time can reduce achievement and that increases in instructional time can increase student achievement (Yeşil Dağlı, 2019; Gromada and Shewbridge, 2016; Holland et al., 2015; Kraft and Novicoff, 2024; Patall et al., 2010; Robinson and Loeb, 2021). Similarly, the tutoring literature shows that "high dosage" tutoring programs that leverage time more explicitly as an input into learning and offer tutoring three or more times per week are more likely to be effective and result in significant learning gains (Fryer Jr, 2017; Nickow et al., 2024) than programs with less frequent tutoring sessions. Just as disruptions can derail traditional classroom instruction and harm academic achievement, disruptions could reduce the efficacy of virtual tutoring as well, though this has yet to be documented in the literature.

Quantifying how much learning time is lost within a classroom has previously presented a substantial challenge for research. Much of the existing literature has relied on the labor-intensive process of classroom observers quantifying pauses in instruction or the use of highly subjective teacher- or student-surveys. Virtual learning creates the opportunity to record and extract data from multiple dimensions of the learning experience with consistent quality across students. Natural language processing (NLP) methods enable researchers to measure the frequency of disruptions more comprehensively and quantify the associated time loss. For this purpose, we use NLP text classifiers to systematically identify different types of disruption in tutor utterances across thousands of virtual tutoring sessions.

Our findings reveal that, on average, disruptions affect 18% of virtual tutoring time at schools in our data, with significant variation across school sites, group size, and student characteristics. The most disruptive campus site in our sample had an average of 21% session time disrupted, which is 50% more than the least disruptive campus, 13.5%. This difference persists after controlling for student characteristics and group size. We find that students tutored in pairs experience significantly more disruptions than those tutored individually, resulting in a seven percentage points ($p < 0.001$) reduction in usable instruction time.

2 Data

We use a large dataset of audio and video recordings of virtual tutoring sessions originally collected for a randomized controlled trial evaluating the effectiveness of an early literacy tutoring program in an individualized (1:1 student-tutor) or paired (2:1 student-tutor) format. Our data includes over 26,000 sessions that took place between November 2022 and May 2023, of which 19,448 were conducted 1:1 and 7,150 were conducted 2:1. We received separate recordings for the tutor and student(s) present in each session, with metadata that includes student and tutor identifiers and their timestamps for logging in and off of each session. A small subset of sessions are conducted with a higher student-tutor ratio, possibly due to tutor absences. We observe 1,357 students and 192 tutors participating in virtual tutoring during this time, spread across 12 different campuses of the same school system. Tutoring took place with students in kindergarten, first, and second grade, and focused on learning literacy skills.

We received school administrative data to link tutoring sessions with student characteristics and achievement scores from the beginning and end of the year. The administrative data we use includes variables such as student grade, date of birth, race/ethnicity, gender, whether the student received free or reduced-price lunch or was otherwise indicated as economically disadvantaged based on the receipt of other public assistance, whether they had an Individualized Education Plan or 504 Plan, whether they were designated as an English learner, and their availability for tutoring within the school day. It also includes student scores on standardized literacy evaluations Dynamic Indicators of Basic Literacy Skills (DIBELS) University of Oregon (2021) and MAP Reading. We supplement these data with information about the implementation setting in each school, typically in the classroom or in a common space, and school enrollment numbers.

3 Methods

To analyze the impact of disruptions on instruction time loss, we used a RoBERTa-based text classifier for types of disruption typical to virtual tutoring sessions and annotated tutor transcriptions at scale. This classifier allows us to identify when a disruption occurs during a tutoring session and estimate the time spent on each type of disruption based on session transcript timestamps.

In a qualitative exploration of student and tutor audio recordings, we identified significant differences in the quality of the recordings for tutors and students. Tutors are more often connected to tutoring sessions from an environment with almost no background noise, and tutors tend to speak clearly into the microphone. Student audio recordings from the same tutoring sessions reveal a wide range of background noise levels in the student environment. Student speech clarity is also variable, which can be attributed to a combination of student age (younger students tend to articulate less) and interference from background noise from the environment. As a result, Automatic Speech Recognition (ASR) generated transcripts for students contained a high degree of hallucinations and inaccuracies, whereas tutor transcripts presented a high degree of accuracy. This difference led us to focus our analysis solely on tutor transcripts.

3.1 Text Classifier

3.1.1 Disruption Categories

Based on a qualitative analysis of a small sample of tutor transcripts, we defined a disruption annotation framework to distinguish the types of disruptions that can occur in the context of in-school, individualized, or paired virtual tutoring. The categories included in our framework are the following.

- *Background Noise*: The tutor is prevented from delivering instruction because there is too much noise. This includes noise caused by adults or students in the room with the target student. This is often indicated by a tutor asking the student to repeat themselves. Example: "I cannot hear you. It is so loud."

- *Tech Problem*: The tutor is prevented from delivering instruction because the technology underlying the session (e.g., headphones, Internet, video, microphone) is not working as intended. This is sometimes indicated by a tutor who does not think they can be heard or seen, or a tutor repeating themselves. Example: "You can't see me? I'm not sure why you can't see me."

- *Student Disruption*: The tutor could be delivering instruction or building rapport during this time period, but the student is doing something that prevents this from occurring. This includes when the student purposefully uses otherwise functional tech to prevent instruction from occurring. Example: "[Student], there's no writing right now. So just leave your thing. I'm asking a question."

- *Second Student*: The tutor could be delivering instruction or building rapport during this time period, but they are instead asking or commenting on the absence of a second student or greeting that second student when they arrive. Example: "So [Student A], I want to let you know what I told [Student B]."

- *Substitute Tutor*: The tutor could be delivering instruction or building rapport during this time if the regular tutor was present, but the regular tutor is absent and so the substitute tutor has to spend time introducing themselves. Example: "I will be your tutor for today only because your regular tutor is not available today, so maybe she'll be here on Thursday, all right?"

- *Not a disruption*: Tutor is delivering instruction or building rapport during this time. Example: "So for our first activity, we are going to be reading the words on the board."

Disruption category	Count
Not a Disruption	1,948
Tech Problem	239
Student Disruption	140
Background Noise	81
Second Student	63
Other Stop to Usable Time	21
Substitute Tutor	9
Total	2,501

Table 1: Frequency of disruption classes of annotated tutor utterances

3.1.2 Human Annotation

We used Label Studio Tkachenko et al. (2020-2022), an online labeling interface to annotate utterances from tutor transcripts. We randomly selected five utterances from transcripts of 255 tutoring sessions, with the constraint that at least one utterance had to come from the first two minutes of the tutor and student appearing in the session. In the labeling interface, we displayed the selected utterances along with the three utterances before and after to give context. Annotators were able to select one of the above categories for each utterance and also indicate when a line in the context around the utterance showed a different disruption. Disruptions in the context lines were later annotated using the same system. For this submission, we focus on 2,501 labels of 1,488 unique utterances, annotated by 22 annotators. The frequencies of each disruption category annotation are shown in table 1. For these annotations, we observe a Krippendorff's alpha of 0.71 when considering all disruption classes, and of 0.74 when considering only whether or not an utterance was a disruption, which suggests moderate inter-rater reliability.

3.1.3 Training and Performance

Using our transcript annotations as a training set, we trained a multi-class natural language classifier to predict the disruption categories of transcribed tutor utterances. We fine-tuned the RoBERTa model (Liu et al., 2019) for our disruption classification task, using 80% of our 2,501 annotations as the training set. To account for the naturally occurring imbalance in category representation in our corpus, where the majority of utterances is expected to be classified as "Not a Disruption," we oversampled disruptive utterances in our training set by a factor of 3. The remaining 20% of observations were held out for use in a test set. The utterances input into the model were also surrounded by the context utterances, three before and three after, consistent with the protocol used for human annotation.

We used our held out utterances to create a test set of 289 unique utterances with labels agreed on by at least two human annotators. The model we train

Category	Precision	Recall	F1 Score	ROC-AUC
Not a Disruption	0.93	0.86	0.89	0.90
Tech Problem	0.60	0.79	0.68	0.94
Student Disruption	0.50	0.72	0.59	0.92
Background Noise	0.35	0.43	0.39	0.74
Any Disruption	0.67	0.83	0.74	

Table 2: Disruption classifier metrics on test set

achieves an accuracy of of 80% on the test set. When not considering the disruption category, and instead only whether or not the utterance indicated a disruption, the accuracy improves to 85%. Further evaluation metrics for the disruption classifier are presented in table 2; only the four largest categories are shown, as many of the categories were not well supported in the test set due to the small size of the dataset.

3.2 Predictors of Disruption

We estimate the following regression model for each category of disruption as separate dependent variables. Our unit of observation is the session level. We include an indicator for 2:1 tutoring that equals 1 for sessions attributed to a pre-assigned pair and 0 otherwise; a vector of student characteristics indicators including economically disadvantaged status, English learner designation, student with disability designation, recorded race Black, recorded ethnicity Latina/o/x, and recorded gender female, which equals 1 when the characteristic is present for at least one student in the session; school-specific indicators (s); and grade-specific (g) indicators. We use these estimates to explore how variables representing student characteristics and how each school in our sample relates to the different types of disruption. We cluster standard errors at the student-pair assignment to the RCT level to account for sessions from the same student or student-pair not being independent of each other.

$$y_i = \beta_0 + \beta_1(2:1t)_i + \mathbf{X}_i\boldsymbol{\gamma} + \delta_{s(i)} + \lambda_{g(i)} + \epsilon_i$$

We estimate a second group of regressions to investigate the relationship between implementation characteristics and disruptions, also at the session level. We included the following independent variables in addition to student controls: where and at what time of day tutoring took place in the school, number of students being tutored simultaneously at the school, number of students enrolled at the school as a proxy for infrastructure capacity, and the share of simultaneous sessions as a proxy for strain on infrastructure capacity.

Category	Predicted session portion
Not a Disruption	0.820
Tech Problem	0.077
Student Disruption	0.064
Background Noise	0.021
Second Student	0.009
Other Stop to Usable Time	0.006
Substitute Tutor	0.003

Table 3: Prediction portion of session time by disruption category

3.3 [Next steps] Disruption and tutoring effectiveness

Since our data was collected for an RCT that was not designed to estimate the effect of disruption or instructional time on tutoring effectiveness, we will adopt an instrumental variable approach to account for unobserved confounding factors between the disruptions experienced by a student and their achievement scores at the end of the school year.

4 Results

4.1 Predicted Disruption

We used the classifier to predict disruption categories for unannotated sessions. We retrieved the tutor transcripts corresponding to the out-of-sample student audio files described above and passed them as input into the classifier. Using the normalized classification output for each utterance and the time between utterances, we approximated the portion of time spent on disruptions of each category during sessions. The mean results over the unannotated sample are shown in table 3. The estimates from our model show that almost one-fifth of each session is spent dealing with disruptions. When we examine the breakdown of disruption time by category, we see that environmental disruptions, such as tech issues, background noise, and substitute tutors, account for about 10% of usable time. Disruptions caused by students account for another 7%.

4.1.1 Environmental and Behavioral Disruptions

We used our custom text classifier to annotate each tutor utterance and calculated the estimated amount of time dedicated to each disruption category based on utterance timestamps. Across our sample of 21,937 sessions, we estimate an average of 18% of tutoring time is spent on disruptions, with the two largest categories being tech problem (7.7%) and student disruption (6.4%).

Tables 4 and 5 present the regression estimates for the variables representing session conditions and student characteristics. Coefficients in these regressions

are interpreted as percentage point difference in time used for the disruption category of interest. In table 4, column one indicates that time spent on utterances classified as "Not a disruption" are negatively correlated with 2:1 tutoring sessions, with usable instruction time decreasing by 7 percentage points ($p < 0.001$) compared to 1:1 tutoring sessions. On the other hand, having at least one female student in the session is associated with an increase of usable instruction time by 0.9 p.p. ($p < 0.05$). These differences in usable instruction time appear to be explained in great part by time spent on "student disruptions" (column two), which can be attributed to behavioral stereotypes of student characteristics, such as gender and race/ethnicity, and is expected to increase with the number of students.
In table 5, we show the results for disruptions attributed to the tutoring environment and unrelated to student behavior. Column one shows that tech problems during the tutoring session are less likely to occur for second grade students, the oldest group of students in our sample and likely most experienced and able to use the necessary technology correctly. We also find significant variation in how much tech disruption is associated with specific schools. Similarly, in column two, Background noise identified by our text classifier is only associated with specific schools, indicating that in this sample of sessions, certain schools are providing better conditions for virtual tutoring.

Tables 6 and 7 show the results for regressions estimating the relationship between overall disruption and tech disruption only respectively, and tutoring implementation variables. We find that schools that implemented tutoring in a common space of the school tend to have slightly less disruption. This result ceases to be significant once we include variables that represent tutoring time of day for overall disruption, but persists for tech disruption, indicating that common space tutoring in our sample is associated with better technology or technology setup than the classrooms. Once we include tutoring time of day, we observe that sessions earlier in the day tend to be significantly less disruptive than mid-day sessions for overall disruptions, and sessions later in the day tend to be significantly more disruptive. The former is not significant for tech disruption, while the latter is.

5 Discussion and Next Steps

Our study relies on tutor speech to identify the occurrence of disruptions, which is likely an underestimate of the true occurrence of disruptions. By using language as a marker, we are able to identify cases when the usable instruction time is clearly interrupted. However, we are not able to identify and measure situations when the quality of instruction is reduced due to a persistent disruption that cannot be avoided. For example, during our qualitative analysis of audio recordings, we noticed excessive background noise in several student recordings, which represented a challenge for understanding student speech and would likely affect the tutor ability to respond to mistakes. However, our estimates from tutor speech result in only 2.1% of session time disrupted due to background noise. Future work should incorporate multi-modal elements to compose a more comprehensive measure of instruction quality potential.

The next steps in this study include investigating whether disruption can be associated with more or less effective tutoring through student learning outcomes. We will use an instrumental variable approach commonly known as the judge leniency design, which will allow us to approximate the effect of environmental disruption (tech and background) on student learning during tutoring isolated from the student's own disruptive tendencies.

References

R. G. Fryer Jr. The production of human capital in developed countries: Evidence from 196 randomized field experiments. In *Handbook of economic field experiments*, volume 2, pages 95–322. Elsevier, 2017.

A. Gromada and C. Shewbridge. Student learning time: A literature review. *OECD Education Working Papers*, 2016.

P. Holland, P. Alfaro, and D. Evans. Extending the school day in latin america and the caribbean. *World Bank Policy Research Working Paper*, (7309), 2015.

M. A. Kraft and M. Monti-Nussbaum. The big problem with little interruptions to classroom learning. *AERA Open*, 7:23328584211028856, 2021.

M. A. Kraft and S. Novicoff. Time in school: A conceptual framework, synthesis of the causal research, and empirical exploration. *American Educational Research Journal*, 61(4):724–766, 2024.

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692, 2019. URL `http://arxiv.org/abs/1907.11692`.

A. Nickow, P. Oreopoulos, and V. Quan. The promise of tutoring for prek–12 learning: A systematic review and meta-analysis of the experimental evidence. *American Educational Research Journal*, 61(1):74–107, 2024.

E. A. Patall, H. Cooper, and A. B. Allen. Extending the school day or school year: A systematic review of research (1985–2009). *Review of educational research*, 80(3):401–436, 2010.

C. D. Robinson and S. Loeb. High-impact tutoring: State of the research and priorities for future learning. *National Student Support Accelerator*, 21(284): 1–53, 2021.

M. Tkachenko, M. Malyuk, A. Holmanyuk, and N. Liubimov. Label Studio: Data labeling software, 2020-2022. URL `https://github.com/heartexlabs/label-studio`. Open source software available from https://github.com/heartexlabs/label-studio.

University of Oregon. 8th edition of dynamic indicators of basic early literacy skills (dibels®), 2021. Available: `https://dibels.uoregon.edu/`.

Ü. Yeşil Dağlı. Effect of increased instructional time on student achievement. *Educational Review*, 71(4):501–517, 2019.

A Appendix

A.1 Tables

	Not a Disruption	Student Disruption	Other Stop to Usable Time	Second Student
Intercept	0.825***	0.059***	0.009***	0.004***
	(0.013)	(0.006)	(0.001)	(0.001)
2:1 tutoring	-0.071***	0.030***	0.001*	0.012***
	(0.005)	(0.003)	(0.000)	(0.001)
Economically disadvantaged	-0.009	0.003	0.000	0.000
	(0.007)	(0.003)	(0.001)	(0.001)
English learner	0.006	-0.004	-0.001	-0.000
	(0.005)	(0.002)	(0.000)	(0.000)
SWD	-0.004	0.005	-0.001	0.001
	(0.011)	(0.006)	(0.001)	(0.001)
Black	-0.009	0.002	-0.001	0.002
	(0.008)	(0.004)	(0.001)	(0.001)
Latina/o/x	0.001	-0.002	-0.001	0.001
	(0.008)	(0.003)	(0.001)	(0.001)
Female	0.009*	-0.006**	0.000	0.001*
	(0.004)	(0.002)	(0.000)	(0.000)
Kindergarten	-0.008	0.005	-0.001*	0.000
	(0.006)	(0.003)	(0.000)	(0.001)
Second Grade	0.010*	-0.004	0.001**	-0.001
	(0.005)	(0.002)	(0.000)	(0.000)
School A	-0.014	-0.000	-0.002*	-0.001
	(0.010)	(0.004)	(0.001)	(0.001)
School B	0.002	0.002	-0.001	0.000
	(0.010)	(0.005)	(0.001)	(0.001)
School C	0.012	0.001	-0.003**	0.000
	(0.010)	(0.005)	(0.001)	(0.001)
School D	0.029**	-0.005	-0.003***	-0.002**
	(0.010)	(0.004)	(0.001)	(0.001)
School E	0.004	-0.003	-0.001	-0.001
	(0.011)	(0.004)	(0.001)	(0.001)
School F	-0.019	0.005	-0.002*	-0.000
	(0.011)	(0.005)	(0.001)	(0.001)
School H	-0.003	0.003	-0.002*	-0.000
	(0.011)	(0.005)	(0.001)	(0.001)
School I	0.039***	-0.000	-0.003***	-0.001
	(0.009)	(0.004)	(0.001)	(0.001)
School J	-0.017	0.003	0.000	0.001
	(0.011)	(0.005)	(0.001)	(0.001)
School K	-0.032***	0.007	-0.002*	0.002
	(0.010)	(0.005)	(0.001)	(0.001)
School L	0.018	0.010	-0.002**	0.001
	(0.010)	(0.005)	(0.001)	(0.001)
Observations	21937	21937	21937	21937
R^2	0.071	0.047	0.008	0.073
Adjusted R^2	0.070	0.046	0.007	0.072

Note: *p<0.05; **p<0.01; ***p<0.001

Table 4: Regression results for behavioral disruptions

	Tech Problem	Background Noise	Substitute Tutor
Intercept	0.083***	0.019***	0.001***
	(0.008)	(0.002)	(0.000)
2:1 tutoring	0.022***	0.005***	0.000***
	(0.003)	(0.001)	(0.000)
Economically disadvantaged	0.005	-0.001	0.000
	(0.005)	(0.002)	(0.000)
English learner	-0.000	-0.001	-0.000*
	(0.003)	(0.001)	(0.000)
SWD	-0.003	0.002	-0.000*
	(0.005)	(0.002)	(0.000)
Black	0.003	0.004*	0.000
	(0.004)	(0.002)	(0.000)
Latina/o/x	-0.003	0.003	0.000
	(0.004)	(0.002)	(0.000)
Female	-0.004	-0.000	0.000
	(0.003)	(0.001)	(0.000)
Kindergarten	0.005	-0.001	0.000**
	(0.003)	(0.001)	(0.000)
Second Grade	-0.008**	0.001	-0.000
	(0.003)	(0.001)	(0.000)
School A	0.013*	0.003	0.000
	(0.006)	(0.002)	(0.000)
School B	-0.003	-0.001	0.000
	(0.007)	(0.002)	(0.000)
School C	-0.006	-0.004***	-0.000
	(0.006)	(0.001)	(0.000)
School D	-0.014*	-0.005***	-0.000
	(0.006)	(0.001)	(0.000)
School E	0.004	-0.003*	-0.000
	(0.007)	(0.001)	(0.000)
School F	0.017*	-0.001	0.000
	(0.009)	(0.001)	(0.000)
School H	0.004	-0.001	-0.000
	(0.007)	(0.002)	(0.000)
School I	-0.028***	-0.006***	-0.000
	(0.006)	(0.001)	(0.000)
School J	0.013	0.000	-0.000
	(0.008)	(0.001)	(0.000)
School K	0.024***	0.002	0.000
	(0.006)	(0.002)	(0.000)
School L	-0.026***	-0.001	-0.000
	(0.005)	(0.002)	(0.000)
Observations	21937	21937	21937
R^2	0.040	0.025	0.003
Adjusted R^2	0.039	0.024	0.002
Note:			*p<0.05; **p<0.01; ***p<0.001

Table 5: Regression results for environmental disruptions

	Disruption			
Tutored in common space	-0.019***	-0.016***	-0.016***	-0.009
	(0.005)	(0.005)	(0.005)	(0.005)
Simultaneous tutoring sessions			0.000	-0.000
			(0.001)	(0.001)
Total school enrollment			-0.000	-0.000
			(0.000)	(0.000)
sessions / enrollment			0.039	0.133
			(0.419)	(0.410)
Tutoring began 7-10				-0.016*
				(0.007)
Tutoring began 14-17				0.013*
				(0.006)
Student-level covariates		X	X	X
Observations	21937	21937	21937	21937
R^2	0.005	0.019	0.023	0.027
Adjusted R^2	0.005	0.019	0.022	0.026
Note:				*p<0.05; **p<0.01; ***p<0.001

Table 6: Regression results for disruption on tutoring context information Student-level covariates include grade, gender, race, economically disadvantage indicator, EL, SWD. Simultaneous tutoring sessions calculated as the number of sessions starting per hour at each school.

	Tech Problem			
Tutored in common space	-0.012***	-0.011***	-0.011***	-0.007*
	(0.003)	(0.003)	(0.003)	(0.003)
Simultaneous tutoring sessions			0.000	-0.000
			(0.000)	(0.000)
Total school enrollment			-0.000	-0.000
			(0.000)	(0.000)
sessions / enrollment			0.301	0.301
			(0.280)	(0.275)
Tutoring began 7-10AM				-0.006
				(0.004)
Tutoring began 2-5PM				0.012**
				(0.004)
Student-level covariates		X	X	X
Observations	21937	21937	21937	21937
R^2	0.004	0.010	0.014	0.018
Adjusted R^2	0.004	0.009	0.013	0.017
Note:	*p<0.05; **p<0.01; ***p<0.001			

Table 7: Regression results for tech problems on tutoring context information. Student-level covariates include grade, gender, race, economically disadvantage indicator, EL, SWD. Simultaneous tutoring sessions calculated as the number of sessions starting per hour at each school.

A.2 Figures

Figure 1:

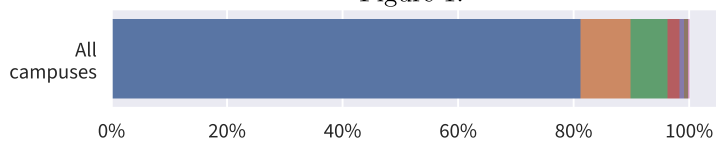


Figure 2:

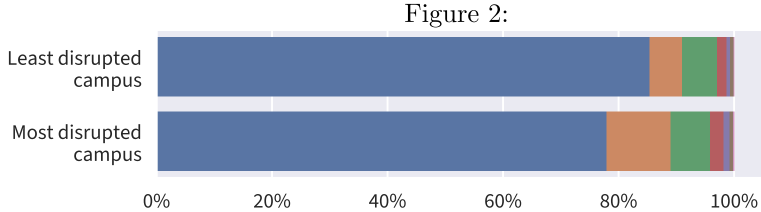


Figure 3:

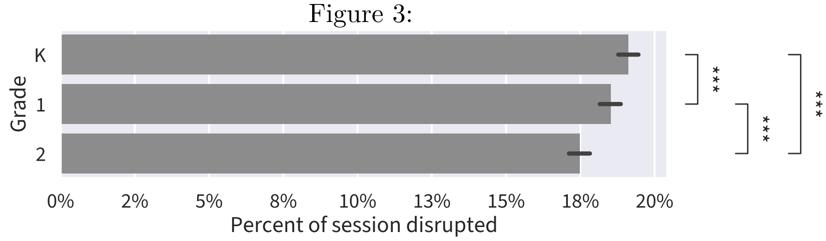
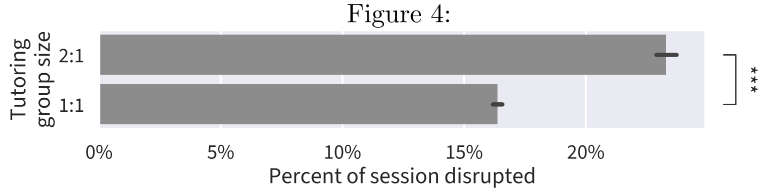


Figure 4:


Tutoring group size
2:1
1:1
***
0%
5%
10%
15%
20%
Percent of session disrupted