Efficacy Study of the Science Notebook in a Universal Design for Learning Environment

Jennifer W. Yu[1], Xin Wei[1], Kate Ferguson[1], Annie Fikes[1], Tracey Hall[2], and Jose Blackorby[2]

[1] Center for Learning and Development, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94205

[2] CAST, 40 Harvard Mills Square, Ste. 3, Wakefield, MA 01880

Abstract

Background: Research literature indicates that science notebooks can be used to support the active science learning process and the development of scientific literacy (Klentschy, 2005; Hargrove & Nesbit, 2003). However, paper-and-pencil-based science notebooks present barriers to students who struggle in the learning process because they require relative proficiency in reading and writing. Without sufficient skills in these domains, students are unable to use notebooks to support the development of deep understanding of the content. The Science Notebook for a Universal Design for Learning Environment (SNUDLE) was created to better access the use of science notebooks than traditional science notebooks for struggling and unmotivated students.

Purpose: This paper will describe the preliminary findings from Year 1 of a two-year efficacy study that addresses the following research hypotheses: 1) compared to fourth grade students using traditional paper-based science notebooks in business-as-usual (BAU) classrooms, students in SNUDLE classrooms will significantly increase their science content knowledge and motivation in science; and 2) a subsample of students with disabilities and students who are English language learners will experience significant positive interaction effects of the SNUDLE intervention.

Setting: The study focused on 4th-grade students in one large, urban school district. The participating district has 23 elementary schools with over 45,000 students. Nearly 36% of students are English learners, 81% receive free or reduced-price lunch, and 8% receive special education services. Less than 4% of 4th-grade students are identified as White.

Participants: The study sample was comprised of 717 students (394 intervention, 323 comparison) across seven elementary schools. All 4th-grade students participating in inclusive general education science classes were eligible to participate in the study. Demographics, including disability and ELL status were obtained from student records.

Intervention: Like traditional science notebooks, SNUDLE provides students with space to collect, organize, and display observations and data. However, using a universal design for learning (UDL) framework and a digital tablet platform, SNUDLE differs from traditional science notebooks in several key ways (Figure 1). First, SNUDLE was purposefully designed to lower construct-irrelevant barriers to science learning. For instance, text-to-speech technology is built directly into the notebook with real-time highlighting to support simultaneous access to auditory and visual processing. Second, pedagogy is built into the interface design, guiding students and teachers in the process of active science learning. For instance, students are prompted to think about making direct reference to their data and observations and to use relevant vocabulary. Finally, teachers are prompted and supported to provide feedback that may include corrective information, information to clarify ideas, or encouragement to engage in the scientific process.

In the study, the SNUDLE intervention was embedded with content from a commercially-available science curriculum already used by the schools. The study was conducted during the first 18 weeks of school, covering nine concepts from the curriculum.

Research Design: This efficacy study used a randomized control trial design based on classroom level randomization. Thirty-one 4th grade teachers were recruited and agreed to participate for 2 years with the expectation that they would teach in either the SNUDLE or BAU condition for the first year and then the opposite condition in the second year. This crossover design was feasible because the intervention's reliance on technology minimizes the threat of contamination.

Data Collection and Analysis: Curriculum-based unit quizzes, a validated assessment of broad science knowledge called the Measures of Academic Progress (MAP), and a student-reported Motivation for Science (MFS) Inventory were collected at baseline and post-test.

To test our hypotheses, we conducted hierarchical linear model (HLM) regressions with students nested within classrooms within teachers to compare the SNUDLE intervention group with the BAU group from pre- to post-test. Teacher were modeled as a random effect, and district and cohort were modeled as fixed effects. Covariates in the model included baseline measurements to reduce residual variability.

Results: Table 1 shows the demographic characteristics of the study sample compared to the overall school district. Characteristics by intervention and comparison group are found in Table 2. Table 3 describes the student assessment scores at baseline and posttest for the MAP, MFS, and summed unit quiz scores.

Primary estimates of the intervention effect were derived from the intent-to-treat analyses. Table 4 demonstrates that the treatment group did not differ from control group in MAP and total quiz scores after controlling for baseline demographic characteristics and achievement scores. Table 5 demonstrates that intervention students scored higher on the MFS efficacy subscale than the comparison group (Cohen's d=0.15, $p < .10$).

We examined the moderating effect of the intervention for students with disabilities and language spoken at home. A consistent positive interaction effect of the intervention's impact on the total quiz score (d = 0.89, p <0.01; Table 6) and for all four MFS subscales (d =2.53, p <0.01; d =1.64, p <0.05; d =3.13, p <0.05; d =2.05, p <0.001; Table 7) were found. We also found a positive trend in the impact of the intervention among students whose home language was Spanish (d =0.19, p <0.10; Table 7).

Conclusions: The preliminary findings from the first year of a two-year efficacy study provide evidence as to the efficacy of universally designed approaches to support students' active science learning, particularly among struggling students with disabilities or for whom English is not a first language.

Of note, a natural disaster occurred at the beginning of this study that closed all schools in the district, delaying study implementation by over a month, limiting the dosage of the intervention, and creating challenges with access to the technology. Despite these limitations, the current data lead to rich opportunities for ongoing research. Specifically, there is a need for continuation of this research to add further outcomes, confirm the positive outcomes in academic, motivation, and self-efficacy for the student subgroups seen in year one of the study, and investigate possible relationships of students' SNUDLE usage data to positive outcomes. We intend to address these and future research considerations as we engage in the second year of the study in order to support the building of students' understanding and sense-making skills in science education at the elementary level.

References

Klentschy, M. (2005). Science notebook essentials. *Science and Children, 43*, 24–27.

Hargrove, T. Y., & Nesbit, C. (2003). *Science notebooks: Tools for increasing achievement across the curriculum* (ERIC Document Reproduction Service Number ED 482720). Columbus, OH: ERIC Clearing- house for Science Mathematics and Environmental Education.

Figures and Tables

Figure 1. Theory of Change

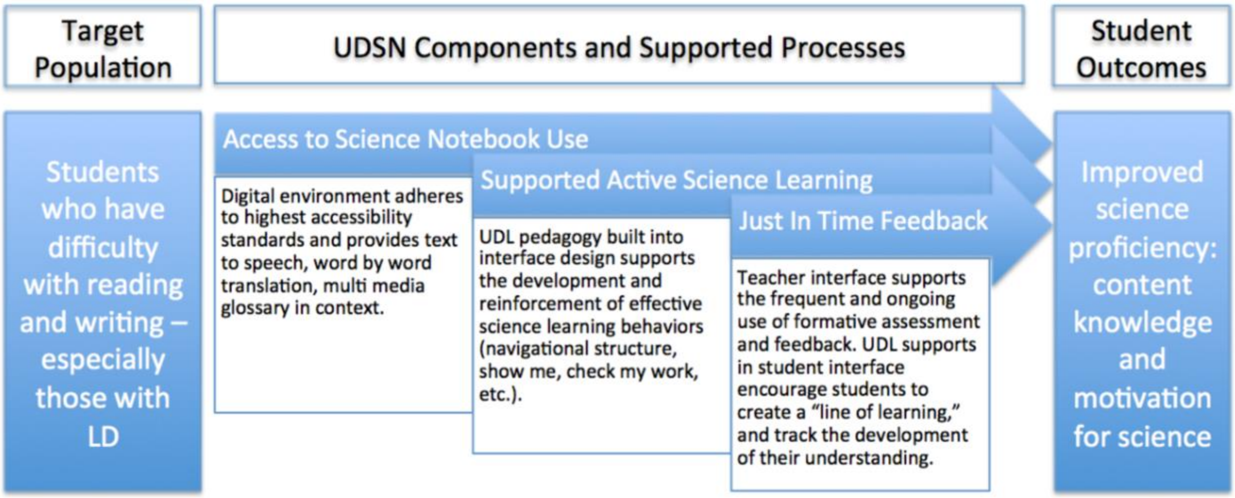

Table 1. Student demographic characteristics for participating district and study sample

Characteristics	Overall District (N=17,000) %	Study sample (N=717) %
Male	49.0	48.0
Race	52.7	45.0
Hispanic	52.7	45.0
White	4.1	34.6
African American	28.7	38.4
Asian	12.3	13.1
English language learner	42.7	25.0
Free/reduced price lunch	82.6	82.4
Has individualized education plan	7.2	6.0

Table 2. Baseline demographic characteristics of study sample students

Student characteristics	Intervention Group %	Comparison Group %
Male	49	47
Race		
White	32	38
Hispanic	38	54
African American	46	29
Asian	13	13
Free or reduced lunch status	82	83
Dual Language Learner status	25	25
Home language		
English	51	36
Spanish	31	47
Disability status	7	4

Note. Treatment sample size is 394 and control sample size is 323. There were no missing data on demographic variables for 375 out of the 394 treatment students and for. 323 out of the 394 control students.

Table 3. Student assessment scores by assignment

Student assessments	Treatment			Control		
	N	Mean	SD	N	Mean	SD
STAAR reading (pretest for MAP posttest)[1]	345	1410.15	159.47	270	1426.9	153.54
MAP Posttest	345	197.92	11.99	270	198.5	10.78
Motivation for Science Pretest: efficacy	302	3.36	0.61	230	3.34	0.54
Motivation for Science Posttest: efficacy	302	3.45	0.54	230	3.38	0.52
Motivation for Science Pretest: interest	301	3.12	0.45	229	3.13	0.44
Motivation for Science Posttest: interest	301	3.16	0.41	229	3.12	0.36
Motivation for Science Pretest: desire for challenge	300	3.01	0.48	230	3.09	0.4
Motivation for Science Posttest: desire for challenge	300	3.07	0.47	230	3.08	0.4
Motivation for Science Pretest: comfort using computer	301	3.17	0.6	231	3.13	0.58
Motivation for Science Posttest: comfort using computer	301	3.35	0.59	231	3.3	0.54
STAAR reading (pretest for total quiz score)[1]	350	1411.31	158.8	278	1428.24	153.25
Total quiz scores	350	40.03	15.79	278	40.4	17.12

[1]State of Texas Assessments of Academic Readiness (STAAR) in reading was used as the pretest measure for MAP and for the unit quizzes.

Table 4. Regression analysis results on student academic achievement in science outcome

Predictors	MAP		Total quiz score	
	β	SE	β	SE
Intercept	197.47***	1.83	40.56***	3.27
Pretest	0.05***	0.002	0.05***	0.003
Treatment	0.79	1.25	3.18	3.28
Male	1.93**	0.63	2.12*	0.85
American Indian/Alaskan Native	0.27	1.08	1.64	1.52
Asian	-0.06	1.64	5.79*	2.24
Black African American	-0.86	1.29	2.06	1.76
Native Hawaiian/Pacific Islander	-5.33	3.92	-7.78	5.38
Hispanic	-0.57	1.24	0.93	1.68
FRL	-0.38	0.84	-2.66*	1.16
DLL	-1.36	1.01	-6.24***	1.50
Home language English	-0.48	1.26	-6.29**	1.80
Home language other	1.91	1.61	-0.60	2.22
IEP	-4.01***	1.41	-3.22	2.05
504 plan	-3.32†	1.88	-5.74*	2.60

Note. Benjamin-Hochberg multiple comparison adjustment was not included in the analysis because the treatment impact was not significant.
*p<.05, **p<.01, ***p<.001

Table 5. Regression analysis results on motivation for science outcome

Predictors	Efficacy		Interest		Desire for challenge		Comfort using computer	
	β	**SE**	**β**	**SE**	**β**	**SE**	**β**	**SE**
Intercept	3.48***	0.11	2.95***	0.08	3.06***	0.09	3.29***	0.11
Pretest	0.34***	0.04	0.35***	0.03	0.30***	0.04	0.51***	0.04
Treatment	0.08†	0.04	0.04	0.03	0.03	0.04	0.04	0.05
Male	0.05	0.04	0.01	0.03	-0.06†	0.04	0.07	0.04
American Indian/Alaskan Native	-0.03	0.07	-0.02	0.05	0.006	0.06	-0.13†	0.07
Asian	-0.13	0.11	-0.02	0.08	-0.09	0.09	0.04	0.11
Black African American	-0.03	0.09	0.04	0.06	0.09	0.07	-0.01	0.08
Native Hawaiian/Pacific Islander	0.15	0.35	0.28	0.25	0.13	0.29	0.14	0.35
Hispanic	-0.05	0.08	0.11†	0.06	0.13†	0.07	-0.05	0.08
FRL	-0.01	0.06	0.09*	0.04	-0.02	0.05	0.07	0.06
DLL	-0.03	0.05	0.005	0.04	-0.13**	0.05	0.09	0.06
Home language English	-0.09	0.08	0.05	0.05	-0.06	0.06	-0.11	0.08
Home language other	-0.08	0.11	0.0003	0.08	0.17†	0.09	-0.08	0.11
IEP	-0.17	0.10	0.01	0.07	-0.14	0.08	-0.02	0.10
504 plan	-0.22	0.16	-0.37***	0.11	-0.20	0.13	-0.21	0.15

Note. Benjamin-Hochberg multiple comparison adjustment was not included in the analysis because the four subscales of the MFS are in different domains.
†*p*<.10, **p*<.05, ***p*<.01, ****p*<.001

Table 6. Moderator analysis results of Measures of Academic Progress and content-based quiz score by disability status and language

Variable	MAP		Total quiz score	
	Impact estimate	**SE**	**Impact estimate**	**SE**
Disability status	1.01	2.64	14.69**	4.72
Dual language learner	0.31	1.75	1.58	4.78
Home language – English	0.78	1.48	3.06	3.70
Home language – Spanish	-2.27	1.56	1.12	3.72
Home language – Other	3.77	1.64	5.22	4.29

†*p*<.10, **p*<.05, ***p*<.01, ****p*<.001

Table 7. Moderator analysis results of Motivation for Science subscales by disability status and language

Variable	Efficacy		Interest		Desire for challenge		Comfort using computer	
	Impact estimate	**SE**	**Impact estimate**	**SE**	**Impact estimate**	**SE**	**Impact estimate**	**SE**
Disability status	1.34**	0.40	0.74*	0.34	1.22*	0.42	0.90***	0.23
Dual language learner	-0.001	0.08	0.002	0.08	-0.08	0.08	-0.02	0.09
Home language – English	0.05	0.08	-0.04	0.06	0.03	0.07	0.04	0.07
Home language – Spanish	0.10†	0.06	0.07	0.05	0.07	0.05	0.04	0.08
Home language – Other	0.01	0.10	0.04	0.10	-0.10	0.10	-0.04	0.09

Note. Benjamin-Hochberg multiple comparison adjustment was not included in the analysis because the four subscales of the MFS are in different domains.
†*p*<.10, **p*<.05, ***p*<.01, ****p*<.001