DESIGNING MATHEMATICS LEARNING ENVIRONMENTS FOR MULTILINGUAL STUDENTS: RESULTS OF A REDESIGN EFFORT IN INTRODUCTORY ALGEBRA

DISEÑANDO AMBIENTES DE APRENDIZAJE EN MATEMÁTICAS PARA ESTUDIANTES MULTILINGÜES: RESULTADOS DE UNA INVESTIGACIÓN DE DISEÑO EN ALGEBRA INTRODUCTORIA

William Zahner
San Diego State University
bzahner@sdsu.edu

Kevin Pelaez
San Diego State University
kpelaez10@gmail.com

Ernesto Calleros
San Diego State University
ernestocalleros@gmail.com

We describe results of a long-term design experiment focused on promoting mathematics learning among multilingual ninth graders classified as English Learners. The intervention at a linguistically diverse public high school in the US focused on a unit introducing concepts related to linear rates of change. We analyze results from a curriculum-aligned pre- and post-unit assessment used to document student learning across each design cycle. The main result is that students in both the Pre-Intervention and Redesigned classrooms made gains on the pre- and post-unit assessments. However, student gains on the assessments were higher in the Redesigned classrooms than in the Pre-Intervention classrooms. Additionally, on the assessment in the Redesigned classes, students classified as English Learners made larger gains than their non-EL peers, and the majority of the gains occurred on conceptually-focused items.

Keywords: Algebra and Algebraic Thinking, Design Experiments, Curriculum, Equity and Diversity

Objective

Research on the mathematics education of emerging multilingual learners[1] has shown that such students often experience procedurally-focused instruction (Callahan, 2005), are provided low cognitive demand tasks (de Araujo, 2017), and may have limited opportunities to engage in disciplinary and discourse practices (Zahner, 2015; Moschkovich & Zahner 2018). Recently, mathematics educators have engaged in design research to study how to transform the learning environment in linguistically diverse classrooms with the goal of promoting more robust forms of student learning (e.g., Chval et al., 2014; Prediger & Zindel, 2017). These design efforts have yielded promising results and frameworks for integrating mathematics and language learning. In this report, we describe results from a four-year design research effort that took place in a linguistically diverse ninth grade mathematics classes in an urban secondary school in the US. We analyze trends observed in the student responses to a curriculum-aligned assessment used to evaluate the efficacy of the design effort.

Our overarching research question is: To what extent did the design effort meet its goal of promoting student learning? In particular, we address the three specific questions

1. What was the effect of the redesign effort as measured by student response patterns on curriculum-aligned assessments?
2. How did the assessment results differ for students classified as ELs and those not classified as ELs (non-ELs)?

[1] In our research context, multilingual students who are learning the language of instruction are classified as "English Learners" (ELs). We use ELs when describing students as they are classified by the school. However, when referring to this group of students more generally, we will use "multilingual students" to highlight the assets of students learning the language of instruction.

In: Sacristán, A.I., Cortés-Zavala, J.C. & Ruiz-Arias, P.M. (Eds.). (2020). *Mathematics Education Across Cultures: Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Mexico. Cinvestav / AMIUTEM / PME-NA. https:/doi.org/10.51272/pmena.42.2020

3. On what kinds of problems, conceptual or procedural, in the assessments of the Redesigned lessons did the students make the largest gains?

Framework

This design research (Cobb et al., 2003) project utilized the Academic Literacy in Mathematics (ALM) framework (Moschkovich, 2015, Moschkovich & Zahner, 2018), a sociocultural framework for analyzing and designing mathematics learning environments for emerging multilingual students. The ALM framework highlights that developing academic literacy entails developing mathematical proficiencies (in particular, procedural fluency and conceptual understanding), engaging in disciplinary practices, and participating in mathematical discourse. In our design efforts, we focused on developing students' conceptual understanding of the slope of a linear function as representing a rate of change (Lobato & Thanheiser, 2002). Based on prior research, we know that many students develop procedure-bound understandings of linear rates of change. For example, students can calculate a rate of change given a well-ordered table, but then they may struggle to interpret the rate in a given problem context (Lobato & Siebert, 2002). Therefore, one of our goals was to develop students' conceptual understanding of a linear rate of change as a multiplicative relationship between quantities.

Building on insights from prior design efforts in linguistically diverse mathematics classrooms (Chval et al., 2014; Prediger & Zindel, 2017) and the ALM framework, our redesign efforts were based on three guiding principles: (a) aligning a conceptual focus and problem contexts across the unit to minimize linguistic complexity, (b) integrating mathematical language goals linked to the conceptual focus, and (c) incorporating language supports in daily lesson activities to allow all members of a linguistically diverse classroom to engage in classroom discussions. In support of this design research effort we collected a wide array of data and we are conducting both quantitative and qualitative analyses of these data. In this report, we report on quantitative evidence of student learning measured in pre-unit and post-unit assessments. These assessments were designed by the participating teachers and researchers to align with the content of the unit, and to target both conceptual and procedural knowledge. In this sense, the assessment provides data related to the first component of the ALM framework (Moschkovich, 2015).

Data and Methods

Design Cycles

The classroom-based design research entailed three main phases and two design cycles, shown in Figure 1. In Phase I, Pre-Intervention, the researchers observed ninth grade mathematics classes at City High across a unit on linear rates of change. Students in the observed classes completed a pre-unit assessment at the start of the unit and a post-unit assessment at the conclusion of the unit. Additionally, during Phase I, the researchers conducted clinical interviews with selected students from the observed classes. Due to space restrictions, the interviews are not discussed further here.

At the conclusion of Phase I, the teachers and researchers analyzed the Phase I data with the goal of redesigning the unit on linear rates of change. The Redesigned lessons were then pilot tested in Phase II during Teaching Experiment (TE; Prediger, Gravemeijer, & Confrey, 2015). The teaching experiment lessons were taught in an after school setting, which allowed the researchers and teachers to make adjustments in the lesson designs from week to week. The TE lessons were designed using the project design principles outlined above. The teaching of the lessons in the TE was the conclusion of the first iteration of the design cycle and the start of the second iteration.

Finally, in Phase III, the team of teachers and researchers analyzed the data from the Phase II TE lessons and redesigned the unit lessons from Phase II for use during the regular school day. The participating teachers implemented the Redesigned lessons during their typical school day in Phase

III. During this second iteration of the design cycle, the classroom lessons were observed and the students in the Phase III classes completed a pre-unit and post-unit assessment (paralleling the Pre-Intervention observations from Phase I). In the results reported here, we focus on the assessments that were administered at the start and conclusion of the classroom observations in the Pre-Intervention (Phase I) and Redesigned (Phase III) units.

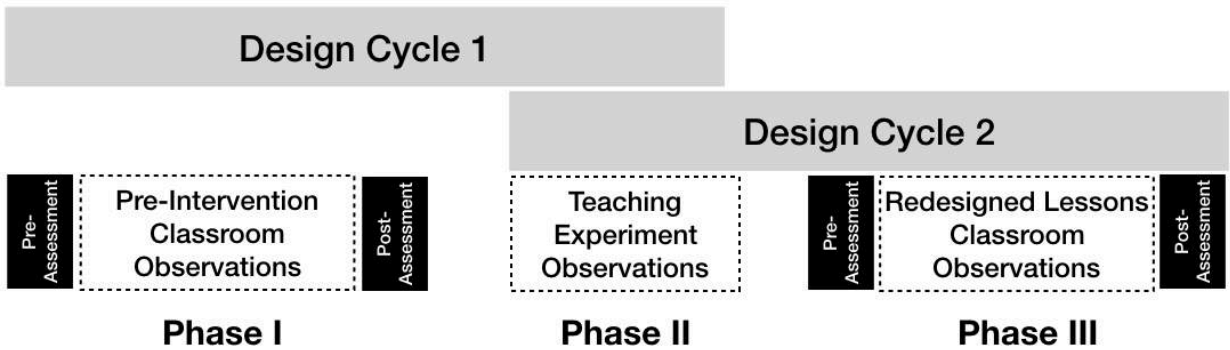

Figure 1: Overall project design and phases of design cycles.

Setting and Participants

Data collection took place in ninth grade mathematics classes at City High, a large, urban high school enrolling a linguistically diverse group of students. City High is located near the US-Mexico border, and it serves a relatively large number of recent immigrant and transborder students. Throughout the study, approximately 30% of City High students were classified as ELs, and an additional 50% of the students were formerly classified as ELs. Over 75% of the students were identified as Latinx in school demographic data, and most students classified as ELs spoke Spanish as their primary language. Additionally, the majority of City High students were from households with low socioeconomic status.

Six teachers participated in the design effort across the four-year project. In this analysis, we focus on the assessment results for the students of Mr. S because he was the only teacher who participated in all phases of the design effort. Mr. S had six years of experience at the start of the study, holds a teaching credential in mathematics, is bilingual in Spanish and English, and has taken some courses on teaching English learners in the content areas as part of his mathematics teaching certification program. He participated in the study to further develop his expertise teaching mathematics to multilingual students.

We included all students in Mr. S's focal classes who took the pre-unit and post-unit assessments during the Pre-Intervention (n=18) and Redesigned (n=28) phases of the research. We note that the assessments were designed to align with the respective units. Therefore, we were not able to use the same items on the assessments for the Pre-Intervention and Redesigned lessons. To account for these differences in the assessment design, we converted the total scores on each assessment to the proportion of correct answers.

Assessment

The pre- and post-assessments were designed by the researchers in collaboration with the teachers. In both the Pre-Intervention and Redesigned units, the assessments included a mixture of procedurally-focused items that did not require explanation, as well as conceptually-focused items characterized by prompts for explanation. Aligning with the design guidelines for the ALM framework (Moschkovich & Zahner, 2018), the focus of the intervention was developing students' conceptual understanding of rate of change. The conceptual items reflected this focus and assessment goal. Figure 2 shows an example of a procedural item (top) and conceptual item (bottom) from the Phase III assessment. In the Results section, we compare patterns of student responses to the conceptually-focused and procedurally-focused items from Phase III.

1) Deondra is saving to buy a new bike. She starts off with $10 in her account and saves $20 a month babysitting. She writes an equation that represents the total amount of money (y) she has saved in any given month (x):

$$y = 20x + 10$$

How much money did Deondra save after 6 months?

❑ $20
❑ $100
❑ $120
❑ $130
❑ $150

6) Connie ran 45 meters in 10 seconds. Jen ran the same speed as Connie, but Jen ran a different amount of time and distance.

a) Which of the following are possible for the time and distance that Jen ran? **Select all that apply.**

❑ 90 meters in 5 seconds
❑ 54 meters in 12 seconds
❑ 22.5 meters in 5 seconds
❑ 90 meters in 20 seconds
❑ 50 meters in 15 seconds
❑ 4.5 meters in 1 second

b) Explain: How can you check if Jen's distance and time makes the same speed as Connie?

Figure 2: A procedural item (top) and a conceptual item (bottom) from the Redesigned assessments.

The assessments were scored using a rubric by the research team. In general, students were awarded points for correct responses as well as valid explanations (on the conceptual items). For the analysis of assessment gains, we used comparisons of means and paired effect sizes (Cohen, 2013) to identify the direction and magnitude of change from the pre-assessment to the post-assessment in each phase. For the comparison of student performance on the conceptually-focused versus procedurally-focused items, we also compared means and measured paired effect sizes. We used paired effect sizes because we had both pre-assessment and post-assessment data by student. The details of these analyses are in the Results.

Results

We note a caveat at the outset of this section: we had relatively small sample sizes. These small sample sizes are a result of engaging in intensive work with a small group of teachers across multiple years. Our sample sizes were reduced by teacher turnover across study years and student absences/mobility within each design cycle. These realities, while not ideal for statistical research, are byproducts of working in our local school setting. In light of the small samples, we used Hedges' correction whenever presenting claims about growth as measured by effect sizes.

Research Question 1: What was the effect of the redesign effort as measured by student responses on the pre- and post-assessments?

To answer RQ1, we compared results of the pre- and post-assessment gains in Pre-Intervention and the Redesigned units. Table 1 shows the pre- and post-assessment mean scores along with the paired effect size to measure the magnitude of the gain.

Table 1: Pre- and post-assessment means for the Pre-Intervention and Redesigned units.

	Pre-assessment	Post-assessment	Gain	Paired effect size (Magnitude)[1]
Pre-Intervention (n=18)	0.42 (0.17)	0.57 (0.18)	+0.15	0.80 (L)
Redesign (n=28)	0.24 (0.16)	0.44 (0.19)	+0.20	1.12 (L)

Note: [1]S: $|d|<0.20$, M: $0.20 < |d| < 0.50$, L: $|d| > 0.50$ (Cohen, 2013)

In summary, there was significant growth in both the Pre-Intervention and Redesigned units as measured by the pre- and post-assessments. The effect size comparing the magnitude of the pre-post gain was larger in the Redesigned unit. While the direction of the difference in effect sizes was favorable for our design effort, given the small sample size, we were not able to use a statistical test such as an ANOVA to compare the difference in effect sizes.

Research Question 2: How did the assessment results differ for students classified as ELs and those not classified as ELs (non-ELs)?

Table 2 shows the pre- and post-assessment mean scores broken out by Pre-Intervention and Redesigned Lessons and student EL classification. As with Table 1, we included a paired effect size to illustrate the magnitude of the differences (Cohen, 2013). The data in Table 2 indicate that while there was significant growth from the pre- to post-assessment scores in the Pre-Intervention unit, the effect size was larger for non-ELs (0.81) than for Els (0.58). That is, during the Pre-Intervention unit, the students who were not classified as ELs benefitted more from the business as usual teaching. During the Redesigned unit, however, the pattern in effect size was reversed. That is the effect size was larger for ELs (1.44) than for non-ELs (0.94). In terms of the mean gain scores, ELs in both the Pre-Intervention and the Redesigned lessons had larger gains than their non-EL classified peers. One interpretation of this result is that the intervention led to larger gains for ELs than non-ELs. However, this result is moderated by the fact that ELs in the Redesigned unit pre-assessment started relatively low compared to the non-ELs.

Table 2: Pre- and post-assessment mean scores by EL classification for the Pre-Intervention and Redesigned units.

		Pre-assessment	Post-assessment	Gain	Paired effect size (Magnitude)
Pre-Intervention					
	Non-EL (n=14)	0.45 (0.16)	0.59 (0.18)	+0.14	0.81 (L)
	EL (n=4)	0.33 (0.21)	0.49 (0.21)	+0.16	0.58 (L)
Redesign					
	Non-EL (n=20)	0.27 (0.17)	0.45 (0.20)	+0.18	0.94 (L)
	EL (n=8)	0.17 (0.11)	0.43 (0.19)	+0.26	1.44 (L)

RQ3: On what kinds of problems, conceptual or procedural, in the assessments of the Redesigned unit did the students make the largest gains?

Our final research question examines patterns in student responses to the pre- and post-assessments from the Redesigned unit. The assessment for the Redesigned unit included five conceptually-focused and three procedurally-focused items. Conceptually-focused items included a prompt for explanation while the procedurally-focused items were multiple-choice questions with no prompt for further explanation. To conduct the analysis of student results based on item type, we calculated a subscore for each type of item (e.g., a subscore for the three procedural items, and a subscore for the five conceptual items). Table 3 summarizes the results by type of question (Conceptual versus Procedural).

Table 3: Pre- and post-assessment mean scores by type of item for the Redesigned unit.

Type of Item	Pre mean (SD)	Post mean (SD)	Growth	Paired Effect Size (Magnitude)
Procedural	0.47 (0.22)	0.44 (0.26)	-0.03	0.12 (S)
Conceptual	0.20 (0.16)	0.44 (0.21)	0.25	1.27 (L)

Table 3 shows there was little change in the mean for the procedural items from the pre-assessment to the post-assessment. However, there were significant gains in the conceptual items. This result aligns with the focus of the redesign efforts on fostering student development of conceptual understanding related to linear rates of change.

Discussion & Conclusion

In this report we have described quantitative findings based on pre- and post- unit assessments that were used as part of a design research effort. The goal of the design research effort was to create a classroom learning environment in a linguistically diverse secondary mathematics classroom where students learn to reason about linear rates of change. The main findings presented in this report are:

1. Students in the Redesigned unit showed larger gains as measured by paired effect size with Hedges correction,

2. ELs in the Redesigned unit had larger gains when compared to their non-EL counterparts. The gains by Els in the Redesigned unit were also relatively larger than their gains in the Pre-Intervention lessons, and
3. In the assessment of the Redesigned unit, the students had more growth on the conceptually-aligned items than on the procedural items.

While we have shown that the paired effect size was larger in re-designed unit, we are not able, at this point, to make a causal claim about learning linked to our design efforts. However, we can hypothesize that some of the growth observed in our analyses can be traced back to the design principles and the instructional activities that were developed across the two design cycles. For example, in the Pre-Intervention lessons, we observed that most students had limited opportunities to solve conceptually demanding tasks during their regular mathematics classes. Therefore, in the Redesigned unit, we intentionally included conceptually demanding tasks, along with targeted linguistic supports for students classified as ELs, to allow all students in the Redesigned classroom an opportunity to solve more conceptually-focused problems. For example, during one of the Redesigned lessons students were challenged to make up a story related to a graph showing distance and time. Then they were tasked with solving non-routine problems about rates of change in other graphs. These instructional tasks, which consistently included prompts for students to explain their reasoning and justify responses, were similar to the conceptually-focused item in Figure 2.

One question that remains for our design efforts is why the students in the Redesigned unit did not show gains on the procedural items. When we ran the analysis for RQ3, we expected to see growth on both the conceptual and the procedural items. However, looking at the results in Table 3, there was little change on the procedural tasks from the pre-assessment to the post-assessment. One possible explanation for this lack of gain is that the learning opportunities within the Redesigned unit did not include many tasks focused on developing students' procedural fluency. Reflecting on the data in Table 3 prompts us to consider whether the learning tasks in the Redesigned lessons overemphasized a conceptual focus and did not include enough opportunities for students to develop procedural fluency.

Despite the limitations induced by the small sample size of this study, the results presented here highlight potential avenues for further development and research. This is of particular interest since there are relatively few studies, like the one presented here, that have engaged in design research focused on meeting the needs of linguistically diverse students. One potential follow-up study might be to use the materials developed in this project in a study with a larger number of teachers and students. Such a study, particularly one using a more controlled design, could test of the robustness of the results presented here.

Another avenue for follow-up on this research, which we are undertaking presently, is qualitative analyses of the design effort. In these analyses we are examining the forms of student reasoning that were developed during the Redesigned lessons and tracing those forms of reasoning back to the classroom learning environment. For example, we are examining the quality of student explanations on the conceptually-focused items, and then connecting the student reasoning observed on the assessments to the classroom observations (paralleling the analysis of Zahner, 2015). This analysis will allow us to understand *how* a linguistically diverse group of ninth grade students developed particular forms of reasoning about linear rates of change.

Acknowledgments

This research was supported by a grant from the National Science Foundation (#1553708) to William Zahner. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of The National Science Foundation.

We are grateful for the support of research assistants Yessika Gamala and April Zuniga and the collaboration of the teachers and students who are participating in this study.

References

Callahan, R. M. (2005). Tracking and high school English Learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305–328. https://doi.org/10.3102/00028312042002305

Chval, K. B., Pinnow, R. J., & Thomas, A. (2014). Learning how to focus on language while teaching mathematics to English language learners: A case study of Courtney. *Mathematics Education Research Journal*, *27*(1), 103–127. https://doi.org/10.1007/s13394-013-0101-8

Cobb, P., Confery, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*, 9–13.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences.* Academic Press.

de Araujo, Z. (2017). Connections between secondary mathematics teachers' beliefs and their selection of tasks for English language learners. *Curriculum Inquiry, 47*(4), 363–389. https://doi.org/10.1080/03626784.2017.1368351

Lobato, J., & Siebert, D. (2002). Quantitative reasoning in a reconceived view of transfer. *The Journal of Mathematical Behavior, 21*(1), 87–116. https://doi.org/10.1016/S0732-3123(02)00105-0

Lobato, J., & Thanheiser, E. (2002). Developing understanding of ratio as measure as a foundation for slope. In B. Litwiller (Ed.), *Making sense of fractions, ratios, and proportions: 2002 yearbook* (pp. 162–175). National Council of Teachers of Mathematics.

Moschkovich, J. N. (2015). Academic literacy in mathematics for English Learners. *The Journal of Mathematical Behavior, 40*, 43–62. https://doi.org/10.1016/j.jmathb.2015.01.005

Moschkovich, J. N., & Zahner, W. (2018). Using the academic literacy in mathematics framework to uncover multiple aspects of activity during peer mathematical discussions. *ZDM Mathematics Education*, *50*(6), 999–1011. https://doi.org/10.1007/s11858-018-0982-9

Prediger, S., Gravemeijer, K., & Confrey, J. (2015). Design research with a focus on learning processes: An overview on achievements and challenges. *ZDM Mathematics Education*, *47*(6), 877–891. https://doi.org/10.1007/s11858-015-0722-3

Prediger, S., & Zindel, C. (2017). School academic language demands for understanding gunctional relationships: A design research project on the role of language in reading and learning. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(7b), 4157–4188. https://doi.org/10.12973/eurasia.2017.00804a

Zahner, W. (2015). The rise and run of a computational understanding of slope in a conceptually focused bilingual algebra class. *Educational Studies in Mathematics*, *88*(1), 19–41. https://doi.org/10.1007/s10649-014-9575-x