THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

COLLEGE OF INFORMATION SCIENCES AND TECHNOLOGY

Understanding Undergraduates' Perceptions of Student Debt Visualizations: A Human-Centered
Approach

GLENN HUBBARD
SPRING 2023

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Applied Data Sciences
with honors in Applied Data Sciences

Reviewed and approved* by the following:

Syed Billah
Assistant Professor of Information Sciences and Technology
Thesis Supervisor

John Yen
Professor of Information Sciences and Technology
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

Data Science, as a field, relies heavily on the translation of analytic results into visualizations to aid in the explainability of these results. Visualizations, such as graphs and infographics, serve as a tool to quickly interpret quantifiable results, however the design of these elements is often performed without a proper assessment of the data interpretation abilities of the target audience. Moreover, without consulting a particular audience, the effectiveness of one's visualization can be lost within this audience, since everyone's background may impact their understanding differently.

This paper focuses on the design elements that work towards the creation of effective and broadly understandable data visualizations that relate to student debt. Student debt is of particular interest as this form of debt if applicable to millions within the United States. Student debt has been labeled as a financial burden that impacts home ownership and financial stability, which can have a cascading effect for the rest of the U.S. economy. While efforts are being made within the Biden Administration to cancel a portion of the U.S.'s student debt burden, these efforts are currently being held back in the Federal Court System. With no relief in sight, learning ways for individuals to best tackle and understand the impact of their debt burden is of critical importance for personal and societal financial stability.

Through an interview-based IRB-approved study of Penn State University – University Park undergraduates, this paper will explore the aspects of effective debt visualizations and how one's background impacts their understanding of both student debt and data visualization interpretations. Finally, design guidelines for stakeholders in the student debt industry will be presented for the assistance in the creation of future visualizations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Chapter 1

Introduction

Student debt is an evolving crisis. Home ownership rates, especially amongst young people, exists as a negative inverse relationship with student debt—as debt burdens increase, home ownership rates decrease. Home ownership has proven links to wealth holdings within a household, which can be an indicator of financial stability (Turner 2009). Despite home ownership being an effective tool for creating a socially and financially stable future—and student debt being an impediment to the ability to buy a home—there is currently little legislative success for improving the student debt crisis (Studentaid.gov 2022). While home ownership isn't the only way to shore up financial stability, it consistently presents itself as a major driver of total economic activity within the United States (National Association of Realtors 2009). With cost-of-living increases across the United States occurring at unprecedented rates, regardless of ability to buy a home, the ability to afford a stable life and repay a student loan is increasingly difficult, simply due to the opportunity cost of the dollar (SSA.gov 2022). With a stalemate coming out of the Biden Administration's plan for debt forgiveness, we must consider alternative strategies to protect the financial stability of college graduates within the status quo.

One of these strategies, which this paper will focus on, involves students becoming more educated on their individual loan burden. Students need effective tools to help them finalize their loan borrowing decision, as well as grasp the minutia of this decision and how it will impact their life in the future (Perna 2017). To better develop this strategy, this paper will focus on reviewing the results of an IRB-approved study (conducted by the paper's author), which seeks to explore

how individuals interpret student debt-related data visualizations and the relationship between an individual's knowledge of student debt and these interpretations.

Data visualizations provide an effective means to rapidly convey information and allow information to analyzed and interpreted (Unwin 2020). Despite data visualizations becoming increasingly common in all aspects of daily life, little work is being done towards exploring the design of the visualizations and influence of these visualizations on student debt perceptions. This paper aims to explore these perceptions within an undergraduate population as well as identify the key components of debt visualizations that impact undergraduate's understandings.

I am particularly interested in this exploration, since prior to reaching the conclusion of my undergraduate career, I failed to explore my options for debt repayment. I often told myself that it would be a 'later issue', but as I am reaching graduation, that 'later issue' is now on the horizon. While I had used simplistic debt calculation and repayment tools when deciding to take out loans, these tools often painted a picture I wasn't comfortable with, and since I wanted to go to college, these tools would not have swayed my decision. Throughout all the tools that I used, there was a severe lack of effective graphs. If graphs did exist (which wasn't always the case), they often existed in a vacuum—exemplified by incredibly bland aesthetics and no external educational materials. As an 18-year-old, I wasn't thinking about home ownership, retirement, or family planning. Now, these topics are top of mind. While I can't say for certain if alternative or improved data visualizations would have swayed my decision in a different way, I believe that these alternatives are worth exploring.

Despite my own experiences, millions of students around the United States make the decision to take on debt for their education before understanding just how much debt they might acquire. Colleges are often not upfront with the total cost of an undergraduate education and can

fail to label what money is 'free'—such as a grant—and what money needs to be paid back. These discrepancies, and the financial sleuthing needing to be performed by potential students and their families, highlights even greater troubles for new students, perhaps leading them into a debt burden they did not fully understand before they had signed their promissory notes (Nadworny 2022). Moreover, financial literacy surrounding student debt can potentially create long-term fiscally responsible behaviors, which is of growing importance as financially literacy comes at a cost—the average American loses $1,800 annually to easily correctible financial mistakes (Moorcraft 2023).

Research Questions

This paper delves into three research questions. The first question strikes directly at the motivations for this thesis.

RQ1: *"Is there a gap in undergraduate knowledge of student debt, and if so, what is this gap"?*

This question is context-dependent, meaning that prior exposure to student debt, classes on financial literacy, personal debt situations, and general numerical aptitude can play a role in determining how and why some students understand debt differently than others.

RQ2: *"What aspects of current visualization design for student debt situations are most important for enhancing one's understanding of debt"?*

This question was chosen as it might be able to provide insights into why certain visualizations are more effective than other ones at increasing understanding of debt, which should be of critical interest for stakeholders in the student debt industry such as educational institutions, loan providers, and the Federal Department of Education.

RQ3: *"What are potential improvements to current data representations that can improve undergraduate understandings of student debt"?*

With this question, I hope to develop guidelines that can advise stakeholders on ways in which they can improve their current visualization apparatus.

Following this section, Chapter 2 will discuss the related works surrounding student debt and data visualizations, which will be used to paint a picture for why this research is critical as well as provide context for findings in the subsequent chapters. Chapter 3 will discuss the methods used to design, collect, and analyze data for this study. Chapter 4 will review these findings, while Chapter 5 will apply these findings towards the larger body of visualization literature. Chapter 6 will conclude this thesis.

Chapter 2

Background and Related Work

This thesis draws from disciplines such as human-computer interaction, finance, and psychology to craft a more precise understanding of how undergraduate students grasp financial visualizations. Prior work has touched on this intersection, often aiming to understand how quantitative ability and prior financial literacy affect one's financial behavior, however most of these works either don't include student debt as the financial product of study or fail to include assessments of visualizations. This thesis will bridge this gap using student debt as the financial product of interest with the end goal of creating a framework for more effective visualization designs.

Financial Literacy

While institutions of higher education are mandated by the Higher Education Act of 1965 to provide loan counseling to students, this counseling is often focused on showing students basic information regarding their loan, rather than providing information that might increase their financial literacy regarding their loan (McCarthy, 2015). The loan tool provided by the Department of Education—Federal Student Aid *Entrance Counseling*—is the recommended tool for Penn State undergraduates (Penn State Office of the Bursar, 2022). According to McCarthy, this tool "does not provide students with the information necessary to increase their overall financial literacy which is key to helping reduce debt". The tool covers information such as the costs of a college education, repayment options, interest, and loan options, as depicted in **Figure 1**. The tool includes an *extremely* limited visualization component and exists as a primarily text-based educational platform. The tool is available to demo at studentaid.gov (Department of Education, 2022).

Acknowledging the limitations of the widely used federally provided entrance counseling tool, the Department of the Treasury outlines a series of best practices for institutions of higher education regarding their financial literacy offerings for students (U.S. Financial Literacy and Education Commission, 2019). These suggestions, which include mandated financial literacy courses, peer educational programs, financial literacy as a component of the core curriculum, and enhanced communication with students, are not mandatory to incorporate (Dept. of Treasury 2019).

The completion of *Exit Counseling* tool provided by the U.S. Department of Education is required for all graduating students who have taken federal loans. This tool is external from the student's undergraduate institution; however, the completion result of the tool is shared with the student's institution.

Figure 1: Entrance Counseling Tool

Figure 2: Exit Counseling Tool Pie Chart

Figure 2 depicts a visualization provided by the Exit Counseling tool. This picture includes my total loan balance owed to my loan servicer. Within the exit counseling tool, there are no visualizations provided by the tool that relay time to pay off my loan, the interest rate of the loan, or any form of line/bar chart. If I wanted to translate my loan information into a bar chart, line graph, or any other visualiation type, I would have to export their loan data into another tool, such as Excel or an online graph generator.

In a study of a major public university in Massachusetts, it was found that undergraduates had an "alarmingly low level of financial literacy" (Artavanis, 2020). Artavanis also found there was significantly increased financial literacy within STEM and business majoring students when compared to students in other majors, which Artavanis claims "may underscore the role of financial knowledge and numeracy". This deficit in knowledge is then used to show that these students (those who are not STEM or business majors) "are more likely to underestimate future student loan payments" (Artavanis, 2020). With these discrepancies in financial literacy—

especially knowledge surrounding student debt—the status quo set forth by major institutions is not improving debt understanding and can have real-world consequences.

Student debt rests at the intersection of financial quantitative literacy and financial institutional literacy (Nye, 2013). Financial quantitative literacy involves reasoning with quantitative information about a given financial product to make a financial decision. An example of a question that involves quantitative reasoning might involve determining one's time to pay off a student loan with a given interest rate and principal balance. Financial institutional literacy involves understanding "markets, instruments, and institutions" (Nye, 2013). Here, questions regarding the difference between private and public loans, default law, or what college investment accounts exist. The bulk of prior work regarding student debt and financial literacy centers itself on quantitative reasoning, however, having both quantitative literacy and institutional literacy has obvious benefits. The previously mentioned *Entrance Counseling* tool provided by the Department of Education includes education for both domains, however, the tool as established is not enough to bolster the education of students.

Debt Calculations

When weighing a loan's viability, the terms principal, interest, balance, and interest rate are commonplace. In the below photo, the Exit Counseling tool provided by the Federal Department of Education provides explanations of these terms for undergraduates who are nearing the culmination of their time at college.

Figure 3: Exit Counseling Tool Key Terms

Key loan terms are the basis for any calculations regarding a loan—such as the time to payoff a loan and monthly payment amount. Both of these calculations have immediate value for undergraduates and recent graduates—they can help determine one's financial plan for the future as well as identify the viability of different life paths. For instance, if a student only knows they

are taking out $65,000 in student loans, then they might view this number as an ensemble value, when in reality it is the combination of many monthly payments.

Loan Details	
Loan amount	$65,000.00
Annual interest rate	4.99%
Loan period in years	10
Start date of loan	3/19/2023

Loan Summary	
Monthly payment	$689.11
Number of payments	120
Total interest	$17,692.98
Total cost of loan	$82,692.98

Figure 4: Example Loan Amortization Calculation

Without inputting their loan's key values into a calculator, such as the one depicted above, an individual may not know that over a course of ten years, they will be required to pay almost $700 a month, which can a substantial portion of one's monthly income. Moreover, unless a student plugs in their loan's key values into a calculator and understands the relationship between these values—such as how interest rate affects total interest and subsequently the total cost of the loan—they might never come across their key values until it comes time to start paying their dues.

The primary way in which a loan is calculated is through an amortization calculation. The formula includes three key values (*P=Principal, i=interest rate, n=total number of payments*). These values result in *A*, which is the periodic payment value (how much should be payed daily, monthly, or yearly).

$$A = P \frac{i(1 + i)^n}{(1 + i)^n - 1}$$

While pre-formulated calculators do exist, if a student were to calculate their own monthly payment amount or total cost of loan, they would be required to utilize the above calculation. Some students may have seen the amortization formula during a math class in middle or high school, however they likely would not have the formula on hand for a quick calculation about their student loan burden. Thus, they would likely have to review relevant educational material pertaining to loan calculations. Building upon our previous example, when a student borrows $65,000, and they use a ten-year loan term, they must first calculate their *(n)* total number of payments. Over a ten-year period with monthly payments, n = 10 years $\times$ 12 months, or 120. Next, the individual would need to assess their monthly interest rate, which is calculcated by dividing interest rate i by 12. For this example, a 4.99% annual interest rate would result in a monthly interest rate of around 0.00415. After both steps, the individual would be able to plug value into the above formula to receive their monthly loan payment information. Without any visualizations to go along with this calculation, even with the use of a pre-formulated calculator, the complexity of this calculation is intimidating. There are pre-calculation steps to take on before the final calculation can even be processed, and without realizing this, the final monthly payment is extremely skewed and incorrect. In the spirt of openness, I made a mistake while calculating the loan example used above during my first run through of the calculation. I did not realize interest rate needed to be divided by twelve, and it took me a few minutes to figure out why my scratch calculation was far off from the one provided by the calculator. Despite this, there is more to a loan than just a key value. If you know the loan costs you around $700 a month for ten years, that number is just a figment of

your memory. It exists, without context, as a cold and empty piece of text. Without a visualization, the gravitas behind that payment amount might not be realized.

Quantitative Ability

In a study comparing text, volumetric, and linear educational depictions of compounding interest, the text and volumetric tools outperformed the linear tools (Hubbard, 2016). This study identified that linear tools might be 'cognitively taxing' and that the text and volumetric tools may have worked better at challenging prior misconceptions. The study notes the importance of tackling misconceptions as they create barriers for quantitative analysis. Importantly, this work provides considerable support for the explorations within my thesis by describing that "experimental evaluation of the impact of different pedagogies yields interesting and informative results for both theory and practice" and that it is important to study "whether the efficacy of the tools [visualizations] extends to the realm of debt literacy."

Expanding the imperative behind this thesis, "the vast majority of the possibilities in the design space will be ineffective for any specific usage context. In some cases, a possible design is a poor match with the properties of the human perceptual and cognitive systems" (Munzner, 2015).

The graphs most used to display financial information are line and bar graphs, yet as identified Hubbard and Nye, these displays require high levels of numeracy and quantitative literacy. The translation of these abilities and the application of financial topics can be quite challenging for those who are less experienced in these areas and prove that not everyone "profits from standard visual displays" (Galesic, 2010). Interestingly, the individuals who use standard visual displays daily, such as people employed in a financial or STEM field, are likely the same individuals crafting visual displays for public consumption. Therefore, a gap in experience and understanding exists between those who are creating the visualizations and the public.

Grounded Theory

This paper builds its analytic procedure through the grounded theory process derived from Barke Et Al. This process builds upon an evaluation of qualitative data by forgoing a predetermined hypothesis and instead building a post analytic theory. This theory takes the form of codes and themes, which are founded based on participant statements. This process is applied by identifying themes and trends in association with participant verbal responses to questions in a semi-structured interview that covers their demographic information, perceptions of student debt, and analysis of debt visualizations. Throughout the course of the interviews, trends appeared, which identified two clear patterns for the interviews. For participants whose demographic information (year, major, experience with debt, experience with graphs) was profound, their interviews often occurred the most quickly and participants needed the least prompting or educational assistance for an interpretation of a visualization. For participants whose demographic information indicated the opposite, the interviews commonly took the longest due to a need for additional explanations of debt or visualization concepts. While both patterns were important in determining how undergraduate students interpret debt visualizations and perceive student debt, it was through the participants who took the longest to explore the graph were the most critical gaps in debt knowledge revealed and the most crucial components of debt visualizations identified.

Chapter 3

Understanding The Perceptions of Undergraduate Student Debt

The aim of this thesis is to enhance our understandings of what makes an effective data visualization for student debt data. To do this, I explore student's prior understanding of student debt, their own debt situations, and their ability to analyze visualizations designed around debt. Within this methods section, I will explore research philosophy, participant recruitment and demographics, interview question design, interview procedure, apparatus and stimuli, and data analysis.

Method

This thesis will follow the interpretivist research philosophy. The core of this thesis is supported by an IRB-approved interview study. This study aims to understand the unique experiences of each of these participants that is leading them towards a specific understanding of the visualizations presented in the study. For instance, some individuals may not have had prior financial literacy education or may have had a benefactor take out their loans for them, and this thesis aims to explore how these relationships are linked to one's understanding of a debt visualization. Moreover, this thesis aims to ask how these debt visualizations can be improved for greater educational quality and accessibility. The study takes its form through **twelve semi-structured interviews** that focused on Penn State – University Park undergraduates.

Participants

Participants were recruited through a combination of convenience sampling and snowball recruitment (otherwise known as word-of-mouth recruitment). The only criterion for recruitment was active undergraduate status at The Pennsylvania State University – University Park. Besides this requirement, participants could come from any major, have or not have student loans, or be any year-level.

The intended sample size for the study was twelve participants, which was achieved. This sample size was determined such that it provided breadth for age and educational diversity yet prevent resource constraints. Prioritizing age and educational diversity was such that quantitative and financial ability could remain varied between the participants. As described in the Chapter 2, Hubbard and Nye both suggest that STEM and business majors have greater ability to grasp financial concepts. To ensure potentially extrapolatable results, this study aimed to contrast differences in educational background. Participants were not compensated for their time.

Participant #	Undergraduate College	Upper/Under Class Status	Loans (Y/N)	Parent 1 Bachelor's Obtained (Y/N)	Parent 2 Bachelor's Obtained (Y/N)
P1	Engineering	Underclass	No	Yes	No
P2	Human Health and Development	Upperclass	No	Yes	Yes
P3	Business	Upperclass	Yes	No	Yes
P4	Communications	Underclass	No	No	No
P5	Education	Underclass	Yes	Yes	Yes
P6	Business	Underclass	No	Yes	No
P7	Engineering	Upperclass	No	Yes	Yes
P8	Liberal Arts	Underclass	Yes	No	No
P9	Business	Upperclass	No	No	Yes
P10	Business	Underclass	Yes	Yes	Yes
P11	Engineering	Upperclass	Yes	No	Yes
P12	Engineering	Upperclass	No	Yes	Yes

Table 1. Participant Demographics

Participants are classified by their academic college, rather than major, to help protect participant identity. Moreover, participants are classified by their status as either upperclass (third-

year or fourth-year) or underclass (first-year or second-year). By grade-level divisions, the interview pool was represented with half the pool being underclass and upperclass. Just under half of the participants (N=5) had taken out student loans, while just over half of the participants (N=7) had not taken out any student loans. Approximately one-third of the interview pool was an engineering major, while another third were business majors. The remaining third included education, liberal arts, and communication majors.

The Fisher's Exact Test is used to determine if there is a correlation between STEM/Business/Other majors and loan status. Due to the 2x2 dimensional limitation of the Fisher's Exact Test, I grouped together STEM and Business majors together for the purposes of the calculation and the documented relationship of these groups in the literature review regarding debt literacy and numeracy. The Fisher's Exact Test was calculated with Python's SciPy library.
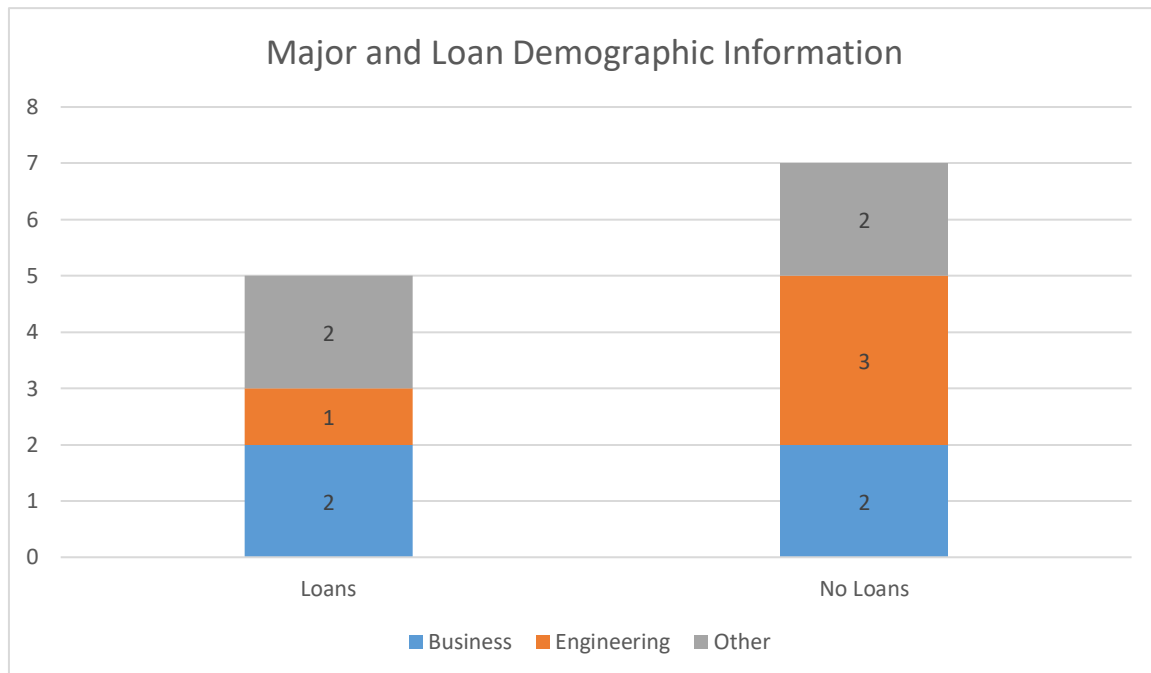
Figure 5. Major/Loan Demographic Information

With a p-value of 1 and a significance level of 0.05, the result does not show a significant relationship between loan status and major within the current sample pool.

Rather than using heteronormative parental terms (mother and father), I chose to use Parent 1 and Parent 2 as the parental designation. This study does not pursue gender as a potential signifier to loan status or ability in interpreting graphs, and subsequently, I decided to extend this decision to parental demographic information. Since the below table is too small for Chi-Square Analysis and has too many fields for a Fisher's Test, I consolidated the table to include loans and no loans as well as any parent has a degree vs. no parent has a degree. A Fisher's Test is performed on the combined data.

	Parent 1 Bachelor's Degree	Parent 2 Bachelor's Degree	Parent 1 No Bachelor's Degree	Parent 2 No Bachelor's Degree	Row Totals
Student Loan	3	4	3	2	12
No Student Loan	4	4	2	2	12
Column Totals	7	8	5	4	24

Table 2. Loan Status and Parental Education
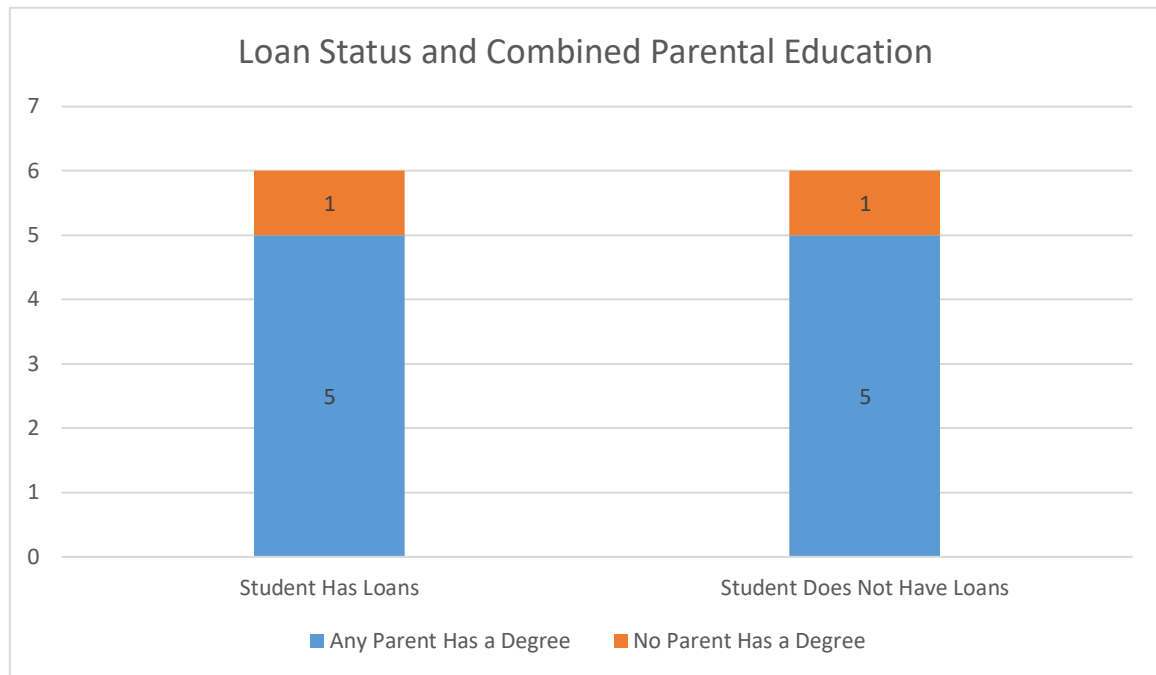
Figure 6. Loan Status and Combined Parental Education

With a p-value of 1 and a significance level of 0.05, the correlation between student loan status and parental bachelor's degrees within the sample pool is *not significant.* Despite this, abundant research shows that parents who have had higher education do tend to make larger incomes, thus allowing their children a more debt-free education (Walsemann 2017). However, this conclusion cannot be drawn from the sample presented within this study due to the small sample size.

Procedure

Interview questions are divided into four categories. While the interviews were conducted in a semi-structured format—meaning that the structure of the interview was flexible and allowed for clarifying questions—the overarching categories remained throughout all interviews. First, I collected background and demographical information from the participants. Information such as their educational background, loan history, and post-graduation plans was included in this section.

These pieces of information can indicate how one's own education and loan history might impact their understanding of the debt visualizations. Second, I collected information about the participants' financial institutional literacy. This information included questions surrounding one's understanding of loan types and decision processes related to taking out a loan. This information can indicate if the participants' knowledge of student debt was specific to their own experiences or if their knowledge was more global and hence more financially institutionally literate. Third, I collected information regarding participants' experience with data visualizations to indicate prior experience level using or creating visualizations. I also collected information regarding prior experience with financially related visualizations. Fourth, I asked participants to describe their understanding of three visualizations related to debt. Participants were asked questions that would require them to describe the content of the graph, how the design impacted their understanding, and how this type of visualization might impact their financial decision-making process. This information was collected to assess how and why undergraduates perceive debt visualizations and how these visualizations can be improved to be more inclusive. This interview protocol was repeated for all twelve participants.

Each interview was recorded using the Zoom recording feature. For the analysis of data, the Zoom-produced transcript was reviewed and annotated. Demographical lines of questioning were included in these transcripts, however the answers to these questions were also typed in a document for later review by the researcher. Graphical analysis lines of questioning were only recorded for the resulting transcript.

Participants have varied educational backgrounds, interests, and experiences. In practice, this educational background divide is not accounted for when students apply for loans and is not accounted for when students must determine how and when they should pay those loans back. For

some people with prior financial education and strong quantitative ability, continued education on the topic of student loans may be redundant. For others, continued education might mean the difference between paying that loan off early or accruing thousands of dollars of interest. When interviewing participants, it rapidly became clear that some individuals had a much better grasp on student loans and debt than others and that this limitation extended to their ability to interpret the debt visualizations. To ensure that each interview could result in relevant, interesting, and nuanced data, I would, when needed, provide participants with an oral education presentation of background information on student debt.

This presentation was tailored to fit the educational gap of the participant. While a scripted presentation would, in theory, ensure that every participant had the same background knowledge, the shift in formality—from a conversational semi-structured interview tone to the reading of instructional material—was deemed inappropriate for this study. Moreover, the tailoring of the instruction to everyone mitigated the risk of didactic teaching, which could create unwanted feelings of hurt for the participants. The choice to provide oral instruction solely, rather than include a pamphlet or similar material, was so that the participants could refer to the existing visualizations in the interview material to enhance their understanding of debt.

Apparatus and Stimuli

The most used visualization designs for loan information are line and bar graphs. I will explore how and why these designs perform and fail to perform through an analysis of student perceptions of these graphs and how those perceptions relate to student debt.

Constant Loan Payment Over Time

Figure 7. Stimulus 1

Figure 7 identifies the first stimulus that will be presented to participants. This graph design was chosen for its relative simplicity. Most individuals, regardless of graphical literacy, would likely have interacted with a graph of this sort during their tenure at Penn State or in high school. Thus, this graph is used as a baseline to assess graphical literacy for the participants. Moreover, as this graph is lacking a legend, differing colors, a description, or any form of marks or interactivity, this graph offers participants some of the most robust opportunities to expand on what graphs they prefer and why they prefer them. This graph was designed by the author in a Python Jupyter notebook using the Plotly graphical library. Besides the title of the graph, there is nothing else relating this graph to student debt. Responses to this graph will be assessed to further assess debt literacy and see if students are able to apply extraneous concepts and relevant knowledge.
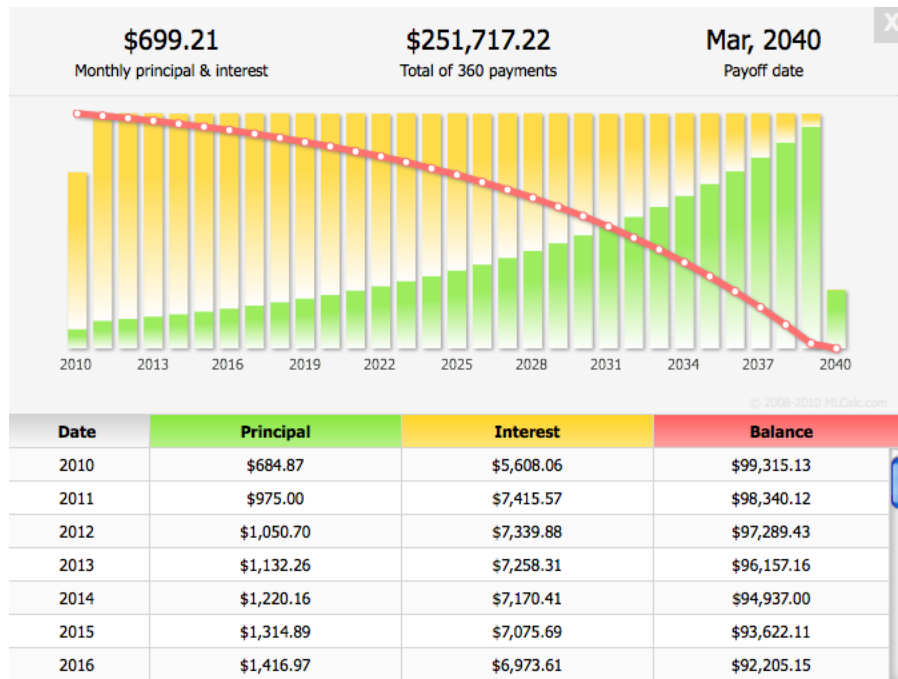
Figure 8. Stimulus 2

Figure 8 identifies the second stimulus that will be shown to participants. This graph is static, meaning that the numbers on the page are not able to be manipulated. This type of graph is similar to the types of graphs that would be presented if one were to use either a loan payment tool or graphical tool to map their loan data. The three key elements of a loan are included in this graph, which are referenced in detail within Chapter 2. One's knowledge of these elements, and their ability to apply that knowledge to their perceptions of the graph will provide critical insight into their overall understanding of debt. As a higher-level graph than the first stimulus, this graph includes multiple marks, tabular data, colors, and summary information. Participants will be tasked

with notating the improvements or alterations they believe will increase the usefulness of this graph, which may present itself as a more difficult task when contrasted with the first stimulus.

Simple Loan Calculator

Loan Details	
Loan amount	$20,000.00
Annual interest rate	5.50%
Loan period in years	1
Start date of loan	2/18/2023

Loan Summary	
Monthly payment	$1,716.74
Number of payments	12
Total interest	$600.83
Total cost of loan	$20,600.83

Legend: Principal Paid, Interest Paid, Loan Balance
Chart axis: $14,000.00, $12,000.00, $10,000.00, $8,000.00, $6,000.00, $4,000.00, $2,000.00, $0.00
Years: 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029

Pmt No.	Payment Date	Beginning Balance	Payment	Principal	Interest	Ending Balance
1	3/18/2023	$20,000.00	$1,716.74	$1,625.07	$91.67	$18,374.93
2	4/18/2023	$18,374.93	$1,716.74	$1,632.52	$84.22	$16,742.41
3	5/18/2023	$16,742.41	$1,716.74	$1,640.00	$76.74	$15,102.41
4	6/18/2023	$15,102.41	$1,716.74	$1,647.52	$69.22	$13,454.90
5	7/18/2023	$13,454.90	$1,716.74	$1,655.07	$61.67	$11,799.83
6	8/18/2023	$11,799.83	$1,716.74	$1,662.65	$54.08	$10,137.18
7	9/18/2023	$10,137.18	$1,716.74	$1,670.27	$46.46	$8,466.90
8	10/18/2023	$8,466.90	$1,716.74	$1,677.93	$38.81	$6,788.97
9	11/18/2023	$6,788.97	$1,716.74	$1,685.62	$31.12	$5,103.36
10	12/18/2023	$5,103.36	$1,716.74	$1,693.35	$23.39	$3,410.01
11	1/18/2024	$3,410.01	$1,716.74	$1,701.11	$15.63	$1,708.90
12	2/18/2024	$1,708.90	$1,716.74	$1,708.90	$7.83	$0.00

Figure 9. Stimulus 3

Figure 9 is on the surface quite like **figure 8**, however this figure links directly to an Excel workbook provided as a loan calculator template by Microsoft. Participants will have the opportunity to manipulate values within the workbook, allowing them to 'learn as they go', providing insights into the thought process of the participants. Moreover, I will be able to learn how interactivity within a visualization impacts one's understanding of debt. Additionally, the

contrast between having no interactivity and interactivity will be crucial in understanding the importance of interactive components within a visualization.

Data Analysis

This study was conducted through an inductive approach with the goal of exploring collected qualitative data to learn what, how, and why individuals interpret debt visualizations and how these visualizations can be improved.

Interviews were processed with the initial coding approach to qualitative coding. This approach was derived from Johnny Saldaña's 2009 working titled *The Coding Manual for Qualitative Researchers.* Each interview was analyzed for critical themes and unique comments. Repeated themes were labeled, and nuanced comments, which were specific to each participant, were highlighted to provide context to these higher-level themes. These items were analyzed to find commonly held positive and negative beliefs by the participants, as well as information that could help explain the experiences of individuals.

Information about prior financial education or debt literacy is crucial to assess the first research question within this paper, while analytical questions will be critical for assessing the second and third research questions.

Debt literacy was assessed through a series of pre-test questions that asked participants about their prior experience with student loans, their knowledge of student loan types, and their experiences with prior financial literacy education. Graphical literacy was assessed through a series of pre-test questions that asked participants about their prior experience with using and making visualizations as well as their familiarity with financial visualizations.

Chapter 4

Findings and Results

This section will analyze the results of the conducted study through significance tests between potential relationships within the sample data as well as a qualitative review of participant statements and sentiments. This chapter will be broken down by each of three research questions and followed up by an observation and reflection of the human factors that may have contributed to participant perceptions and visualization evaluation.

RQ1: *"Is there a gap in undergraduate knowledge of student debt, and if so, what is this gap"?*

Participant #	Financial Ed. (Y/N)	General Knowledge of Student Loans (Y/N)	Knowledge of Student Loan Type (Y/N)	Knowledge of Federal Dept. of Education's Involvement(Y/N)	Experience Creating Visualizations (Y/N)	Experience Using Visualizations (Y/N)	Debt Literacy (High, Medium, Low)	Graphical Literacy (High, Medium, Low)
P1	Yes	Yes	No	No	Yes	Yes	Medium	Medium
P2	No	Yes	No	No	Yes	Yes	Low	Medium
P3	Yes	Yes	Yes	No	Yes	Yes	High	High
P4	No	Yes	No	No	No	No	Low	Low
P5	No	No	No	No	No	Yes	Low	Low
P6	Yes	Yes	Yes	No	Yes	Yes	High	High
P7	Yes	Yes	Yes	No	Yes	Yes	Medium	Medium
P8	No	Yes	No	No	Yes	Yes	Medium	Medium
P9	Yes	Yes	Yes	No	Yes	Yes	High	High
P10	Yes	Yes	Yes	No	Yes	Yes	High	Medium
P11	No	Yes	Yes	No	Yes	Yes	Medium	Medium
P12	No	Yes	No	No	Yes	Yes	Low	Medium

Table 3. Participant Knowledge and Experiences

The first research question presented in this paper aimed to assess if there was a gap in undergraduate knowledge of student debt. In the above table, answers from the interview

participants were grouped into six categories that relate to their experiences with financial education, visualizations, and student debt. Half of the participants had not had any financial educational classroom experience that they could recall. Most of the participants had either created a visualization or used a visualization in the past. Some participants asked if visualizations in statistics class counted, as this was where their experience primarily rested. Half of the participants were not able to describe the difference between a private and a public student loan. No participants were able to accurately describe the Federal Department of Education's involvement in the student loan industry.

Derived by both explicit participant responses and interviewee experiences during the interview, Debt Literacy and Graphical Literacy were assessed. If participants had mentioned they had little to no experience with debt or visualizations or they were not able to interpret the visualizations without external stimuli (interviewer explanations), they were given a 'Low' classification. The 'Medium' and 'High' classifications were reserved for when participants mentioned greater experience, or they were able to interpret the visualizations with little to no external stimuli. One out of every three of participants had either low, medium, or high debt literacy. Only one participant had low graphical literacy, more than half of the participants (N=7) participants had medium graphical literacy, and only two participants had high graphical literacy. Neither the Chi-Square nor the Fisher's Exact test are viable options to assess relationships within this data. Chi-Square demands larger field values, while Fisher's Exact demands a 2x2 matrix. Despite this, I will be grouping medium and high graphical literacy to allow for a Fisher's Exact test to occur, in the event a relationship can be found within these groupings. This grouping will occur due to the reasonings behind these distinctions: low graphical literacy was derived through participants having very limited graphical literacy, while participants with medium and high

graphical literacy had much greater levels of understanding. The same protocol will apply to the

debt literacy tables below.

	Low Graphical Literacy	Medium Graphical Literacy	High Graphical Literacy	Row Totals
Experience Creating Visualizations	0	7	3	10
No Experience Creating Visualizations	2	0	0	2
Column Totals	2	7	3	12

Table 4. Experience Creating Visualizations and Graphical Literacy

The Fisher's Exact task for the above graph identified a p-value of 1.0, which is above the

designated significance level of 0.10, showing that there is no significant relationship between

experience creating visualizations and graphical literacy within the sample data.

	Low Graphical Literacy	Medium Graphical Literacy	High Graphical Literacy	Row Totals
Experience Using Visualizations	1	7	3	11
No Experience Using Visualizations	1	0	0	1
Column Totals	2	7	3	12

Table 5. Experience Using Graphs and Graphical Literacy

The above Fisher's Test results in a p-value of 0.166, which is above the 0.10 significance

level. While this does not support a strong relationship within the data, it may show that at least a

weak relationship is present between experience using visualizations and graphical literacy.

	Low Debt Literacy	Medium Debt Literacy	High Debt Literacy	Row Totals
Has Financial Education	0	2	4	6
Has No Financial Education	4	2	0	6
Column Totals	4	4	4	12

Table 6. Financial Education and Debt Literacy

With a p-value of 0.002 and a significance level of 0.10, there is a clear relationship between debt literacy and financial education present within the data when tested under a Fisher's Exact Test.

	Low Debt Literacy	Medium Debt Literacy	High Debt Literacy	Row Totals
Has General Knowledge of Student Loans	3	4	4	11
Has No General Knowledge of Student Loans	1	0	0	1
Column Totals	4	4	4	12

Table 7. General Student Loan Knowledge and Debt Literacy

In the above Fisher's Test, the resulting p-value is 0.36, which is above the significance level of 0.10, but does identify a very weak relationship between general knowledge of student loans and debt literacy within the data.

With all this mind, we can see that there is a relationship between knowledge of student loans and debt literacy, as well as a relationship between financial education and debt literacy. There was no relationship present between experience creating or using a data visualization and graphical literacy. However, with the weak relationship present in knowledge of student loans and financial education, we can assess at the very least, there is a gap in undergraduate knowledge of student debt. This gap existed most prominently in participant ability to describe loan type, as well

as describe the involvement of the Federal Department of Education for student loans. When paired with the varied level of prior financial education, this combination identified that most crucially, the participants' gap rested in their ability to explain the key values of a loan, how to interpret these values on a visualization, and how a loan exists as a financial product in both the public and private sectors.

RQ2: *"What aspects of current visualization design for student debt situations are most important for enhancing one's understanding of debt"?*

The second research question presented in this paper aimed to assess what aspects of current debt visualizations are important in enhancing understandings of student debt. For this question, key themes emerged. These themes were derived from questions pertaining to the visualization testing portion of the interview sequence.

Themes (Current Debt Visualizations)
Color Associations
Core Data Relationships
Accessible Organization
Key Value Portrayal

Color Associations proved to be one of the most effective tools for enhancing understanding among the participant pool. Participants most associated the color 'green' with money, which they found put them in the proper mindset to assess loans. This was described by participants stating that:

"Green was a decent choice. So … at first you're [going to] think of the

principle as … Green. When I think Green, I think money, right?" (P3)

Another participant followed up this same sentiment, however built upon the relationship between color and money further by integrating interest into the discussion.

"I think in modern depiction of … money … green and yellow and red

also is all … used, but Green, you know, sometimes refers to how much

you have. Red refers to like, how much you're losing, or you're

expensing. The interest or yellow is like … no change or very minimal

change, or something of like inches accumulating". (P6)

Besides just the association of color to monetary concepts, participants identified how the color choice made them feel, which included feelings of cautiousness and negativity. Emotions of happiness were reserved for stimuli 3's characters in the upper right-hand corner. One participant mentioned the following:

"A lot of times talking about money, especially when it comes to … loans

and student debt, it's [kind of] … a taboo topic among some people…

So, I guess the green yeah kind of make me a little bit cautious". (P1)

The same participant stated that "yellow" made them cautious:

"Yellow it … immediately made me … more cautious, especially when

it's talking about … interest and being that, when you're looking at this

like graph or … the calculator it doesn't, make me as cautious, mainly

because it's more of a like a lighter green, even though it's still talking

about like interest in general". (P1)

In P7's account, red is viewed negatively:

"I guess the fact that the balance line, [is] red is kind of it makes me think

it's more like a negative than it is." (P7)

In combination with the coloration of the visualizations, P2 offered a different perspective by discussing how playfulness and characterization of the visualization made them feel towards a given graph:

"I don't know. I think the first thing that I thought of was that it's very …

visually … appealing … it's more fun... I definitely think people up in the

corner … the little like illustration." (P2)

While main color groups did prove useful to the participants in identifying key values on the visualization, the shading of individual colors was not as preferred. P7 suggested not to have *"interest and principal the same … like both of them as green"*, while P11 suggested that they *"would add more contrasting colors rather than just two shades of green".*

Outside of perceptions of color, participants found consistent usage of color was a beneficial component of the visualization design:

"I really like how … the colors on the legend match really well with the graph, and especially with the red balance line, it's really easy to kind of see how that changes". (P9)

This notion was shared by other participants and was identified as improving the approachability of the visualization:

"It's like as soon as I see it [legend], I already have an association. I don't feel like I have to go searching for it … so I almost know what to expect even before I've looked at the graph. So, I just think in general it makes it a lot more approachable". (P10)

Core Data Relationships were also found to be a useful component of the current visualization regime. Within the second and third visualizations, a table of data is presented. While the tables by themselves was not described as valuable, when they were combined with a visualization, the participants found that the table did enhance their understanding:

"The chart on the bottom with … the year, … the date, principal, interest, [and] balance. I like to see direct relationships like that… I like it because it kind of answers my questions before I ask [about] the loan

details of summary. It also has on the bottom … the same balance,

principal, [and] interest that was in the last graph that I liked." (P8)

While this preference wasn't completely universal with P5 claiming that it's *"very confusing, there's a lot at once"*, there was more alignment with how P4 stated *"the [table and graph], [they] need each other"*.

Accessible Organization also became pronounced through the interviews. This theme was defined through notions of general "cleanness" of the visualizations, as well as notions of clarity. For instance, when axes went unlabeled or currency was not designated, participants were quick to point out these deficiencies. P12 notably stated that without *"labeling the X and Y axes, [they're] not really sure what this graph [was] trying to depict"*. This theme was consistent across most participants, with P8 stating that in the third visualization, they thought *"the graphics are just better. To me it looks more organized".* Through the interviews, it became clear that clearly labeled graphs, organized tables, and certain designs of graphical marks was paramount. These graphical marks—which are perceived as lines and bar chart components on the visualizations—were often a point of critical confusion for the participants. While some participants called for separating stacked bar charts into multiple individual bar graphs, most participants found that cumulative bar charts (those that represent a sum of values) rather than proportion-based bar charts (those that represent two values as a proportion of a value) were significantly easier to interpret. P10 most clearly identified this notion by saying *"This makes a lot more sense to me, because I see interest … cumulatively".*

Key value portrayal also plays into accessibility of visualization design. Within this study, key value portrayal describes how key numerical values are identified and represented on a graph.

For instance, this might refer to interest rate, loan balance, payoff date, or monthly payment amount. Participants found that these values most rapidly helped them understand the visualizations, and more than that, had the greatest impact on how they applied the graphs to their potential future actions. "*The monthly principal, interest, the pay off, and … the total 360 payments … stand out,*" and later spoke on the fact that their future actions might be *"looking into a different major… or apply for scholarships".* (P5) This theme was further amplified through participant statements on how key value portrayal set up their understanding of the visualization:

> *"I'm looking at … the total 360 payments. I feel like everything at the top*
> *is kind of … prepping me for what I'm about to see is … real nice. Those*
> *are … the most digestible numbers on the graph, and they're placed at*
> *the top".* (P10)

key value portrayal also had the ability to create an emotional impulse from the participants. P1 found that *"it's a little intimidating … 20 years away. So, I think it would maybe put into perspective how well and that can affect you".* The evident display of important pieces of information can impact graphical understanding as well as debt perceptions. In Chapter 2, the Exit Counseling tool is shown to provide information regarding key values. **Participants brought up no knowledge of these key values from their time interacting with the entrance or exit loan counseling tools.**

Within current visualization design, **Color Associations** provided participants with the ability to assess the visualization and apply their knowledge of outside concepts, however varied associations between the graphs—which is indicative of inconsistent color schemes the design

regime as a whole—increased the confusion of participants. **Core Data Relationships** relate to how participants relate tabular data to a visualization, and within the current visualization regime, participants found that when combined with a visualization, tabular data can improve their understanding of debt. **Accessible Organization** relates to how well put together a graph is—including a readable legend, clean configurations, and inclusive mark design. Participants found that when graphs included a well-organized legend, did not overwhelm them with too much information, and could separate meaning between marks, the graph did provide valuable information. However, when these instances were violated, participants found that visualizations were more detrimental to their understanding. **Key Value Portrayal** builds upon the themes of Core Data Relationships and Accessible Organization by describing how participants enjoyed how a loan's key values were displayed prominently, as it prepared them for their analysis of the visualizations.

    RQ3: *"What are potential improvements to current data representations that can improve undergraduate understandings of student debt"?*

    This section will answer the third research question presented in this thesis as well as cover guidelines for further development of student debt visualizations. Two themes emerged that help to answer the third research question.

Visualization Improvement Themes
Enhanced Interactivity
Context-Specific Representations

Improvements exists in a two-fold relationship: ability to expand on information within a visualization and implementation of user-specific data. First, for **Enhanced Interactivity**, participants across the board recommended the ability to hover over data and have this expand available information.

> *"If you were to hover over the line, you would see the exact values at the line. I know sometimes that's really important and kind of useful when looking at ... a line graph to see ... where the actual values are, because sometimes it's hard to tell. And I'd say definitely, being able to click on ... any of the any of the axis' and getting a ... description popping up would definitely be more clear for me trying to understand the graph."*
> (P11)

A vast majority of participants (N=10) shared this sentiment. **Enhanced Interactivity** also involved an accessibility component: participants found that the ability to have a sliding feature and a dropdown menu to select values was important.

> *"A sliding scale where you can see the graph ... You can go up and down that line ... I think, having labels above every little point with a dollar amount would look very clustered. But I think if we add a sliding scale, I think then the user could easily see the exact dollar amount they want."* (P3)

This participant built upon the concept of accessibility by desiring for an up-down value changer within the calculator visualization to reduce user error.

"Having like an up and down [value changer] [could limit] error [for] [increasing] the loan, so you don't accidentally type 10%, and get scared". (P3)

Secondly, **Context-Specific Representations** showcased a desire for participants to include information relevant to them. Often, this involved participants wanting to include their own loan values within the calculator presented in the third visualization.

"I'm not just looking at ... an example of a loan. I can apply my actual loan details or my future loan details ... It gives me ... a good guide as to ... how long each loan would take to payoff. And I think that's ... a really useful tool to have before, and even during like looking into a loan or getting a loan. I think if I were to look at many different loans. I think I would definitely use this tool to ... put in the details and see how long it would take me to pay off, and how much interest it's accruing, and I think I would probably use a tool like this to put in every single possible loan that I [am seeing] data and put my own loan amount into it, and see ... how long it would take. And then I would use that as the basis of my decision to ... go with which loan I want to actually pick". (P11)

With these themes in mind, we can begin to describe improvements to the current graph regime. We know that for the theme of **Color Associations** having relationships that tie into the legend as well as societally recognized symbology is important. Thus, future visualizations might want to include colors that represent money (green). The use of other colors should be used such that they have coexist and build upon understanding. For instance, participants suggested that blue represented decreasing values, as did red. Including both colors, on the same graph, or in the same information set, might prove confusing for rapid color associations. Color was presented as a key driver of understanding, and thus the choice of color is paramount to enhanced understanding. **Core Data Relationships** showcases the data-centric drive within the interview pool. Participants found that while tabular data or numerical information by themselves was of little value, when that data was combined with a visualization, they could explore the visualization further. We should encourage further exploration, and this combines with the theme of **Context-Specific Representations** as it allows for individuals to explore their own potential loan data in-depth. **Accessible Organization** targets the status quo of current visualization design. Debt visualizations are often not used in the day-to-day proceedings of individuals, let alone undergraduate students. Thus, having visualization designs that are difficult to read, whether that be based on poor labeling, niche use of marks, or just rapidly compiled visualizations, should be avoided. Students may not have had prior education on how to read these graphs, or even on the loans they may be taking out. Visualizations that immediately dissuade them from further exploration into their financial future are a detriment to these students. Students found that **Critical Value Portrayal** helped to promote an accessible understanding of the graphs. They noted that these values are the pieces of information that stood out the most to them. Likely, these values are the first thing a student would notice when looking at a visualization, and they should be used to drive their attention to the key

factors of their loan. For instance, while a 5% interest rate might not seem like a large difference to an 8% interest rate, the total interest accumulation may be a very large difference. Thus, highlighting pieces of information that can impose greater financial hardships down the line should be an absolute priority. Participants found that interactivity was a key desire for advancing the current visualization regime. For some, this interactivity involved hovering over pieces of information and getting greater context, while for others it involved the ability to directly manipulate the graph (rather than typing numbers into a calculator) to see how different loan balances and interest impacted how the visualization appeared. Interactivity expands upon the exploration of debt for students—it allows them to compare loans, compare time to pay down their loan, and contrast this with their interests and behaviors. Perhaps this takes the form of a change in major or a focus on applying for more scholarships, but despite the form of this future behavior, if graphs can be designed to promote a greater exploration of opportunities, students may have the opportunity to see the impact of their loans in a greater light.

Observations and Reflection

As a qualitative study, it is important to assess how participants react during interviews. This reaction can further teach us how individuals perceive visualizations and student debt, so that we might better be in touch with individual realities.

While not universal, I could tell that certain participants were withdrawing from the interview as time went on. This would take the format of one worded or shortly phrased responses. P4, P9, and P12 presented the most with these behaviors of withdraw from the interview.

Besides this, we can break the human factors presented within the interview into four categories, which are presented below.

Human Factor Categories
Reasons For Taking Loans
Parental Situation
Voluntary Ignorance
Loan Conscientiousness

Within this study, participants described that they took out loans out of necessity, or that the decision to take our student loans was not wholly their own decision. Thus, **the reasons for taking on loans** were typically something that had to be done, rather than something that was a want. This finding does make logical sense, as when given the choice between having a debt-free education and a debt-ridden one, most rational individuals would choose to have the debt-free education.

"It wouldn't have affected it only because it was kind of a necessity to take them out and it was the best option. Obviously, no one likes taking debt on… [and there's] good debt and bad debt and not that student loan is good debt, but I think the Federal loans that I took were the best version of get I could have taken." (P3)

One participant's father forced their hand into taking out a loan, as described by the participant:

"Taking a loan wasn't. Really my idea. I didn't really want to, so this

[visualization] probably would have reinforced that." (P8)

Parental situations also played a hand in taking out loans, albeit with a weak relationship.

	Either Parent Has a Bachelor's Degree	No Parent has a Bachelor's Degree	Row Totals
Loan	4	1	5
No Loan	6	1	7
Column Totals	10	2	12

Table 8: Chi-Square Test and Loan Status with Either Parent's Education Status

Using a Chi-Square test, we can obtain a value of 0.069, which is slightly larger than our significance level of 0.05. Thus, while there might not be a strong relationship between parental degree status and loan status within this dataset, there is at least a weak relationship which might suggest that parents with degrees are more financially well off and have experience navigating student loans or institutions of higher education. Despite this, in the same Fisher's Exact Test performed in Chapter 3, there is no evidence of a relationship.

Participants with loans commonly had behaviors that associated with the **Voluntary Ignorance.** This behavior, which is demonstrated through the computational term Ostrich Algorithm, is the practice of ignoring problems to deal with later (Wikipedia 2023). This notion was often identified by individuals claiming that they had not often viewed their loan debt summary, which is a tool provided by the Pennsylvania State University to offer students the opportunity to check on their current debt situation.

"Because of the current environment, where there is no payments ...

[they've] been paused ... since COVID ... I haven't looked at it a lot. I

feel like I would have looked at it more had that not been the case, but

I'm starting to look at it now even more just because graduation is

getting closer. But throughout college I rarely looked at it." (P3)

With the participants in mind, the idea of Voluntary Ignorance was not out of fear of the loans, but out of lack of alternatives to the loans. If debt was necessary to obtain an education, the participants, given the pause in interest due to COVID relief, may have found that reviewing their debt consistently during their undergraduate tenure to be a repetitive and unnecessary task. There was a notable exception to this behavior, with P10 explaining how they took meticulous spreadsheet recordings of their current loans; however, this person was an outlier. This individual did have High Debt Literacy.

Participants were **conscientious about their loans** or lack thereof. Conscientiousness looks to see if participants are aware about their own student loan situation's impact on their own finances or those of other people, whether they have loans or not.

"It is pretty much … a privilege that I also don't know this, and I don't

have to … pay … much more like other people so … I recognize my

privilege, but I would like to be more educated on it." (P2)

This conscientiousness was a common notion, and was cleanly expanded upon by another participant:

"A lot of people pay for college and that people have a very tough time

paying it back, because they don't necessarily get as much money coming

out of college as they expect, or they end up having to go to more school

and taking on more debt." (P9)

The above two quotes describe participants that are aware of how loans impact other people, even though they themselves don't have any student loans. Participants with loans noted how these loans impact their potential future actions:

"Maybe like looking into a different major or … maybe apply for scholarships." (P5)

We can see that participants are aware of student debt's impact on other individuals and themselves, and even though it might not be top of mind for this participant pool, debt is something they are accounting for.

With these themes in mind, design guidelines began to present themselves. These guidelines involve a combination of multiple themes.

One of the foremost guidelines was **Color Standardization.** The colors presented within the visualizations of this study offered participants the ability to rapidly associate values on the graphs. Participants found that the colors presented drove their associations. For instance, green tied in an association to money, while blue and red provided the perception of a decreasing value. Thus, while green did provide participants with a notion that the graphs involved monetary values, green itself as a color choice for marks on the visualizations was not as effective. Standardized color schemes are important as they would help students understand visualization from multiple mediums, rather than having to 'relearn' how to interpret a visualization each time they come

across a new one. Therefore, having consistent coloration within a visualization and perhaps within the domain of student debt is paramount to increase comfortability with debt visualizations and enable ease of access.

The next guideline that presented itself was **Thorough Organization**. Participants found that when graphs went unlabeled, or the legend was not easily accessible, or when the graph was cluttered (proportional bar charts contrasted to cumulative stacked bar charts), they had a more difficult time digesting the visualization. Visualizations should prioritize the ability to rapidly read these graphs and the ability to easily locate critical pieces of information that promote understanding the visualization. Graphs should have values that are easily identifiable, and that when a visualization is produced, it should prioritize a 'simple and clean' look, as simplicity reigned king throughout this study.

Participants also often requested informational materials. Thus, **access to educational materials** within the visualizations is also an important guideline. Through this study it became clear that many individuals lacked knowledge of student debt. This involved not understanding how key values such as interest, balance, and principle were related. Offering education on these concepts within a visualization can be helpful as participants noted that their knowledge of these concepts improved following the study. While the education I provided was based on request of the participants and to ensure they could attempt an interpretation of a visualization, asynchronous education on a graph could be a longer-term solution to promoting greater awareness of student debt. This might involve offering a popup explanation of how interest works or even a quick video tutorial on an amortization calculation. Despite the format, further educational opportunity needs to be included in the design of debt visualizations.

In the age of smartphones and JavaScript enabled websites, interactivity within visualizations is commonplace today. Despite this, **interactivity driven design** should be ensured within all debt visualizations. Participants consistently requested interactive features. Most often, this involved manipulating loan values, hovering over visualization components, or having ways to specify the graph for a specific timeframe. Visualization design should be highly interactive as this allows for students to further their exploration of loan opportunities. If the graph is engaging, it might mean the student compares loans, identifies alternative educational pathways, or pursues greater financial aid, all of which can have long-run benefits for the students' financial situation.

Finally, participants found that key values or critical pieces of information held the most weight in their understanding of the visualization and debt. Thus, **highlighting key values** is another guideline to take into consideration. Participants found that some of the most useful pieces of information within the visualizations presented were not on the graphs themselves. Participants often did not realize that, depending on the loan, it might take decades to pay off their debt. Highlighting the payoff date might serve as a rapid wake-up-call for students to realize how their debt might impact their future and how present actions can serve to protect their financial health. Moreover, showing how this payoff date interacted with interest rate and balance proved useful in establishing greater understanding of the visualization. The key values set participants up for more rapid understanding of the visualization, as it showed them what to look for on the graph, as well as the magnitude of the loan.

Chapter 5

Discussion

Pre-attentive vision played a critical role in participant understanding of the visualizations. Pre-attentive vision consists of four main attributes: color, form, spatial positioning, and movement (Hossain 2018). Throughout this study, these attributes played a major role in how participants perceived visualizations, how rapidly they could understand them, and how educational the visualization could be. Color has been discussed throughout this paper as having been important to participant understanding of the visualizations (Munzner 2014). Color exists within three channels: luminance, hue, and saturation. While luminance was not referenced by participants within this study, hue and saturation were. First, participants had clear preferences for hue. They preferred some colors over others, often citing the way the color made them feel or an association with a specific topic, such as money. Second, participants found that changing the saturation of the colors was not the most effective tool for conveying information on the visualizations. Stimuli 2 and 3 both changed the saturation, with stimulus 2 offering a shading of the bars and stimulus 3 contrasting two different saturations of green (lighter and darker). Participants found the shading of the bars in the second stimulus to be detrimental to their understanding, and most participants were ambivalent or disliked the saturation choice in the third stimulus. Munzner suggests that saturation is best used for ordinal data, and since the loan data presented was not ordinal, it is clear why participants may have disliked these saturation choices. Since the shading of the bars likely hindered participant ability to *detect boundaries* (a critical component of visualization tasks), it is clear why these participants did not prefer the shading presented within the second stimulus.

Moreover, Munzner builds upon this topic by introducing the *effectiveness principle*, which explains that the noticeability of a given attribute should match its importance. Thus, when participants were presented with colors that they could not identify as more important than the any other (such as two green colors), their confusion becomes clearer and more justified.

Form appears as a limited graphical element within a visualization, such as a line, a bar, a point, or a shape. Moreover, form expands upon these graphical elements by allowing for relative measures such as size to be included within a visualization. Within this study, participants are shown a simple line graph as well as two stacked bar charts integrated with a line chart. Technically speaking, a bar chart can be classified as a line chart expanded into both the vertical and horizontal dimension. Line graphs within solely the horizontal dimension can depict a trend between two variables. Within this study, the line graphs presented in this study depict a total remaining loan balance. When this line mark is combined with a bar graph, participants had a much more difficult time analyzing the material presented to them. When participants were asked if they would prefer if the stacked bar graphs could be separated into two visualizations or if they preferred the line graph by itself, participants with the least prior experience with visualizations found this option to be preferable. Munzner describes separability as a crucial component of visualization design. Within separability is the notion that marks should not clash—a horizontal mark should not be in the same field as a vertical mark. Within the second and third stimuli, this clash occurs. The line graph moves downward and towards the right, while the bar graph component moves upwards and two the right. The bar graph also includes two differing colors. Thus, while these marks are clearly distinct from one another, unless the participant had a prior clear understanding of debt topics, they likely were required to perform greater levels of critical thinking to interpret the meaning behind the two marks and their directional trends, thus prompting their dislike of the combinatorial graph

features. *Pattern detection*, another key visualization task, was recognized when participants discussed marks. Participants were able to identify marks moving in a direction, however understanding why that directional trend was happening was more difficult. For the first stimulus, understanding the pattern of the line was simpler than for stimulus two and three, as most participants were able to explain that the visualization depicted a loan balance decreasing over time. However, when this line was combined with a stacked bar graph, their ability to understand the reasoning behind the patterns was diminished. Pattern detection (or trend analysis) is a key driver behind visualization design. Rather than looking at a list of data and determining the direction in which that data is headed, a visualization offers a rapid way of assessing this directionality. Participants were overwhelmed by the competing direction of the line and bar graphs. Moreover, participants were confused why the second stimulus had a stacked bar that existed as a ratio between principal and interest, as many of them had never seen a stacked bar graph presented in this manner before. Participants without topic specific knowledge did not know that interest payments are greater at the beginning of a payment period than at the end of one, and thus were not able to pick up on this pattern within the graph. Therefore, it is evident that when the visualizations relied on topic-specific knowledge for detecting a given pattern, participants had a more difficult time interpreting the visualization. Moreover, the competing directionality of the marks identified another key point of difficulty for participants, as it required combining multiple competing key values and debt topics.

Spatial positioning refers to the location of a given mark on a visualization. While participants did not particularly notate the location of marks on the visualizations, they did often comment about the location of the 'critical values' such as interest, balance, and payoff date. Participants found this information most valuable, and found that when this information was

clearly visible, it enhanced their understanding the visualization. Munzner describes how cognitive load can impact the usefulness of a visualization. Since short-term memory is limited, the ability for participants to access the critical values relating to the proved even more important. Moreover, participants often asked for a hover feature within their interactivity demands, with this feature identifying the value of interest or even perhaps explaining how this value applies to the visualization and debt. This demand falls in line with Munzner's criteria for limiting cognitive load. Thus, while spatial positioning in this study was not topical for the marks within the visualizations, it absolutely had value for providing key information about the loan, such as identifying the most key values. This falls in line with the notion of a *visual search,* a visualization task. When the key values were dominant and clearly notated, participants were drawn to these pieces of information. Interestingly, participants even got confused as to if this information was the legend (which it was not). Even with this confusion, participants still found this information to be very valuable, since it was the most important information for their understanding of the visualization and the loan. By looking at the graph, participants would not be able to describe the monthly payment or interest rate without performing serious calculations, however, since this information was described to them in a clean manner, they would be able to learn everything they needed about the loan. When asked if they preferred this information by itself or alongside a graph, participants always said they preferred it with the graph, which amplifies the importance of including both key values and a visualization when describing student loans.

Finally, the pre-attentive attribute *movement* was not applicable to the visualizations presented within this study. All the visualizations did not include any moving components within them; thus, this study cannot attest to how movement may affect participant understanding.

Cleveland and McGill's 1984 framework for accuracy of understanding elementary tasks within a visualization is apparent within this study. In order of most accurate to least accurate, they first describe *positioning along a common scale*. All the visualizations presented within this study utilize a common scale for the values presented, which is shown through participant ability to identify that the visualizations are occurring on a standardized time X-axis. Their next task, *non-aligned scales*, was not applicable to this study. Following this, *length, direction, and angle* are introduced. Participants had greater difficulty understanding why certain bars were longer than others (particularly within the second stimulus), as well as why the line and stacked bars were in moving in opposite directions. This falls in line with Cleveland and McGill's assessment of accurate understanding. Next, *area and volume* were not applicable to this study, however, *curvature* was. The second and third stimuli presented both included a curved line, albeit the third stimulus' line's degree of curvature was significantly reduced when compared to the second stimulus'. Participants were able to detect that this line was decreasing as it represented the overall balance decreasing, however as to why the curvature of the line (and thus the slope) changed, participants provided no context. Finally, *shading and color saturation* were the most contentious tasks within this study. Participants had different interpretations and opinions regarding the same colors, indicating that an accurate assessment based on color universally was not possible. There was consensus that green equated to money and that red indicated a decrease or something malign, however, when multiple shades of the same color were presented, this consensus grew even muddier. Cleveland and McGill's ranking of the accuracy in understanding elementary tasks does seem to be apparent within this study, which reiterates the need for visualization design that can maximize an individual's understanding of a visualization through only the most effective and generalizable use of these tasks.

Limitations

This study was limited to participants from Penn State University – University Park. This study attempted to cover a wide variety of majors and undergraduate years. Despite the varied sampling within this study, only twelve participants were included. Thus, the sample size of this study provides its greatest limitation to generalizability.

Snowball (word-of-mouth) recruitment, while reputable and time-effective, does have its downfalls. First, individuals were recruited through connections to the researcher, and thus, all individuals recruited for this study were within two degrees of separation from the researcher. Since this thesis studies undergraduates at Penn State, this method isn't entirely without fault, however extrapolation to other universities, high school seniors, or graduate students may be limited.

While this study is an *exploration* of student debt visualizations and education, the scope of the study may be outside that of the research—with only twelve participants and the use of three visualizations, important lessons can be learned, however these lessons may not be able to answer the research questions presented within the introduction. That designation is up to the reader, however the author believes that whether the research questions are answered is of small consequence, since an exploration into the field of student debt visualizations still occurs and lessons are still learned.

Chapter 6

Conclusion

When undergraduates apply for public student loans, educational institutions are only mandated to perform entrance and exit loan counseling. For many students, this is the first and last time they interact with any form of debt education. Within this study, serious gaps in undergraduate knowledge of student debt are identified. While unsurprising given the lack of educational opportunities provided to undergraduates regarding their student loans, this lack of education clearly impacted students' ability to interpret and explain debt situations, including debt visualizations. Given the status of the U.S. Student Debt Crisis, further education is absolutely needed.

This study showed that current debt visualizations have effective components, but are at large, most ineffective resources for students. These visualizations lack the educational capacity needed for students to learn about their debt, and moreover, can be quite complicated, which might turn students away from pursuing a greater exploration of ways in which they can mitigate their debt. Visualizations need to be accessible to students, meaning they must be able to be rapidly ingested and offer students pathways to learn exactly the implications of what is being described on the visualization.

This study further described ways in which the current debt visualization complex can be improved. These suggestions can hopefully be used by stakeholders—such as the U.S. Department of Education, parents, educational institutions, and debt lenders—to create visualizations that prioritize education and exploration.

Student debt has lifelong consequences. Ensuring students can understand exactly the situation they are heading into should be a critical prerequisite to taking out a loan. Without better resources, debt mismanagement and financially reckless behavior will continue to maintain their position in the status quo.

Exploring student debt within this thesis has grown my own knowledge of the situation. I've learned that my peers are taking out loans out of necessity, as college was the only option they or their families had explored. I've learned that debt can be incredibly scary, and that because it is scary, individuals don't think too often about it. I've learned that unless someone actively engages with financial or quantitative material, they might never engage with a debt or financial visualization during their educational tenure. We know that STEM and business majors earn more money in their lifetimes on a consistent basis more than any other degree fields. I found these individuals needed the least prompting and educational explanations of the visualizations during my interviews. Thus, perhaps more education targeted towards people who have less experience with financial topics and quantification is necessary.

Student debt isn't going anywhere, and an overhaul of the student debt industry is unlikely. However, changing ways in which we educate ourselves about student debt may help turn the tide of this crisis, and I think it's worth giving it a shot.

Appendix A

Interview Questions

Background and Demographics

1. "What is your major?"

    1A. "If you have any minors, what are they?"

2. "What is your family's education background?"

3. "What is your current undergraduate year?"

4. "What classes have you taken at that might discuss debt or financial products?"

5. "If you can remember, what was discussed in the classes you mentioned?"

6. "If you have taken student loans, in what range will this total value fall into by the time you graduate. 0-10,000, 10,001-20,000, 20,001-30,000, 30,000+"

7. "What are your post-graduation plans?"

8. "What post-graduation salary will you be aiming for?"

9. "How do you believe this salary compares to other individuals in your field?"

Financial Institution Literacy

1. "What is your understanding of student debt?"

2. "Can you describe the difference between private and public student loans?"

3. "Can you describe the structure of public loans within the Department of Education?"

4. "If you have taken loans, what tools have you used to plan this decision?"

5. "How often do you review a loan debt summary?"

6. "Do present student loans impact your current decision-making process?"

Visualization Experience

1. "What is your understanding of a data visualization?"

2. "What classes have you taken that utilize data visualizations?"

3. "How familiar are you with crafting data visualizations?"

    3A. "What software have you used?"

    3B. "What components of a data visualizations do you find valuable?"

4. "How have you seen or used financial visualizations?"

    4A. "Can you describe these financial visualizations to me?"

    4B. "Have you used financial visualizations when deciding on taking on student debt?"

    4C. "What did these visualizations look like?"

    4B. "Is there anything that stood out in these visualizations?"

Visualization Testing

1. "What do you believe this graph is depicting?"

    1A. "Can you tell me what is on the X-axis and Y-axis?"

    1B. "How do these relate to student debt?"

    1C. "Can you describe the legend?"

        1Ci. "How does the legend relate to student debt?"

    1D. "How do you feel the colors of this graph affect your understanding of the information trying to be conveyed?"

    1E. "How does the design of the graph affect your understanding of the information trying to be conveyed?"

    1F. "If you were to change this graph, what would you change?"

1G. "If you were to add an interactive component to the graph, what would it be?"

1H. "Does this graph cloud, bolster, or have no effect on your understanding of student debt?"

1I. "What use cases would you think this graph could be useful for outside of student debt?"

1J. "If you had seen a graph like this before taking on loans, how might it have affected your decision to take those loans?"

BIBLIOGRAPHY

Amortization Chart. Mortgage Capital Partners, 2017,

    https://www.mortgagecapitalpartners.com/wp-

    content/uploads/2017/06/amortizationchart1.png.

Artavanis, Nikolaos. "Financial Literacy and Student Debt." *Taylor & Francis*, The European

    Journal of Finance, 22 Jan. 2019,

    https://www.tandfonline.com/doi/full/10.1080/1351847X.2019.1711435?casa_token=JGF

    DD-S7_o4AAAAA%3Aj2TjO_Fbvoyg9RoTVkA-

    zyVQj4ie9eyma3d0nOAD2At1wjT2PzHHuUMtbZgZUqe87SRKgAaQcRFB.

Barke, Shraddha, et al. "Grounded Copilot: How Programmers Interact with Code-Generating

    Models." *ACM Program*, vol. 1, Jan. 2018. *Conf. Article 1*.

"Best Practices for Financial Literacy and Education at Institutions of Higher Education."

    Treasury.gov, U.S. Financial Literacy and Education Commission, 2019,

    https://home.treasury.gov/system/files/136/Best-Practices-for-Financial-Literacy-and-

    Education-at-Institutions-of-Higher-Education2019.pdf.

"The Biden-Harris Administration's Student Debt Relief Plan Explained." *Federal Student Aid*,

    2022, https://studentaid.gov/debt-relief-announcement.

Cleveland, William S, and Robert McGill. "Graphical Perception: Theory, Experimentation, and

    Application to the Development of Graphical Methods." *Journal of the American*

Statistical Association, vol. 79, no. 387, Sept. 1984, pp. 531–554.,

https://doi.org/https://doi.org/10.2307/2288400.

"Cost-Of-Living-Adjustment (COLA)." *Cost-Of-Living Adjustment (COLA)*, 13 Oct. 2022,

https://www.ssa.gov/oact/cola/colasummary.html#:~:text=We%20call%20such%20increas

es%20Cost,next%20COLA%20in%20October%202023.

"Entrance Counseling Overview." *Federal Student Aid*, Department of Education,

https://studentaid.gov/entrance-counseling.

Galesic, Mirta, and Rocio Garcia-Retamero. "Graph Literacy." *Medical Decision Making*, vol.

31, no. 3, 2010, pp. 444–457., https://doi.org/10.1177/0272989x10373805.

Hossain, Pri. "Preattentive Processing and How to Use It in Design." *Medium*, UX Planet, 2 May

2018, https://uxplanet.org/preattentive-processing-and-design-e59eba74373e.

"HOW DO I COMPLETE THE LOAN ENTRANCE COUNSELING?" *Psu.edu*, Penn State

Bursar's Office, https://www.bursar.psu.edu/faq/how-do-i-complete-the-loan-entrance-

counseling.

Hubbard, Edward, et al. "Using Online Compound Interest Tools to Improve Financial Literacy."

The Journal of Economic Education, vol. 47, no. 2, 2016, pp. 106–120.,

https://doi.org/10.1080/00220485.2016.1146097.

McCarthy, Patricia C. "Today's College Student: Measuring the Effectiveness of Financial

Literacy Education and Effect on Subsequent Student Debt." *Indiana University of*

Pennsylvania, 2015.

Moorcraft, Bethan. "'Financial Illiteracy Is an Epidemic': Americans Lost an Average of over

$1,800 to Financial Errors in 2022 - Here Are 3 Big Money Mistakes You Could Be

Making Right Now." *Yahoo! Finance*, Yahoo!, 5 Mar. 2023,

https://finance.yahoo.com/news/financial-illiteracy-epidemic-americans-lost-

130000017.html.

Munzner, T. *Visualization Analysis & Design*. CRC Press, 2015.

Nadworny, Elissa. "College Aid Letters Are Misleading Students and Need a Legal Fix." *NPR*,

NPR, 6 Dec. 2022, https://www.npr.org/2022/12/05/1140807450/a-new-gao-report-takes-

a-look-at-college-financial-aid-letters.

Nye, Pete, and Cinnamon Hillyard. "Personal Financial Behavior: The Influence of Quantitative

Literacy and Material Values." *Digital Commons @ University of South Florida*,

Advancing Education in Quantitative Literacy, 2013, http://dx.doi.org/10.5038/1936-

4660.6.1.3.

"Ostrich Algorithm." *Wikipedia*, Wikimedia Foundation, 8 Dec. 2022,

https://en.wikipedia.org/wiki/Ostrich_algorithm.

Perna, Laura W, et al. "Understanding Student Debt: Implications for Federal ... - Sage

    Journals." *Sagepub*, American Academy of Political and Social Sciences, 17 Apr. 2017,

    https://journals.sagepub.com/doi/full/10.1177/0002716217704002.

Saldaña Johnny. *The Coding Manual for Qualitative Researchers*. Sage, 2009.

Simple Loan Calculator. Microsoft, https://pennstateoffice365-

    my.sharepoint.com/:x:/r/personal/gch5121_psu_edu/_layouts/15/doc2.aspx?sourcedoc=%7

    B4A2336D4-8FA5-4F24-802A-

    570AFA16549B%7D&file=Book%204.xlsx&action=editnew&mobileredirect=true&wdTp

    l=TM23154796&wdlcid=1033&wdNewAndOpenCt=1674607288162&ct=167460728816

    2&wdPreviousSession=0e91d444-472b-4289-9758-

    460ef46f065a&wdOrigin=OFFICECOM-WEB.MAIN.TEMPLATES.

Social Benefits of Homeownership and Stable Housing. NATIONAL ASSOCIATION OF

    REALTORS®, Apr. 2012,

    https://www.nar.realtor/sites/default/files/migration_files/social-benefits-of-stable-housing-

    2012-04.pdf.

Turner, Tracy, and Heather Luea. "Homeownership, Wealth Accumulation and Income Status."

    Journal of Housing Economics, Academic Press, 6 May 2009,

    https://www.sciencedirect.com/science/article/pii/S1051137709000217?casa_token=v5oiy

    mswjmwAAAAA%3A44NmIUFfqCUFzGMcq0tWXj1tuwCfeW7CLrH6BTUL2N5D9cO

    sU9NBXdGP62VSDH8R855CNmuKOg.

Unwin, Antony. "Why Is Data Visualization Important? What Is Important in Data

    Visualization?" *Harvard Data Science Review*, PubPub, 31 Jan. 2020,

    https://hdsr.mitpress.mit.edu/pub/zok97i7p/release/4.

Walsemann, Katrina M, and Jennifer A Ailshire. "Student Debt Spans Generations:

    Characteristics of Parents Who Borrow to Pay for Their Children's College Education."

    The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, U.S.

    National Library of Medicine, 1 Oct. 2017,

    https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5926980/.

ACADEMIC VITAE

105 Lantern Lane, Mount Laurel, NJ 08054
Phone: (856) 296-1170 ||| Email: gch5121@psu.edu, glenn.hubbard.career@gmail.com

Education

The Pennsylvania State University, University Park, University Park PA   May 2023
     Schreyer Honors College
     Bachelor of Science in Applied Data Sciences
     Minor in Information Science
Honors in Applied Data Sciences

Honors Thesis   Spring 2023

Understanding Undergraduates' Perceptions of Student Debt Visualizations: A Human-Centered Approach
Thesis Supervisor: Dr. Syed Billah, Assistant Professor of Information Sciences and Technology
Thesis Advisor: Dr. John Yen, Professor of Information Sciences and Technology

Work Experience

Gage Ventures, LLC   Spring 2020-Spring 2021

     Programmed mobile application using React Native/Google Firebase and conducted end-user testing through live demonstrations, allowing for rapid customer feedback and iteration

     Developed web presence using Bootstrap and Amazon Web Services such that the site operated at 99.5% uptime

     Optimized company's search engine ranking through HTML modifications and social media management

     Managed a team of five through a start-up accelerator, ultimately securing $15k in funding

Internship Experience

Information Technology and Data Analytics Intern, The Boeing Company   Summer 2022

     Implemented an automated email-sending platform across Alation and Dataiku to maintain consistent lines of communication and critical parameters to key members throughout The Boeing Company

     Generated statistical analyses on user groups of the Alation Data Catalog for the creation of targeted retention programs to ensure continued use of the Catalog and a growing userbase

     Cultivated a deeper understanding of airplane technology, aerodynamics, FAA regulation, and company processes including manufacturing, safety, sales, airline business models, and corporate mergers/venture capital

Grants Received
Rock Ethics Institute Fellow   Fall 2019

Awards
Undergraduate Research Award
     Excellence in Information Literacy—*Honorable Mention*   Spring 2022
     $50
Clayton H Schug Forensic Award   Spring 2022
     $250
Penn State Provost Award   Fall 2019-Spring 2023
     $24,000

Scholarships
Schwenger Scholarship   Fall 2021-Spring 2019
     $2,750
Liberal Arts College Endowment   Fall 2019-Spring 2020
     $5,000

Extracurricular Activities
Information Sciences and Technology Consulting Group   September 2020 – April 2021
President

     Collaborated with university officials and corporate recruiters to bring "meet-and-greet" events to members

     Transitioned the Group from limited operating status to an organization with direct lines of communication with university leadership, active client acquisition, and over 60 members

     Jointly developed a performant E-commerce platform for a Pennsylvania clothing business, empowering the business owner to update their inventory, ship nationwide, and take advantage of online payments

Penn State Speech and Debate Society   January 2021-Spring 2023
Debater

     Competing in nationwide Lincoln-Douglas competitions, previously earning the privilege to compete at the National Forensics Association Championship twice, and guiding the team towards a nationally recognized award

     Scouring legal literature to craft speech arguments centered around constitutional law and current events while operating as a research lead for arguments related to autonomous systems and drone technology

     Educated on debate theory, philosophy, and community ethics in a selective for-credit course designated for team members