Novice Programmer Planning: A Grounded Theory Approach

Jonathan Wellons, Dr. Julie Johnson
Vanderbilt University

Dr. Renee McCauley
College of Charleston

Context: Planning in Software Design

Programming classes differ in their emphasis on planning. Those that encourage it often do so in very different forms (e.g., writing the program comments in advance of any code versus a top-down modularization). It is widely known that planning is critical to programming success for non-trivial projects and that students may experience deceptive initial success with unplanned small projects. Given that programming is an essential skill in many engineering disciplines and that an enormous diversity of educational approaches are employed, it is natural to ask if there is a better way to teach planning.

Kuhl and Goschke[1] proposed a model for self-regulated learning which includes the steps of goal setting and planning. They believed that the model was recursive and that learners would return to this task again and again as they experienced internal feedback from the products they generated. We proposed this study in order to observe and explain planning in a broader sense among novice programmers. Due to the relative lack of literature in the field, we believe it would be helpful to analyze qualitative data to suggest a set of emergent theories. Such a theory and the subsequent results of its testing (not a part of this study) could lead to the development of tools to help students improve and refine this critical self-regulatory skill.

Our central question is: What theories can we formulate that could ultimately lead to better teaching practices in program planning?

Grounded Theory

Grounded theory is a form of qualitative research based on the formation of theory from data. The theory that results may be tested in subsequent quantitative research. This methodology consists of 5 steps: 1) gathering of **data**, 2) isolating meaningful points, or **codes**, 3) grouping of codes into **concepts**, 4) constructing **categories** that highlight the relationship between concepts, including a **core category** and 5) the suggestion of one or more **theories** which explain the data and may be tested in future research. Grounded theory researchers begin by casting a wide net to capture a diverse and multi-dimensional set of data which may be fertile ground for theories[2].

Grounded theory is *Emergent Research* rather than *Hypothesis-Testing Research*. Our goal is to generate ideas to be rigorously tested, without preexisting bias or theory. We asked open-ended questions and subjects were encouraged to discuss any aspects of their programming experience that they deemed meaningful.

"What is" Versus "What works"

Practitioners divide Research on Teaching into two kinds: "What is" and "What works". "What is" research focuses on observations about current conditions and processes in the learning environment. "What works" research tests and measures alternative teaching practices. This is a "What is" project attempting to suggest causes and effects of student program planning for further study.

Organization of Data

Our raw data consists of 11 interviews with students in Engineering Science 140. The interviews' total length is approximately 50,000 words. The interviews were refined into approximately 400 codes, divided over 14 concepts. The table below shows example codes taken from the data along with the corresponding concept. The codes are data points, which in the context of the interview, reveal a possible connection to planning.

Example Code	Concept
Initial plan was a short, vague list of tasks	Initial Planning
Plan failed because MATLAB doesn't handle long strings	Plan Adequacy
Program was constructed of fragments adapted from in-class examples	Programming
Searched for help on Google, e.g., "how to write a for loop"	Help Sources
Wrote program in 3-4 parts which were tested separately and combined at the end	Testing
"My goal was to receive an A or B"	Goals
It took 2 to 2.5 hours to finish. Had thought it would take 1 hour	Time Needed
"First day was too much, second day I started to understand, it clicked on the third day"	Class Experience
Programming background consists of using the Starcraft map editor	Programming Background
Had some advanced Math, "not good at it" but "loves it"	Quantitative Background
Calls self a "number cruncher" in everyday life	Hobbies
"Why go for the extra credit when I don't understand the basics"	Ambition
Rather than write a time-consuming brute-force program, solved the permutation puzzle visually out of 362,880 possibilities	Lateral Thinking
Influenced lab partners to use pseudocode in the future	Personality

Similar concepts are clustered into categories. The 14 categories that naturally arose in the above table are shown with their category in the table below.

Concepts	Category
Initial Planning, Plan Adequacy	Planning
Programming, Help Sources, Testing	Programming Methodology
Goals, Time Needed, Class Experience	Goal Setting and Achievement
Programming Background, Quantitative Background	Background
Hobbies, Ambition, Lateral Thinking, Personality	Personal

The Planning category is naturally the best choice to be deemed the core category as it is the focus of our work. Relationships between this category and others gives rise to theories.

Results: Hypotheses

Our goal in grounded theory is to creatively examine the apparent connections between categories to suggest theories which can explain the data before us.

Two families of theories suggest themselves: Those that suggest a student's background or personal details are correlated with a tendency to plan and those that appear to show a higher quality of educational experience for students who plan.

Beginning with the first family, the Quantitative Background category appears to be well-correlated with a student's propensity to plan. In particular, enjoyment of Mathematics is a better predictor than the student's claimed skill. Therefore we formulate:
Theory I: Students with enjoy Mathematics show a greater tendency to plan their programs

A similar correlation can be seen in subjects who enjoy pastimes, such as logic and strategy games and scientific reading.
Theory II: Students with quantitative hobbies are more likely to plan their programs

The second family of theories is strongly represented in the Goal Setting and Achievement category. In addition to expressing their ambition in other ways, students may choose between alternatives that offer different numbers of points in several of the programming assignments in the class. Based on this, we formulate:
Theory III: Students who plan their programs have higher ambitions in the assignment

Based on students' descriptions and assessments of their goals (such as grades, time taken, material mastered), we formulate:
Theory IV: Students who plan their programs are more likely to achieve their goals

Finally, based on wide ranging comments about the quality of a student's learning experience, we formulate
Theory V: Students who plan programs are more likely to report a positive experience in the class

Limitations

The main limitation of qualitative research of this kind is the time required to code and understand the large volume of interview data. It is not clear if more codes or a greater diversity of data sources sould be helpful. The nature of grounded theory places the burden testing the suggested theories on later work. Thus, we naturally cannot say whether the theories are true, which direction of cause and effect is present (if either), the strength of the correlation or how the theories could immediately be useful. Further quantitative research is required with statistical backing. Because the qualitative data gathered has an unlimited number of dimensions and is subject to interpretation, the choice of emergent theories is, by nature, subjective.

Study Design

In order to gather accurate data, volunteers were anonymously interviewed. This year the ES140 class was divided into three successive sections, each lasting approximately 14 lecture periods or 4 weeks. Each section consists of 25-30 students and the same subject material is taught to each section. Members of each section were interviewed.

This course introduces the fundamentals of programming (using MATLAB) within the context of cryptography. The general idea behind the course is to familiarize freshman engineering students with problem solving techniques and tools (such as MATLAB and Excel) while giving them a snapshot of the field of Computer Science and some of its practical applications.

All interviews are voice recorded for later analysis. Volunteers are encouraged to describe their experiences during several assignments, what type of plan they created, how detailed it was, and how it was adapted. Subjects are questioned about their background in programming, their hobbies and other details that could lead to a planning theory. Jonathan Wellons gathered the data, held office hours and guest lectured in 2 or 3 class sessions of each of the 3 sections to develop a rapport with students and introduce the study.

Conclusion and Future Work

In this project, we have applied grounded theory to interviews with novice programmers about their first programs. Through the principles of grounded theory (which leaves the testing of the theories for further work), we have coded, conceptualized and categorized the interview data. Next we have elucidated the connections between categories to generate several plausible theories to explain the data. Five theories are proposed in this project, based on the observed correlation between certain types of personal background and regular activities with the tendency to plan and also on the correlation of planning to several metrics of a successful educational experience.

Future work consists of two kinds: testing and refinement of these theories and the development of "What works" projects to realize the benefits proposed by these theories. Insights in to how students plan will enable us to offer guidance for students in these critical, early programming assignments. Next fall, we will proceed with a "What works" project with measurable objectives to test teaching strategies intended to exploit the apparent advantages suggested by theories III, IV and V.

Further Information

[1] Kuhl, J., & Goschke, T. (1994). A theory of action control: Mental subsystems, modes of control, and volitional conflict-resolution strategies. *In J. Kuhl & J. Beckmann (Eds.), Volition and personality Seattle, WA: Hogrefe & Huber.*

[2] Martin, Patricia Yancey, & Turner, Barry A.. (1986). Grounded Theory and Organizational Research. *The Journal of Applied Behavioral Science, 22(2), 141.*

[3] Creswell, John W. (2006). Qualitative Inquiry and Research Design. *Thousand Oaks, CA: Sage.*

I gratefully acknowledge the support of Dr. Johnson, Derek Bruff, the CFT and the Center for Integration of Research, Teaching and Learning project (NSF Grant No. DUE-0717768). Without this support, this research would not have been possible.