The Impact of Programming Experience on Successfully Learning Systems Analysis and Design

Wang-chan Wong
wcwong@csudh.edu
Information Systems and Operations Management Department
California State University
Dominguez Hills
Carson, CA 90747, U.S.A.

Abstract
In this paper, the author reports the results of an empirical study on the relationship between a student’s programming experience and their success in a traditional Systems Analysis and Design (SA&D) class where technical skills such as dataflow analysis and entity relationship data modeling are covered. While it is possible to teach these technical skills to students without programming experience, the results of the study strongly suggest that students with programming experience complete the course more successfully than those without.

Keywords: Systems Analysis and Design, Programming, IS Curriculum, Problem Solving

1. INTRODUCTION
For many of us formally trained in Systems Analysis and Design (SA&D), it is a logical assumption that some programming experience is a pre-requisite for taking a course in SA&D. However, as we have observed, many IS departments have relaxed their programming pre-requisite requirements, and this de-emphasis is reflected in the 2010 IS curriculum (Topi, et al, 2010). Thus, this current trend prompts the question: Does the lack of a programming background hinder understanding and subsequent success? To answer this, from 2007 to 2013 the author collected the homework scores of 15 SA&D classes from a total of 259 students. Statistical analysis of the data strongly supports the notion that programming experience is important for students to successfully complete the course. There have been similar studies conducted before, but they were typically carried out in a single class with limited sample size and for a short period of time. With a large sample size spanning six years, this study is more definitive and conclusive.

The rest of the paper is organized as follows. Section 2 is a literature review of teaching SA&D, particularly regarding its relationship to programming. Section 3 discusses the background of the study, while the methodology of the empirical study is explained in Section 4. The data analysis results are presented in Section 5. Finally, concluding remarks are given in Section 6.

2. LITERATURE REVIEW
There is a significant gap between teaching and research in SA&D (Bajaj, et al, 2005). This situation is reflected by the low number of research publications in this area. It is particularly true in finding research reports on teaching SA&D.

What do experts and scholars say regarding the relationship between programming and learning SA&D? Most acknowledge that programming is
an essential foundation of SA&D. Yourdon and Constantine’s book on Structured Design, a classic in structured analysis and design, discusses this relationship in the foreword: “...we assume that the reader knows how to code, and is capable of writing “good code”...” (Yourdon and Constantine, 1978, page xvi)

Booch suggests that his object-oriented method is “... most appropriate for courses in software engineering and advanced programming, and as a supplement to courses involving specific object-oriented programming languages” (Booch, 1994, page viii).

Rumbaugh, et al. propose Object Modeling Technique (OMT) as a method to develop object-oriented systems and to support object-oriented programming. They suggest that pre-requisites include “exposure to modern structured programming languages and a knowledge of basic computer science terms and concepts” (Rumbaugh et al., 1991, page x).

Booth’s discussion of “folk pedagogies” in software engineering asserts that students should take a programming course before a design course (Booth 2001).

Similarly, in his study on integrating programming and system analysis, Guthrie concludes that programming is the chicken, and system design is the egg. He demonstrates that a student’s design skill is directly related to their programming skill (Guthrie, 2004).

Studies have been conducted to determine the actual effect of a student’s prior background on their design proficiency. Judith Sims-Knight conducted a small empirical experiment in which she taught high school students and computer science students object-oriented design without programming by using CRC cards. While she found that the high school students were able to adequately handle the design process (the study did not screen whether the students had programming exposure or not), they concluded that the computer science students created more complete designs and demonstrated a deeper understanding of the design process (Sims-Knight and Upchuch, 1993).

In an experiment teaching object-oriented analysis and design (OOAD) to both computer science and math students, Boberic-Krsticev et al. (2013) note that the math students had basic problems mastering the materials, such as having difficulty acquiring fundamental concepts, and even the UML terminology. The deficiency has been attributed to the fact that these students did not have object-oriented programming backgrounds. Even with the computer science students experienced in object-oriented programming, they observe that students created UML diagrams simply for the sake of modeling; both groups of students failed to make connections between the models and their implementations. Similar to Guthrie (2004), the authors recommend that teachers should illustrate implementation in an OO programming language so students may see the connection between models and their implementations.

Chen suggests that DFDs and ERDs are the most important skills an analyst can have (Chen 2006). Only students with programming backgrounds can have a greater appreciation for the design principles that enable them to analyze and design more complex systems.

Serva (1998) argues that the technical tasks in SA&D are comparable to the difficulties of managing a project to its completion. He offers SA&D classes for non-IS students, but the coverage of the course is to simulate the difficulties of management within an IS environment and not to teach formal systems analysis and design.

Ultimately, many IS instructors, professionals, and practitioners alike prefer students to have programming experience before taking an SA&D course (Stack Overflow 2013).

3. BACKGROUND AND EMPIRICAL STUDY

The studies surveyed in Section 2 support the notion of learning programming before analysis and design. However, they are mostly small-scale experiments conducted for a short period of time. In this paper, the author is reporting on a study spanning over six years with a comfortable sample size, providing a more conclusive and definitive response to the issue.

The SA&D course in this investigation is offered by a public university -- a comprehensive urban university primarily serving a metropolitan area. Information Systems, just like other business majors such as accounting, finance, marketing, etc., is a concentration of the College of Business, instead of a single major. CIS 372 Analysis and Logical Design is a required course...
for IS and IS-related majors such as Information Systems Security, and Global Logistics and Supply Chain Management. However, the department has dropped the programming prerequisite for CIS 372. The only pre-requisite of CIS 372 is CIS 370, an introductory course on Information Systems Theory and Practice. CIS 370 is also a required course for all business undergraduate students. However, it does not provide any programming training besides a few Excel and ACCESS exercises with no coding required. As it is, CIS 372 has enrolled non-IS related students ranging from general business, marketing, human resources, finance, sports, entertainment and hospitality management, international business, accounting, biology, music, criminal justice, and sociology, in addition to computer science and information systems students. In fact, none of the non-IS students in this study had any experience with programming; even though IS and IS-related students are advised to take an introductory programming class before taking CIS 372, some do not follow this advice.

4. EMPIRICAL STUDY - METHODOLOGY

<table>
<thead>
<tr>
<th>Topics</th>
<th>Assignments*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to Analysis and Design</td>
<td></td>
</tr>
<tr>
<td>Analyzing the Business Case</td>
<td>Analysis of a Business Case</td>
</tr>
<tr>
<td>Managing Systems Projects</td>
<td>Project management, Gantt Diagram in MS Project</td>
</tr>
<tr>
<td>Output and User Interface</td>
<td>UI Design</td>
</tr>
<tr>
<td>Requirements Modeling</td>
<td>Requirements Document</td>
</tr>
<tr>
<td>Data and Process Modeling</td>
<td>DFD diagrams in Visio and process descriptions</td>
</tr>
<tr>
<td>Data Design</td>
<td>ERD in Visio and data dictionary</td>
</tr>
<tr>
<td>Development Strategies</td>
<td>Short questions</td>
</tr>
<tr>
<td>Systems Architecture</td>
<td>Short questions</td>
</tr>
<tr>
<td>Managing Systems Implementation</td>
<td>Short questions</td>
</tr>
<tr>
<td>Managing Systems Support and Security</td>
<td>Short questions</td>
</tr>
</tbody>
</table>

*The requirements document, DFD and ERD amount to 25% of the weighted total.

Table 1: Major topics covered and student deliverables in CIS 372

CIS 372 follows the traditional Systems Analysis and Design undergraduate curriculum except that it focuses on the functional (system) approach while the object-oriented approach is covered in another course. We used Systems Analysis & Design in a Changing World by Satzinger, Jackson, and Burd (Satzinger et al., 2007) for several years and then changed to Systems Analysis and Design by Shelly and Rosenblatt (Shelly et al., 2010) in 2010. Topics and major deliverables covered in the course based on Shelly’s book are shown in Table 1.

The course is not project or team-based; students work individually on homework assignments for each chapter. Since the author was interested in determining how well students without programming backgrounds can master technical skills, the homework scores of (1) the requirements document, (2) the dataflow diagram (DFD) design document, (3) the entity relationship diagram design (ERD) document, and (4) the weighted total of each student for the class were collected. The justification to select these four scores and the related hypotheses are as follows:

1. The requirements document is important in any system development, especially for non-IS students. In reality, a non-IS business person will have many opportunities to jointly develop Request for Proposals (RFP) and the requirements document with IS staff. Gaining technical writing skills will greatly improve the quality of the requirements document, RFPs, and any other project-related documentation.

2. DFD design documents are demonstrative of the technical skills that system analysts/developers need to have. These techniques train students how to capture the dynamics and behavior of a system. Needless to say, ERD is important because it trains students to capture the objects being modeled and relationships among them in the system.

3. The weighted total is the percentage of the scores a student receives. It includes all other homework assignments and exams. It is the overall measure of a student’s success in this course.

In the requirements document assignment, students are asked to define both functional and non-functional requirements for certain systems. Students will generate the requirements document similar to the clausal form example as
shown in (Taylor, 2013). Grading for the requirements document is based on clarity in writing, organization, and the ability to capture the major functional and non-functional requirements of the system. For the DFD and ERD assignments, questions are typically taken from the end of chapters in the textbook. The questions are based on small-scale business scenarios that are manageable for a single student. For the DFD assignment, students are asked to draw the context diagram and then decompose it to Level 0 and/or Level 1 diagrams in Visio. Students also need to write high-level process descriptions for each primitive DFD process. Structure charts are not covered, since students without programming experience have difficulty understanding parameter passing and functional decomposition. Grading for the DFD assignment is based on the syntactic correctness of the diagrams, the appropriate logical flow to capture the dynamics and behavior of the system based on the scenario given, and the clarity of process descriptions for the primitive processes.

For the ERD assignment, students are asked to create the crow’s foot model in Visio along with its data dictionary. Grading on the ERD assignment is based on how well the student identifies the entities and their relationships, including meaningful entity names, salient attributes of these entities and appropriate domain or data types for the attributes, correctness of the relationships (cardinalities) for these entities, correct primary key identification and referential integrity enforcement, and correct modeling of the logical and physical data models using Visio.

Since any student may take CIS 372 without fulfilling a programming pre-requisite, for the purposes of this study students were asked about their programming experience during the first class meeting.

**Hypotheses**

The requirements document is in an itemized clausal form by grouping system specifications into categories and subcategories, mirroring the hierarchical structure of a program and the relationships of its components. In an introductory programming class, students have exposure to top-down modular design, structure programming, and/or other programming paradigms such as the object-oriented approach. They are also introduced to the three basic programming constructs: sequence, iteration, and selection. In fact, if a student masters these programming fundamentals, they will easily be able to learn data flow analysis. Furthermore, in programming class, we always emphasize program documentation. When students decompose a program into sub-modules, they need to document the interface, function descriptions, in-line comments, etc. Therefore, the following two hypotheses are posited:

Hypothesis 1 (H1): A student’s requirements document assignment score is positively associated with his/her programming knowledge. Students with programming backgrounds have better scores than students without programming backgrounds.

Hypothesis 2 (H2): A student’s data flow assignment score is positively associated with his/her programming knowledge. Students with programming backgrounds have better scores than students without programming backgrounds.

In a typical introductory programming course, students are exposed to basic data structures such as record, array, files, and relational databases. They will have seen how records are linked and processed. Therefore, the following hypothesis is posited:

Hypothesis 3 (H3): A student’s entity relationship modeling assignment score is positively associated with his/her programming knowledge. Students with programming backgrounds have better scores than students without programming backgrounds.

Finally, if students are able to successfully manage an introductory programming class with the exposures we describe above, they will be able to smoothly transition into learning SA&D. Therefore, this gives rise to the following hypothesis:

Hypothesis 4 (H4): A student’s weighted total is positively associated with his/her programming knowledge. Students with programming backgrounds have better scores than students without programming backgrounds.

The four hypotheses are depicted in Figure 1.
The data were loaded to SPSS version 21 for statistical analysis. Among the 249 students analyzed, 114 were non-IS students, while 135 were IS-related students. All non-IS students in this sample did not have any programming background prior to CIS372, while 14 out of the 135 IS-related students did not take programming classes prior to CIS372. The summary is tabulated in Table 3 below.

<table>
<thead>
<tr>
<th>IS?</th>
<th>No</th>
<th>Yes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS?</td>
<td>114</td>
<td>0</td>
<td>114</td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>121</td>
<td>135</td>
</tr>
<tr>
<td>Total</td>
<td>128</td>
<td>121</td>
<td>249</td>
</tr>
</tbody>
</table>

Table 3: IS/Non-IS with Programming Background Summary

Their descriptive statistics are summarized in Table 4 below.

<table>
<thead>
<tr>
<th>Group Statistics</th>
<th>PROG?</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>REQ</td>
<td>Yes</td>
<td>121</td>
<td>78.5840</td>
<td>22.6500</td>
<td>2.05910</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>128</td>
<td>69.6641</td>
<td>27.5710</td>
<td>2.43702</td>
</tr>
<tr>
<td>DFD</td>
<td>Yes</td>
<td>121</td>
<td>63.9669</td>
<td>29.0330</td>
<td>2.63942</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>128</td>
<td>54.9375</td>
<td>29.8270</td>
<td>2.63637</td>
</tr>
<tr>
<td>ERD</td>
<td>Yes</td>
<td>121</td>
<td>66.1653</td>
<td>27.3890</td>
<td>2.48996</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>128</td>
<td>53.6719</td>
<td>30.0290</td>
<td>2.65425</td>
</tr>
<tr>
<td>Weighted Total</td>
<td>Yes</td>
<td>121</td>
<td>74.3473</td>
<td>13.0270</td>
<td>1.18428%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>128</td>
<td>68.2066</td>
<td>13.9190</td>
<td>1.23032%</td>
</tr>
</tbody>
</table>

Table 4: Descriptive Statistics of the Empirical Study

t-test

The mean scores of the requirements document, DFD design, ERD design, and the overall weighted total in Table 4 reveals that students with programming backgrounds performed better than students without any experience. The consistency and validity of the test scores are justifiable because the assignments have similar degrees of difficulty and were graded by the same instructor for a period of six years. To further analyze the data, independent sample t-tests were conducted in SPSS where the Grouping Variable is the PROG?, and the Test Variables are the REQ, DFD, ERD, and Weighted Total.
Total. The null hypotheses predict that the mean scores of the Test Variables are the same between students who had taken programming classes prior to taking CIS 372 and students who had not.

The t-test results are summarized in Table 5.

<table>
<thead>
<tr>
<th>t-test for Equality of Means</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tail)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>REQ</td>
<td>2.781</td>
<td>247</td>
<td>.006</td>
<td>8.91996</td>
<td>3.20798</td>
</tr>
<tr>
<td>DFD</td>
<td>2.419</td>
<td>247</td>
<td>.016</td>
<td>9.02944</td>
<td>3.73338</td>
</tr>
<tr>
<td>ERD</td>
<td>3.424</td>
<td>247</td>
<td>.001</td>
<td>12.49341</td>
<td>3.64880</td>
</tr>
<tr>
<td>Weighted Total</td>
<td>3.589</td>
<td>247</td>
<td>.000</td>
<td>6.14071%</td>
<td>1.71088%</td>
</tr>
</tbody>
</table>

Table 5: Independent Samples t-test Results (alpha = 0.05)

The t-test results clearly support the four hypotheses shown in Figure 1. It is not surprising that students with programming experience perform better in DFD and ERD than those with no experience. After all, these two technical skills are equivalent to skills used in programming. However, it is interesting to note that students with programming knowledge also outperformed students without programming knowledge in the requirements document assignment. This may be explained by the fact that functional requirements are similar to functional and procedural descriptions in programming exercises. Requirements are written at a higher level of abstraction but are still modular in nature.

**Regression Analysis**

The author further performed a linear regression as a predictive model to measure the potential student completion success of the course. The dependent variable is the Weighted Total, the independent variables are the scores of the requirements document, the DFD and ERD assignment scores, and the control variables are the IS?, PROG? and Class. The regression results are summarized in Table 6.

The regression analysis results in Model 2 show that the scores of the requirements document, DFD, and ERD are all significant at the 0.05 alpha level, confirming the results of the t-test above. Hence, the regression equation is:

\[
\text{WeightedTotal} = 42.355 + 0.175 \times \text{DFD} + 0.154 \times \text{ERD} + 0.118 \times \text{REQ}
\]

The coefficients of the DFD, ERD, and REQ are 17.5%, 15.4% and 11.8%, respectively. Tellingly, while these three assignments amount to only 25% of the total course requirements, the adjusted R² is at 0.484. This suggests that, among all other assignments and tests, the scores of these three assignments alone can explain almost 50% of a student’s overall performance.

### 6. DISCUSSION AND CONCLUDING REMARKS

Results of the empirical study strongly suggest that students with programming experience will complete the course more successfully than those who don’t have experience.

Why is programming so important in learning SA&D? In a paper written by Professor David Gries at a 1974 ACM conference, he points out that general problem-solving is very unique in teaching programming (Gries, 1974). In the same paper, Gries illustrates his arguments and summarizes a four-phase process in problem-solving proposed by Polya in 1945 (Polya, 1945). This process is nearly identical to what is known today as SDLC. See Table 7 below for the comparison.
Polya’s 4-Phase Process | SDLC
---|---
Understand the Problem | Planning
Devise a plan | Analysis
Carry out the plan | Design
Look back | Implementation
Support/Enhancement

Table 7 Poly’s Problem Solving Process vs SDLC

In fact, Poly’s four-phase problem-solving process already suggests the incremental and iterative approach that we currently consider best practice. In Gries’ words, “In a programming course, we attempt to teach the student how to program anything that can be programmed -- that has an algorithmic solution” (Gries, 1974, page 81). This “algorithmic” discipline and training empowers students to solve programs in a systematic way. With this training, students can smoothly transition to software engineering or SA&D, in which they cope with the complexity of solving larger problems in a more conceptual and abstract manner.

Jeffries, et al. studied the processes involved in designing software, and concluded that the decomposition process is central to creating the design (Jeffries, et al., 1981). The process is similar to the stepwise refinement proposed by Wirth; decomposition and stepwise refinement are usually covered in introductory programming class (Wirth, 1971).

The recent IS 2010 curriculum excludes programming from the core requirements even though Systems Analysis and Design (2010.6) remains one of the seven core courses in the guidelines (Topi, et al, 2010). The guidelines proposed in 2010.6 have further replaced technical skills, such as structured and object-oriented approaches, with less programming-oriented business process modeling. It is undeniable that these replaced technical skills are crucial for students intending to further develop their careers in application development and system analysis. In a recent survey conducted by The Economist, 38% of respondents admit that “inadequate technical skill sets” are the biggest challenges for CIOs trying to align technology use with business goals, while only 21% think “inadequate management skills” are the biggest challenge (The Economist, 2013). Another study from the U.K. reports that almost 75% of current IT leaders -- an overwhelming majority -- are unsure that the CIOs of today will still be the right people to lead IT businesses in 2018 (Nice, 2013). Nearly half (43%) of the respondents are concerned by their potential deficiency in technical skills.

Time and time again we are reminded that the critical skills of an IT manager include project management, communication, writing, etc. While we all agree that these soft skills are crucial competencies for an IT manager, they do not necessarily mandate an undergraduate IS curriculum. Realistically, most undergraduate IS students will not be hired for management positions right after graduation. A technical position is still most likely be the first job for many IS graduates. Consider the current IS/IT job markets: reports show that programming and application development is one of the top 10 hottest IT skills for 2013 and 2014 (Pratt, 2012; Brandel, 2014; Simoneau, 2014; Wakefield, 2014). Another report suggests that 60% of the surveyed companies claimed they would hire more developers in 2013 (Pratt, 2012). Most recently, U.S. News & World Report named computer systems analysts as second place in their ranking of 2014’s best jobs (Best Jobs, 2013).

If we intend to maximize the ability of our IS students to successfully find employment, it is our responsibility to properly equip them with sufficient technical skills. For that reason, it is necessary to continue encouraging programming as a pre-requisite to Systems Analysis and Design.

### 7. REFERENCES


http://money.usnews.com/careers/best-jobs/computer-systems-analyst


Robey, Daniel (2001) "Blowing the whistle on troubled software projects", Communications of the ACM (44)4, pp. 87-93.


teaching-systems-analysis-and-design-how-much-programming-experience-is-needed


The Economist (2013). The strategic CIO-Risks, opportunities and outcomes. The Economist Intelligence Unit Limited 2013


