2016

Can You Hack It? Validating Predictors for IT Boot Camps

Courtney Gear
Minnesota State University Mankato

Can You Hack It? Validating Predictors for IT Boot Camps

Courtney C. Gear

A Thesis Submitted in Partial Fulfillment for the Degree of Master of Arts in

Industrial/Organizational Psychology

Minnesota State University, Mankato

Mankato, Minnesota

March 30, 2016

Can You Hack It? Validating Predictors for IT Boot Camps

This thesis has been examined and approved by the following members of the student's committee.

Dr. Andrea Lassiter – Advisor

Dr. Scott Fee – Committee Member

Dr. Daniel Sachau – Committee Member

Abstract

Given the large number of information technology jobs open and lack of qualified individuals to fill them, coding boot camps have sprung up in response to this skill gap by offering a specialized training program in an accelerated format. This fast growth has created a need to measure these training programs and understand their effectiveness. In the present study, a series of analyses examined whether specific or combinations of predictors were valid for training performance in this coding academy. Self-rated, daily efficacy scores were used as outcome variables of training success and correlation results showed a positive relationship with efficacy scores and the logic test score as a predictor. Exploratory analyses indicated a Dunning-Kruger effect where students with lower education levels experience higher overall mood during the training program. Limitations of the study included small sample size, severe range restriction in predictor scores, lack of variance in predictor scores, and low variability in training program success. These limitations made identifying jumps between training stages difficult to identify. By identifying which predictors matter most for each stage of skill acquisition, further research should consider more objective variables such as instructor scores which can serve as a guideline to better asses what stage learners join at and how to design curriculum and assignments accordingly (Honken, 2013).

Keywords: IT boot camps, training effectiveness, predictors

Table of Contents

Background

With the resurgence of the American economy following the Great Recession, not only is it important that jobs are available to Americans, but furthermore it is imperative that Americans are equipped with the training necessary to do those jobs. According to a 2015 Business Wire report, the U.S. Bureau of Labor Statistics projects that hiring of software developers will grow twice the rate of the average for all occupations through 2022, spurring demand for more boot camp-style developer training programs (Schaffer, 2015). There are nearly five job openings for every developer and four times as many openings for programmers as there are qualified people to fill them (Tynan, 2014). Given the large number of information technology jobs open and the lack of qualified individuals to fill them, the need to provide more training opportunities is evident from the creation of programs such as President Obama's TechHire Initiative, which was launched as:

> a bold multi-sector effort and call to action to empower Americans with the skills they need, through universities and community colleges but also nontraditional approaches such as coding boot camps, and high-quality online courses that can rapidly train workers for a well-paying job, often in just a few months." (Office of the Press Secretary, 2015)

In addition to the TechHire Initiative, a variety of coding boot camps have sprung up in response to this skill gap by meeting the needs of employers through a specialized training program in an accelerated format.

History

The coding boot camp phenomenon began in 2012 with Dev Bootcamp, one of the pioneers in the coding boot camp industry (Schaffer, 2015). Commonly referred to as 'boot camps', 'accelerated learning programs', or 'coding academy,' a coding boot camp is defined as an "intensive, accelerated learning program that teaches beginners digital skills like Full-Stack Web Development, Data Science, Digital Marketing, and UX/UI Design" (Eggleston, 2015). There are approximately 67 US and Canada-based coding schools with 13 of those schools reporting no graduates in 2014; those groups do expect to graduate 585 students in 2015 (Eggleston, 2015).

General structure. Course Report, a market research group that publishes reviews, application tips, and interviews to equip students with coding boot camp information, published the 2015 Coding Boot Camp Survey which outlined the current coding boot camp environment and projections for the future year. According to Course Report's survey:

> The 2015 boot camp market will grow by 2.4x, to an estimated 16,056 students in 2015, up from 6,740 in 2014. As a point of comparison, we estimate that there were 48,700 undergraduate computer science graduates from accredited US universities in 2014. The actual market size in 2014 was 6,740 graduates. Average tuition price of qualifying courses is $11,063, with an average program length of 10.8 weeks. The estimated tuition revenue from qualifying US schools will be $172M (up from $52M in 2014), excluding scholarships. (Eggleston, 2015)

Given the growth of coding boot camps and the need for qualified individuals to fill software development roles, there is a need to measure these training programs and understand their effectiveness.

The potential market. The growth of coding academies is likely in response to demand from students and employers alike. When comparing the price of a degree from a university or community college to the experience gained in a coding academy, many students may be drawn by the price and time commitment. According to the American Association of Community

Colleges 2015 Fact Sheet, the average annual tuition and fees of a four-year college is $9,139 and is $3,347 for a community college. Students are beginning to ask the question, why spend two to four years in college, when a similar educational experience can be achieved in a 12 week boot camp? Additionally, many boot camps offer an apprenticeship component so students are paid to continue to learn on the job after completing the training. The time to complete a boot camp is on average three months compared to two to four years at a university.

Moreover for employers, flexibility in curriculum helps employers tailor the program to a specific programming language that is currently in demand. Boot camps can change their curriculum to meet the needs of an employer without issue (Hendershot, 2015) compared to a university that is less likely to change curriculum semester after semester. This previous research leads to the purpose of the present study: to determine the extent to which predictors used to select applicants into the boot camp training program lead to success in the program.

Selection Process

Within the field of Industrial/Organizational Psychology, the identification of predictors and criteria allows us to better understand how one relates to the other. To examine the effectiveness of coding boot camps and understand what makes applicants successful, a variety of predictors have been identified as useful in the selection process. Predictors are individual difference scores that represent knowledge, skills, abilities, or other characteristics. They are examined often in the role they play in employee selection, staffing, and promotion. Predictors are generally categorized into the following three domains: cognitive, non-cognitive, and performance-based (Ployhart, et al., 2006). Research has demonstrated that the combination of all three types of predictors provides the best incremental validity over the use of cognitive predictors alone (Ployhart, et al., 2006).

Cognitive predictors. Cognitive predictors often refer to individual differences that reflect intellectual functioning, academic aptitudes, knowledge, and related cognitive processes. Cognitive predictors examine mental abilities and are strongly linked to general cognitive ability. A meta-analysis conducted by Schmidt and Hunter (1998), reviewed 19 various selection procedures to understand which constructs, measurement methods, and combinations of predictors were most predictive of job performance. In this study, general mental ability was found to be the only procedure that could be generalized across the majority of jobs and maintain the highest validity of .51 (Schmidt & Hunter, 1998). General mental ability was also found to predict job performance, acquisition of on-the-job knowledge, and performance in job training programs more strongly than any other selection method (Schmidt & Hunter, 1998). The addition of a general mental ability test was found to add incremental validity by 24% when combined with other selection measures (Schmidt & Hunter, 1998). This finding showcases the direct effect general mental ability has on job performance as it measures an employee's ability to quickly learn on the job. While a strong predictor of performance, the use of cognitive predictors alone often shows subgroup differences in minority and majority groups which result in unequal hiring rates. Given the large body of evidence supporting general mental ability as a useful predictor, [company] created and utilized a logic test as a cognitive predictor in the selection process. Therefore:

Hypothesis 1. Applicants who score higher on the logic test are more likely to be successful in the training program.

Non-cognitive predictors. Individual differences that reflect dispositions, traits, motivation, choice, and interest are referred to as non-cognitive predictors (Ployhart, et al., 2006). Biographical data is commonly obtained through application blanks and personality questionnaires, which are empirically developed (Schmidt & Hunter, 1998). Biographical data is used in selection to understand past behavior of candidates; therefore, information related to previous jobs held, education, and special skills is typically requested (Ployhart, et al., 2006).

Biographical data alone has substantial validity of .35 for predicting job performance, but when added to general mental ability the incremental increase in validity is small, .01 (Schmidt & Hunter, 1998). This small increase is likely due to the strong correlation of biographical data and general mental ability (.50, Schmidt, 1988). Additionally, biographical data has been shown to predict performance in training programs with a validity of .30 (Hunter & Hunter, 1984). Despite any shortcomings in predictive capability compared to using cognitive ability alone, a key benefit of using biographical data as a predictor is that evidence has shown it to minimize subgroup differences (Pulakos & Schmitt, 1996).

Job experience. Schmidt and Hunter (1998) examined the relationship of previous job experience on similar jobs and found that for work experience less than five years, the correlation of job experience and performance is larger when measured by supervisors (.33) and .47 when measured by a work sample test. However, after five years, experience seems to plateau and additional job experiences do not lead to increased job performance (Schmidt & Hunter, 1998). Furthermore, job experience does not predict performance in job training (Schmidt & Hunter, 1998) so lack of experience does not impede new skill acquisition on the job.

A commonly used method of evaluating previous experience both on-the-job and in training is the point method. Using this method, points are awarded for years or months of job experience, years of relevant educational experience, relevant training experiences, etc. (Schmidt & Hunter, 1988). The point method has found to have a low validity coefficient of .11 (Schmidt

& Hunter, 1988). Furthermore, across many studies, years of formal education does not predict future job performance as it had an even lower validity of .10 (Schmidt & Hunter, 1988). This finding shows that for some jobs, years of formal education does not predict future job performance. Formal education has been a common biographical data source used in many occupations. Considering some coding academies expect a minimum level of education to help candidates perform better:

Hypothesis 2. Applicants with an Associate's Degree will be more likely to be successful in the training program than applicants with less than an Associate's Degree.

Performance-based predictors. Methods such as interviews, work samples and simulations are considered performance-based predictors because they closely resemble the actual job behaviors expected on the job (Ployhart, et al., 2006). To be effective, performance-based predictors need to have high physical fidelity, meaning they most closely resemble performance on the job. Additionally, they need to be consistent in terms of structure, questions asked, and scoring method (Ployhart, et al., 2006). In 1984, Hunter and Hunter found ability-based tests to be the best overall predictor of job performance with a mean validity of .53. Alternatively, they found a mean validity of .14 for other predictors. For the purpose of this study, the coding academy conducts a webpage simulation where they ask applicants to submit their resume in the form of a webpage, and the result is then scored based on the quality of the code. This simulation is an example of a performance based predictor:

Hypothesis 3. Applicants who score higher on the webpage simulation will be more successful in the training program.

Structured interviews. With more recent research suggesting a different conclusion, Huffcutt and Arthur (1994) conducted a reanalysis of Hunter and Hunter's 1984 study with the intention of overcoming several limitations in the previous study. Their findings showed a mean validity of .37 for interviews overall (Huffcutt & Arthur, 1994) and with corrections, a range of .20 to .57 (Huffcutt & Arthur, 1994). Their study examined the effect interview structure has on validity and found that as structure increases so does validity; however they did find a ceiling effect, indicating at a certain point, more structure adds little incremental validity (Huffcutt & Arthur, 1994). In a separate study, Campion, Pursell, and Brown (1988) conducted a study to determine if they could raise the psychometric properties of highly structured interviews even further. They found a corrected interview validity of .56 (Campion, Pursell, & Brown, 1988), which is similar to previous validities of cognitive ability tests, thus suggesting that the structured interview is an equally effective selection technique. Therefore:

Hypothesis 4. Applicants who score higher on the interview are more likely to be successful in the training program.

The identification of predictors as they relate to criteria is an important relationship to consider. In the case of IT boot camps, training success is measured by whether an individual completes the training program and the type of employment opportunity they are given after completion. In order to better identify the criteria, it is important to understand the various training models and how they relate to the structure of IT boot camps.

Training models

The IT boot camp model lies in the premise of expert-novice training programs that focus on a person's ability to quickly acquire the basic knowledge to execute a new skill and then through the use of apprenticeships, fine-tune these skills.

Dreyfus five stage model of skill acquisition. In 1985, Stuart Dreyfus and Hubert Dreyfus proposed a five-stage model of skill acquisition describing the changes that occur when moving from a novice learner to master (Dreyfus & Dreyfus, 1980). The argument Dreyfus and Dreyfus (1985) make is that “skill in its minimal form is produced by following abstract formal rules, but that only experiences with concrete cases can account for high levels of performance” (p.5). This original model described the stages as: novice, competent, proficient, expert, and master (Dreyfus & Dreyfus, 1985). In 2004, the model was modified by Stuart Dreyfus to include the following stages: novice, advanced beginner, competence, proficiency, and expertise. The 2004 model will be referenced going forward and applied to the process of learning software development in the coding boot camp setting.

Novice stage. According to Dreyfus (2004) the first stage is called Novice and is characterized as simply learning rules and following them. Context has not yet been provided to learners. For coding boot camps, this is the pre-work stage before the boot camp begins where applicants are learning the basics of software development and have not yet had any classroom instruction.

Advanced Beginner stage. The second stage, Advanced Beginner, builds on the rules and begins to put context with them, thus teaching the learner instructional maxims or examples of when certain rules do and do not apply. Students that have begun the in class instruction and have also worked through modules to apply their learnings are in this advanced beginner stage.

Competency stage. Stage three, Competency, could be considered the breaking point for many new learners as the learner in this stage feels overwhelmed with how and when to apply the rules to contexts and build their library of knowledge. The learner also experiences emotion, both positively and negatively, and feels remorse when their actions lead to failure. The emotional aspect of this stage is what assists some learners in moving ahead and holding others back. Dreyfus (2004) explains that this emotional investment and taking responsibility for one’s actions, both the successful and unsuccessful ones, strengthens the learner's perspective on what works and does not. Students in the heat of the boot camp training program will remain in this stage as they make mistakes and begin to learn what is required to succeed in this profession. The peak of boot camp will likely take an emotional and mental toll on the learner, but once they push through, their confidence will be restored.

Proficiency stage. In stage four, Proficiency, the learner begins to discriminate between situations and while he/she now sees what needs to be done, he/she still must think about it and then decide how to do it; the action is not yet automatic. For learners in the boot camp, most have likely reached the proficiency stage by the end of formal training in boot camp. They should feel confident in their abilities as they transition into apprenticeships.

Expertise stage. The final stage, Expertise, is characterized by the automation of responses meaning they intuitively know what to do immediately. No longer is the learner thinking about what to do because they have experienced many unique situations and know the appropriate response. As an expert, the learner now focuses on new and unusual experiences because the others no longer require much thinking (Dreyfus, 2004). In this final stage of skill acquisition, the learner has applied his/her learnings to real work situations and responses have become almost automatic when solving issues related to software development. For boot camp graduates, they should move towards becoming an expert at the end of their apprenticeships and as they move forward in their career. The expectation is not that they are experts in software development when they graduate, but rather they have a solid understanding and can troubleshoot problems independently.

The Dreyfus model is an important tool to aid instructors and future managers in understanding where students that are currently in or have recently completed a coding academy program may be. Students may not necessarily move through each stage at the same pace and therefore instructors need to be cognizant of these individual differences and be willing to provide feedback accordingly. For example, at completion of the boot camp, some students may move into an apprenticeship at the proficiency stage, and others may only be at the competency stage, but with additional on-the-job training, they will quickly move to the next stage.

For the present study, a series of analyses will examine whether specific or combinations of predictors are valid predictors of training performance in this coding academy. Figure 1 illustrates this relationship of predictors and the Dreyfus (2004) stages. As an example, stage three in the 2004 Dreyfus model emphasizes the importance of learners pushing through the emotional challenges of learning a new skill. In the case of the IT boot camp, this will play out by how students manage the time commitment, pressure to learn quickly, and acquire the skills dictated by the program. Furthermore, by identifying which predictors matter most for each stage of skill acquisition, these qualifiers can serve as a guideline for instructors to better asses what stage learners join at and how to design curriculum and assignments accordingly (Honken, 2013).

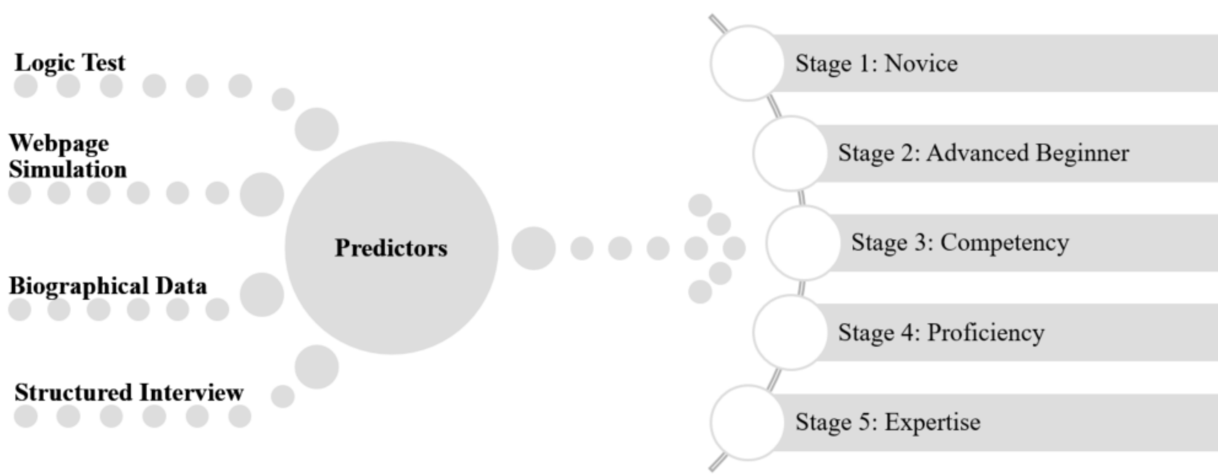

Figure 1. Relationship model of predictors and training stages

Method

Participants

A total of 103 students of [company name] were included as participants in this study. Five cohorts with 19-22 students in each were included beginning with cohort "A" in April 2015 and concluding with cohort "E" in February 2016. Throughout the course of the program, a total of four students dropped out, one from cohort B, one from cohort C, and two from cohort D; therefore, the remaining 99 students were used in analyses. The student make-up was 58% male, 37% female and homogeneous with 3% Black, 6% Hispanic, 6% Asian, and 86% White. No students self-identified as "Islander" or "Native American." A breakdown of demographics by cohort is included in Table 1.

Table 1

Student Demographic Information

Cohort	μ Age	*SD* Age	Female	Male	White	Asian	Black	Hispanic	Other
A	30.63	5.85	11	9	20	0	0	0	0
B	37.10	19.20	9	13	19	1	1	1	1
C	31.28	7.25	6	12	14	1	1	2	0
D	39.14	21.20	6	11	14	2	1	2	0
E	34.22	7.36	22	6	18	2	0	1	2
Totals	34.50	5.85	38	60	85	0	0	0	0
Percentage	---	---	37%	58%	83%	6%	3%	6%	3%

Measures

In order to validate predictors, scores on a variety of cognitive, non-cognitive, and experience-based measures were considered. A multiple-hurdles approach was used, meaning that applicants must pass the logic test, short answer questionnaire, and webpage simulation before moving on to the interview stage. Scores on all measures were used individually as part of an applicant's overall score.

Cognitive predictors. A logic test was administered to all applicants to measure basic cognitive ability. This test was created by [company name] and is part of the three online assessments used in the selection funnel. Scores on the logic test ranged from three to 11 with the mean score across all participants being 8.71. Table 2 shows the minimum, maximum, and mean logic test scores across all cohorts.

Table 2

Logic Test Score Ranges Across All Cohorts

Cohort	Min	Max	μ	*SD*
A	7	11	9.40	1.54
B	3	11	8.18	2.59
C	3	11	8.47	2.29
D	5	11	9.10	1.77
E	5	11	8.45	1.65
Overall	3	11	8.71	2.03

Non-cognitive predictors. A general application with demographic information as well as short answer questions was administered to all applicants. The purpose of the general application was to identify themes related to resilience and grit as well as collect basic information on each applicant. The general application consists of six, short-answer questions, each of which was scored on a one to five scale and then totaled for an overall score ranging from 14 to 30. Table 3 shows the minimum, maximum, and mean short answer questionnaire scores across all cohorts.

Table 3

Short Answer Questionnaire Score Ranges Across All Cohorts

Cohort	Min	Max	μ	*SD*
A	20	30	24.85	3.17
B	18	30	23.36	3.14
C	16	26	22.05	3.01
D	14	27	21.90	3.73
E	14	29	22.95	4.16
Totals	14	30	23.04	3.57

Experience-based predictors. To assess experience, applicants were asked to submit their resume in the form of a webpage simulation which was then scored by a trained rater. The purpose of the webpage simulation was to understand the individual's experience with software development, as well as gauge their interest and willingness to learn and try new skills. Scores ranged from two to five. Table 4 shows the minimum, maximum, and mean simulation scores across all cohorts.

Table 4

Webpage Simulation Score Ranges Across All Cohorts

Cohort	Min	Max	μ	*SD*
A	3	5	3.85	0.81
B	2	5	3.55	1.01
C	2	4	3.26	0.73
D	2	5	3.60	0.99
E	2	5	3.59	0.73
Overall	2	5	3.57	0.87

Additionally, a structured phone interview was conducted with applicants who successfully passed the logic test, short answer questionnaire, and webpage simulation. Interviews were conducted one-on-one by a member of the [company] team. The applicant's responses were subjectively scored as either "meets expectations" or "exceptional," indicating responses were above and beyond. There was low variability in scores as 99 participants were rated as "meets expectations" and four participants (one participant in cohort A and three participants in cohort D) were rated as "exceptional."

Feedback is a critical component of [company's] process, so in addition to the applicant information, daily mood, efficacy, and support scores were used as criteria. As part of the classroom experience in weeks six through eighteen, all students were sent a daily survey (see Appendix A) to voluntarily complete based on their experience that day as it related to their overall mood, confidence in their abilities (efficacy), and how supported they felt by [company] staff. These daily scores were on a five-point Likert scale with one indicating a low score and five indicating a high score.

Procedure

A correlational study using archival data was used to examine the relationship between predictors and criteria for all participants. All information on applicants used was archival data. This design was chosen given the real-life context of the training program. For data collection, all scores related to the application process as well as scores derived throughout training were considered. A discriminatory factor analysis was used to predict which predictors best classify learners by stage in the training process.

Furthermore, it was anticipated that analyses could indicate where jumps between training stages occur thereby informing instructors of the key turning points throughout training.

Finally, results of the study illustrated a profile of the learners that are most likely to be successful in the training program. These results are helpful to the coding academy as they continue to fine-tune their selection procedures and focus on the differentiating predictors that ultimately lead to more successful graduates of the program.

Results

Training success defined

Before testing the hypotheses, the first step was to define success in the boot camp training program. The training program is split into three phases, each lasting six weeks. The first phase involves six weeks of pre-work where students complete self-paced online modules prior to classroom work. Phase two occurs during weeks seven through 12 and is primarily lecture-based training in a classroom setting. During phase two students are expected to complete individual assignments daily and weekly to assess individual progress through curriculum topics. The third and final phase of the training program occurs in the classroom from weeks 13 to 18. However, the work during this final phase transitions from lecture-based to project-based learning where students are expected to apply classroom learnings to real-world projects and work effectively on teams. Results from these daily survey questionnaires were averaged to compute overall mood, efficacy, and support scores for each candidate and then used as outcome variables in the analyses. Additional new variables were computed to look at score means for mood, efficacy, and support for the first half of classroom training as well as the second half of classroom training. The purpose of creating these mean scores for the first and second halves of training was to capture the change from lecture-based to project-based learning. Finally, mood, efficacy, and support scores were averaged from week one of classroom training (Week 1 Mean Score), week two (Week 2 Mean Score), week four (Week 4 Mean Score), week six (Week 6 Mean Score) and the final two weeks (Weeks 11 and 12 Mean Score) of training. These time points were selected based on the curriculum topics provided by [company] and the five stages of the Dreyfus Training Model (Dreyfus, 2004). Students with less than 20 responses over the 60 classroom days were removed, leaving 62 of the total 103 cases remaining.

Correlational analyses

Hypothesis 1. To test the first hypothesis, a correlation analysis was utilized. Correlations among the logic test scores and the mean efficacy scores for week one, week two, week four, week six, and the final two weeks were compared using an alpha level of .05. Results showed logic scores were significantly and positively related to week one mean efficacy scores (r=.38, p=.002), mean efficacy scores from the first six weeks of training (r=.27, p=.03), and overall mean efficacy scores (r=.32, p=.01). Table 5 shows these correlation results across all time periods.

Table 5

Means, Standard Deviations, and Correlations Between Logic Test Score and Efficacy Scores

Variable	N	μ	*SD*	*r*
Logic Score	64	9.00	1.89	-
Week 1 Mean Efficacy Score	64	3.77	0.60	$.38^{**}$
Week 2 Mean Efficacy Score	64	3.44	0.65	.24
Week 4 Mean Efficacy Score	63	3.38	0.62	.18
Week 6 Mean Efficacy Score	62	3.47	0.69	.17
Weeks 11 and 12 Mean Efficacy Score	36	3.77	0.62	.31
First Half Mean Efficacy Score	64	3.45	0.50	$.27^{*}$
Second Half Mean Efficacy Score	62	3.61	0.53	.25
Overall Mean Efficacy Score	64	3.50	0.47	$.32^{*}$

*Significant at $p < 0.05$ level **Significant at $p < 0.01$ level

After significant correlation results were identified, a linear regression analysis was conducted with logic scores as the independent variable and week one mean efficacy score as the dependent variable. Results indicated a significant prediction, ($\beta = .38$, $t(62) = 3.23$, $p = .002$). Logic test scores also explained a meaningful amount of variance in week one efficacy scores, $R^2 = .14$, $F(1, 62) = 10.41$, $p = .002$. A second linear regression of mean efficacy scores from weeks one through six on logic test scores indicated a significant prediction, (β = .27, $t(62) = 2.19$, p = .03). Logic test scores explained a small amount of variance in week one through six mean efficacy scores, R^2 =.07, $F(1, 62)$=4.81, p=.03. A final linear regression of overall efficacy scores means across all 12 weeks indicated a significant prediction, (β =.32, $t(62)$=2.64, p=.01). Logic tests scores explained a meaningful amount of variance in the overall efficacy scores, R^2 =.10, $F(1, 62)$=6.98, p=.01. These results suggest that logic tests scores do predict training efficacy, supporting hypothesis 1.

Hypothesis 2. For the second hypothesis, education level was dichotomized into two groups, the first being participants with below an Associate's Degree, which included participants that self-identified as having "some college" and "high school." The second group included participants with education levels above an Associate's Degree which included the following groups: "Associate's Degree", "Bachelor's Degree", and "Graduate Degree." A total of 27 participants were included in the "less than Associate's Degree" group and 72 participants were included in the "Associate's Degree plus" group. A point-biserial correlation was conducted to examine the relationship of education level and mean efficacy scores for week one, week two, week four, week six, the final two weeks, the first half of training, the second half of training, and overall. Results showed a negative, significant relationship between overall efficacy score and education level (r_{pb}=-.26, p=.04), suggesting participants with lower education levels have higher overall efficacy scores in training. Hypothesis two was therefore not supported and Table 6 shows the results of the point-biserial correlation analysis.

Table 6

Means, Standard Deviations, and Correlations Between Education Level and Efficacy Scores

Variable	N	μ	*SD*	*r*
Education Level	63	3.63	1.15	-
Week 1 Mean Efficacy Score	64	3.77	0.60	-.05
Week 2 Mean Efficacy Score	64	3.44	0.65	-.16
Week 4 Mean Efficacy Score	63	3.38	0.62	-.22
Week 6 Mean Efficacy Score	62	3.47	0.69	-.23
Weeks 11 and 12 Mean Efficacy Score	36	3.77	0.62	-.23
First Half Mean Efficacy Score	64	3.50	0.47	-.04
Second Half Mean Efficacy Score	64	3.45	0.50	-.24
Overall Mean Efficacy Score	62	3.61	0.53	-.26*

*Significant at $p < 0.05$ level **Significant at $p < 0.01$ level

Hypothesis 3. Correlation analyses were also used to test hypothesis three which examined the relationship of scores on the webpage simulation with efficacy scores at various time points throughout training. None of the time points showed a significant relationship; therefore, hypothesis three was not supported. Table 7 below shows results of the correlations conducted.

Table 7

Means, Standard Deviations, and Correlations Between Simulation and Efficacy Scores

Variable	N	μ	*SD*	*r*
Webpage Simulation Score	64	3.66	0.86	-
Week 1 Mean Efficacy Score	64	3.77	0.60	.05
Week 2 Mean Efficacy Score	64	3.44	0.65	.18
Week 4 Mean Efficacy Score	63	3.38	0.62	.08
Week 6 Mean Efficacy Score	62	3.47	0.69	-.13
Weeks 11 and 12 Mean Efficacy Score	36	3.77	0.62	.12
First Half Mean Efficacy Score	64	3.45	0.50	.09
Second Half Mean Efficacy Score	62	3.61	0.53	.00
Overall Mean Efficacy Score	64	3.50	0.47	.09

*Significant at p < 0.05 level **Significant at p < 0.01 level

Hypothesis 4. Correlation analyses were used to test the final hypothesis, which examined interview scores and training success. Results showed no significant relationships; therefore, hypothesis four was not supported. Table 8 displays the correlation results.

Table 8

Means, Standard Deviations, and Correlations Between Interview and Efficacy Scores

Variable	N	μ	*SD*	*r*
Interview Score	64	0.05	0.21	-
Week 1 Mean Efficacy Score	64	3.77	0.60	.02
Week 2 Mean Efficacy Score	64	3.44	0.65	-.07
Week 4 Mean Efficacy Score	63	3.38	0.62	-.12
Week 6 Mean Efficacy Score	62	3.47	0.69	-.08
Weeks 11 and 12 Mean Efficacy Score	36	3.77	0.62	.25
First Half Mean Efficacy Score	64	3.45	0.50	-.08
Second Half Mean Efficacy Score	62	3.61	0.53	.10
Overall Mean Efficacy Score	64	3.50	0.47	.00

*Significant at p < 0.05 level **Significant at p < 0.01 level

Discriminatory Factor Analysis

A discriminatory factor analysis was conducted to identify whether the predictor variables (logic test, short answer questionnaire, webpage simulation, and interview) accurately predict student training success and furthermore, if specific variables predict better than others. Success in the training course was determined by considering the median overall efficacy score for all respondents (3.50). Participants with an overall efficacy score between 0.00 and 3.50 were coded as not successful, and participants with an overall efficacy score between 3.51 and 5.00 were coded as successful. Results of the discriminatory factor analysis were non-significant.

Logistic Regression Analysis

A logistic regression was performed to assess the impact of the predictor variables on the likelihood participants would be successful in the training course. The same categorical dependent variable created in discriminatory factor analysis was used for the logistic regression analysis. The model contained four independent variables (logic test score, short answer questionnaire score, webpage simulation score, and interview score); however, it was not found to be statistically significant χ^2 (1, N=62) = 5.16, p =.27.

Further models were explored to try and identify the likelihood of a participant being successful at various time points according to the Dreyfus model of skill acquisition (Dreyfus, 2004). Figure 2 summarizes the alignment of stages and critical points in time given efficacy scores. Efficacy score means for week one, week two, week four, week six, the final two weeks of training, the first half of training and the second half of training were all dichotomized into successful and unsuccessful using the median scores for each. Table 9 shows the median cut points for each point in time during the training program.

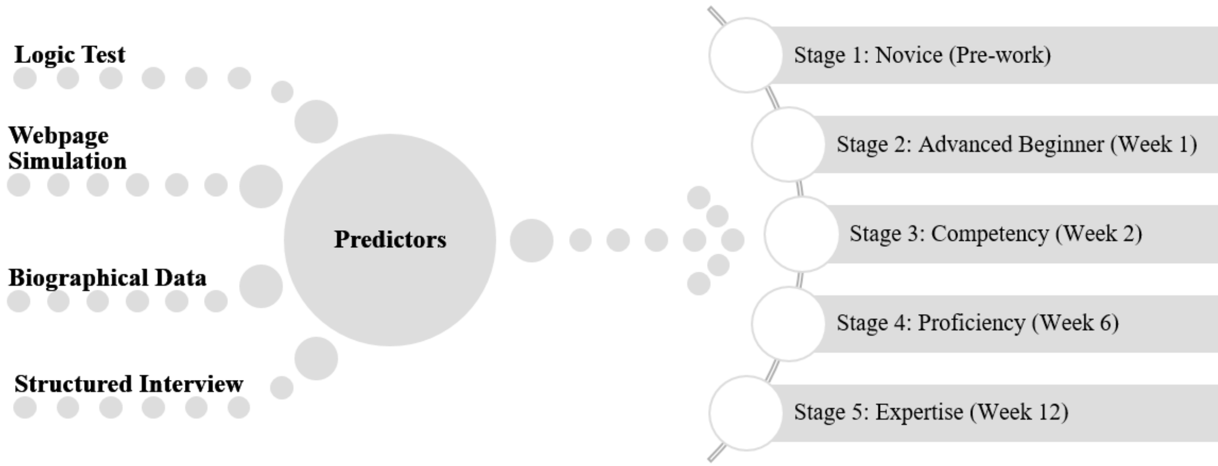

Figure 2. Dreyfus stages aligned to training time points

Table 9

Median Efficacy Scores at Various Points in Time

Point in time	Median Score	Unsuccessful Range	Successful Range
Week 1 Mean Efficacy Score	3.76	0 – 3.76	3.77 – 5.00
Week 2 Mean Efficacy Score	3.43	0 – 3.43	3.44 – 5.00
Week 4 Mean Efficacy Score	3.39	0 – 3.39	3.40 – 5.00
Week 6 Mean Efficacy Score	3.49	0 – 3.49	3.50 – 5.00
Weeks 11 and 12 Mean Efficacy Score	3.77	0 – 3.77	3.78 – 5.00
First Half Mean Efficacy Score	3.45	0 – 3.45	3.46 – 5.00
Second Half Mean Efficacy Score	3.60	0 – 3.60	3.61 – 5.00

Seven additional models were also tested using logistic regression analysis. Each model contained the five independent variables (logic test score, short answer questionnaire score, webpage simulation score, and interview score) and one of the newly created dichotomized variables; however, none of these models were found to be statistically significant. Table 10 summarizes the results of these analyses.

Table 10

Results of Logistic Regression of Training Success on Selection Predictors

Point in time	N	df	χ^2	*p*
Week 1 Mean Efficacy Score	64	4	7.91	.10
Week 2 Mean Efficacy Score	64	4	6.57	.16
Week 4 Mean Efficacy Score	63	4	.75	.95
Week 6 Mean Efficacy Score	62	4	4.71	.32
Weeks 11 and 12 Mean Efficacy Score	36	4	4.48	.35
First Half Mean Efficacy Score	64	4	3.10	.54
Second Half Mean Efficacy Score	62	4	5.51	.24

Exploratory Research

A series of exploratory correlation analyses were conducted with the selection predictor scores (logic test, short answer questionnaire, resume webpage simulation, and interview) and the mood scores for week one, week two, week four, week six, final two weeks, the first half of training, the second half of training, and overall scores. Results showed scores on the logic test were significantly and positively related to the week four mean mood score (r=.30, p=.02). Table 11 illustrates these correlation results.

Table 11

Means, Standard Deviations, and Correlations Between All Predictors and Mood Scores

Variable	N	μ	*SD*	1.	2.	3.	4.
1. Logic Test Score	64	9.00	1.89	-	-	-	-
2. Short Answer Questionnaire Score	64	23.84	2.89	-	-	-	-
3. Webpage Simulation Score	64	3.66	0.86	-	-	-	-
4. Interview Score	64	0.05	0.21	-	-	-	-
5. Week 1 Mean Mood Score	64	3.60	0.59	.11	-.12	.09	-.09
6. Week 2 Mean Mood Score	64	3.42	0.66	.23	-.03	.22	-.14
7. Week 4 Mean Mood Score	63	3.45	0.66	.30*	.04	.12	.02
8. Week 6 Mean Mood Score	62	3.41	0.76	.09	-.16	-.18	-.09
9. Weeks 11 and 12 Mean Mood Score	41	3.24	0.96	-.13	-.03	.15	.20
10. First Half Mean Mood Score	64	3.42	0.50	.20	-.10	.12	-.11
11. Second Half Mean Mood Score	64	3.23	0.73	.20	-.09	.04	-.00
12. Overall Mean Mood Score	64	3.39	0.52	.20	-.08	.12	-.06

*Significant at $p < 0.05$ level **Significant at $p < 0.01$ level

Additionally, a point-biserial correlation was conducted with the dichotomized education variable and mood scores for the following time points: week one, week two, week four, week six, first half of training, second half of training, and overall scores. A summary of correlation results is included in Table 12. Results showed a negative, significant relationship between education level and overall mood for the second half of training (r_{pb}=-.28, p=.03), suggesting participants with lower education levels have higher overall moods in training. These results indicate a Dunning-Kruger effect (Kruger & Dunning, 1999) where less educated students experience a more positive overall affect during training.

Table 12

Means, Standard Deviations, and Correlations of Education Level and Mood Scores

Variable	N	μ	*SD*	r_{pb}
1. Education Level	63	.78	.42	-
2. Week 1 Mean Mood Score	64	3.60	.59	.19
3. Week 2 Mean Mood Score	64	3.42	.66	-.04
4. Week 4 Mean Mood Score	63	3.45	.66	-.14
5. Week 6 Mean Mood Score	62	3.41	.76	-.24
6. Weeks 11 and 12 Mean Mood Score	41	3.24	.96	-.07
7. First Half Mean Mood Score	64	3.42	.50	-.10
8. Second Half Mean Mood Score	64	3.23	.73	-.28*
9. Overall Mean Mood Score	64	3.39	.52	-.18

*Significant at $p < 0.05$ level **Significant at $p < 0.01$ level

Discussion

The purpose of this study was to validate whether predictors used in the selection process do in fact lead to more successful boot camp graduates. Given a more objective training performance criteria was not available, self-rated, daily efficacy scores were used as outcome variables of training success. While efficacy scores did prove to have a relationship with the logic test score predictor, efficacy really only gauges how capable learners feel rather than their actual ability and performance. Instructor scores would likely prove more useful as outcome variables given the research questions and would also likely be less subjective criteria. Given the results of the correlations conducted, it seems the logic test is the only predictor currently used for selection into the boot camp that is valid for predicting a student's success in the training program. This finding is not surprising given previous research on the validity of cognitive predictors. While cognitive ability may be a valid predictor for training performance, it seems there are other unknown facets of a student's profile that make someone successful in this IT boot camp. Perhaps constructs such as positive affect, teamwork, and tenacity should be considered as selection variables to provide a more well-rounded candidate profile.

A secondary purpose of the study was to identify which predictors were related to each of the training stages to better understand the pivotal learning points in the program. Results of the discriminatory factor analyses and logistic regression analyses showed that none of the selection predictors significantly predicted training program success at any of the critical time points identified. More data and further research is needed to better identify these critical time points, as well as better outcome variable such as performance criteria. The lack of variance in attrition rates may have played into these analyses as well. For example, only four of the 103 students that started the program dropped out; therefore, making it much more difficult to identify points in time where learners would typically drop out, thus indicating jumps between training stages.

Results of the exploratory analyses indicated students with lower education levels experience higher overall mood during the training program. This finding is called the Dunning-Kruger effect and states that “those who are less skilled tend to overestimate their abilities more than do those who are more skilled” (Simons, 2013, p. 601). This effect was originally discovered and named in 1999 after a series of experiments conducted by Justin Kruger and David Dunning at Cornell University where they looked at participant’s scores on humor, grammar, and logic as it relates to estimates of test performance and ability. Results showed participants in the bottom quartiles considerably overestimated their performance and ability. This misaligned view was linked to lack of metacognitive competence as once that was increased, participants better recognized their limitations (Simons, 2013). Although education level at the Associate’s Degree did not have a positive relationship with efficacy scores, it may be worthwhile to consider what major or degree specifically those with formal education obtained. For example, a comparison of computer science to non-computer science degrees may indicate that students start the boot camp at a higher level than those with only some college or high school education.

Limitations and Future Research

Several limitations should be noted as drawbacks in the current study. Small sample size and severe range restriction in predictor scores, a side effect of concurrent validation studies, led to low variability and therefore made correlation analyses difficult. Low variability in training program success given the majority of students successfully completed the program makes identifying jumps between training stages difficult to identify. Training success was defined by self-reported efficacy scores which sufficed for this project, but ideally instructor-provided performance scores would be used for future research.

With concurrent validation studies such as this one, range restriction is a common problem, as the study is only considering participants who “passed” the bar to get into the training program; however, this lack of variance in predictor scores, proves challenging during analyses. For interview scores in particular, there were only two ratings provided during interviews, of which 99 students received a “meets expectations” rating and four received an “exceptional” rating. The training program would benefit from a more structured interview that digs deeper on results of the previous selection tests, thus more clearly distinguishing differences between candidates. Additionally, a structured interview could gather information on some of these other constructs that are not assessed in the other predictors to ensure the candidate is best equipped for the program and is the right fit for a particular cohort.

At the time of analysis, 103 students had completed the training. While this was initially thought to be a sufficient sample size, the analysis results suggest more students are needed. As an example, for some analyses, as few as 32 students were considered based on the number of survey responses available for that day. A recommendation for future research includes improving survey response rates by encouraging students to complete the survey every day, not just in the beginning. Alternatively, [company] could measure pre and post efficacy scores for each training module to better identify where students have more foundational knowledge compared to brand-new concepts. Given the voluntary nature of responses, this had a large impact in the data that was valid and available to be used in analyses.

This particular boot camp has recently implemented a more robust evaluation system for daily and weekly student assignments during the course of the training program. While the lack of performance ratings available for this study made finding relationships between predictors and training stages difficult, future research could consider assignment or project ratings and instructor ratings as criteria for training success. Ideally links from selection predictors to success in the training program and finally job status and post-training salary would be drawn and ultimately show whether the boot camp has a valid selection system.

In addition to considering training performance criteria, [company] should consider what distinguishes graduates from the program who graduate in the competency versus proficiency versus master stage (Dreyfus, 2004) and furthermore what employers who hire IT boot camp graduates are looking for. If the expectations is that graduates are simply competent software programmers rather than proficient, that likely leads to a different training design. Are companies hiring boot camp graduates with the expectation they will need additional on-the-job-training or do they expect graduates to be ready to jump into a project? These are questions worth answering and clarifying as boot camp curriculum continues to be fine-tuned.

The final point to consider is what level new students join the boot camp at. While everyone completes the same six weeks of pre-work, there is no quantitative data available, other than the week one mean efficacy scores, to understand a student's confidence with the material prior to pre-work compared to after pre-work. Additionally, the same data is missing from the apprenticeship phase where students are essentially learning on the job. It is unclear what the expectation in either stage is and if a higher level of efficacy or understanding of the material proves more useful to students. To gather a more holistic picture of a student's training experience it would be most helpful to add these additional time points to survey collection.

References

American Association of Community Colleges. (2015). 2015 Fact Sheet. Retrieved from http://www.aacc.nche.edu/AboutCC/Documents/FactSheet2015.pdf

Asher, J. J. (1972). The Biographical Item: Can it be improved? *Personnel Psychology, 25*, 251-269.

Campion, M. A., Pursell, E. D., & Brown, B. K. (1988). Structured Interviewing: Raising the Psychometric Properties of the Employment Interview. *Personnel Psychology, 41*(1), 25-42.

Dreyfus, S. E. (2004). The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science, Technology, and Society, 24*(3), 177-181. doi:10.1177/0270467604264992

Dreyfus, S. E., & Dreyfus, H. L. (1980). A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition. *Air Force Office of Scientific Research*, 1-16.

Eggleston, L. (2015). *2015 Coding Bootcamp Survey.* Retrieved from https://www.coursereport.com/2015-coding-bootcamp-market-research.pdf

Hendershot, S. (2015). Can you hack it? *Crain's Chicago Business, 38*(10).

Honken, N. (2013). Dreyfus Five-Stage Model of Adult Skills Acquisition Applied to Engineering Lifelong Learning. *Paper presented at 2013 ASEE Annual Conference, Atlanta, Georgia*. Retrieved from https://peer.asee.org/19457

Huffcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984) Revisited: Interview Validity for Entry-Level Jobs. *Journal of Applied Psychology, 79*(2), 184-190.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.

Kruger, J., & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

Office of the Press Secretary. (2015, March 9). FACT SHEET: President Obama Launches New TechHire Initiative. The White House. Retrieved from https://www.whitehouse.gov/the-press-office/2015/03/09/fact-sheet-president-obama-launches-new-techhire-initiative

Ployhart, R. E., Schneider, B., & Schmitt, N. (2006). *Staffing Organizations Contemporary Practice and Theory.* Mahwah: Lawrence Erlbaum Associates, Inc., Publishers.

Pulakos, E., & Schmitt, N. (1996). An Evaluation of Two Strategies for Reducing Adverse Impact and Their Effects on Criterion-Related Validity. *Human Performance, 9*(3), 241-258.

Schaffer, R. (2015, March 12). Three Years after Lanching the Coding Boot Camp Industry, Dev Bootcamp Celebrates with an Initiative to Increase Diversity in Tech. *Business Wire*. Retrieved from http://www.businesswire.com/news/home/20150312005185/en/Years-Launching-Coding-Boot-Camp-Industry-Dev#.VhWKBHpVhBc

Schmidt, F. L. (1988). The problem of group differences in ability scores in employment selection. *Journal of Vocational Behavior, 33*, 272-292.

Schmidt, F. L., & Hunter, J. E. (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin, 124*(2), 262-274.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Shane, G. S. (1979). Further tests of the Schmidt-Hunter Bayesian validity generalization procedure. *Personnel Psychology, 32*(2), 257-281.

Simons, D. J. (2013). Unskilled and optimistic: Overconfident predictions despite calibrated knowledge of relative skill. *Psychonomic Bulletin & Review, 20*(3), 601-607.

Tynan, D. (2014). Boom or bust: The lowdown on code academies. *InfoWorld.com*. Retrieved from http://ezproxy.mnsu.edu/login?url=http://search.proquest.com/docview/1496924908?accountid=12259

U.S. Bureau of Labor and Statistics. (2014, January 8). Retrieved from Occupational Outlook Handbook: ttp://www.bls.gov/ooh/computer-and-information-technology/software-developers.htm

Appendix A

Daily Survey Questions

1. Please rate the following questions:

	1	2	3	4	5
How are you feeling?					
How well did you understand the material covered today? (if relevant)					
Did you feel supported by staff today?					

2. Share one thing that you're proud of accomplishing today.
3. What did you like about today? Was there anything you didn't like about today?
4. If you're feeling lost, which topic(s) are you struggling the most with?
5. X wants to have a supportive environment for you. If you felt unsupported today, we'd like to hear a couple of words about your experience.