The effects of interns's former professional experience and GPAs on their received supervision

IBEB Bachelor Thesis 2014

Student name and number: Daniel Djevelekov 343266

Supervisor: Jurjen Kamphorst

Date of submission: 3/7/2014

Abstract

Purpose – The purpose of this paper is to investigate into how cognitive skills develop for interns during their placement and what factors make the internship a valuable tool for both the skill increase of the intern and the benefit of the recruiting firm.

Approach – A microeconomic model will be developed based on cognitive skill acquisition theory and its predictions will be tested against an online survey.

Findings – This paper finds out that interns receive more supervision when they have relatable previous experience, but only if their tasks during the internship are relatively complex. If their tasks are relatively routine, relatable previous experience has the opposite effect – it reduces the amount of supervision received. Another finding is that students receive increasingly more supervision effort from their coaches as the internship progresses.

Implications – The specific findings on the organizational structure of internships can be of high value to both internship supervisors and participants. These findings shed a new light on a popular social issue – whether internships can be considered value-increasing for society. The link between organizational theory and cognitive science we make allows for a new perspective for the approach to examining internships – one that would allow for more general studies and communication between researchers.

Contents

Introduction

In recent years participation in an internship program has become a vital prerequisite for starting almost any career. The National Association of Colleges and Employers (NACE) reports increases of nearly 25% in the number of interns over the past 5 years. In light of this trend, there has been a significant academic interest in developing a theory of the proper organization and carrying out of an internship program. The vast majority of scientific papers focus on one profession or on one specific employer with the most vigorous research being produced in the fields of medicine and education. This approach delivers systematic structural improvements in their respective placement programs by tackling specific issues such as the lenght of the program (Santos & Nunes, 2009) or what the best methods of evaluation are (Nash, Norcross, & Stevenson, 1986). Another type of study is done ,among others, by Benjamin (2001) and Rodolfa et al. (2005) who report on different organizational schemes and compare their effectiveness. Both types of research have high practical and academic value, but they lack generality in the sense that they are only applicable to a very narrow set of cases (their respective fields or specific internship programs). A major backfall of these types of studies is that they are conducted on students who choose similar career paths and come from similar backgrounds, individual differences that interns have are put aside and instead, the focus is on the institutional organization of the internship and what the empirically tested best practices are. The popularity of this inductive approach leads to a lack of academic effort that takes into account multiple spheres of professional development and the multitude of individual differences between participants in the programs. In result, it is increasingly difficult to isolate general best practices for interships as a whole and there is little communication between researchers in each professional field. For these reasons, the following paper will attempt to develop a general theory applicable to many professions which focuses on the individual rather than the organizational structure. Using skill acquisition theory, a microeconomic model will be developed to examine how supervisor-student interaction increases the interns' skill and how these interactions are affected by the intern's learning abilities and former experience. Accepting that any cognitive skill follows similar developmental patterns accross a broad range of disciplines allows for an optimal allocation of training effort depending on every intern's own background and potential.

In the meantime, mainstream media coverage on internship programs had been dwelling on another problem – the social value of internships and whether it creates an ill-inspired trend among employers to abuse their interns for menial tasks without providing them with any real workplace knowledge or higher grade skills. In two newspaper articles by Asoka (2014) and Greenhouse (2010) taking up an internship is compared to slavery and the legality of unpaid internships is disputed. On the other hand, there are also public opinions that internships are a valuable model for the society and the experience gained prepares interns for their future permanent job by teaching them workplace ethics, skills and giving them an orientation into what they might be interested in (Korn, 2013). Although this paper will not focus on the legal dispute or societal value, it will attempt to show a different perspective on internship programs. As mentioned above, both media coverage and academic research take interest in internship programs, but their focus shifts away from the individual and onto institutional factors whether it be the government or the specific employer that offers the internship. This lack of concentration on the interns themselves leads to two problems – first, there's a unsychronized public opinion on whether internship programs as a whole are for the benefit for the interns; and second, academic research on what the best practices are is very segmented and it is difficult to create a single theory or even build on previous research from other professional fields.

For these reasons, the aim of this paper is to present an individual-focused approach to internship structure based on skill acquisition theory. A microeconomic model is developed to compare interns by their previous experience and learning abilities and predictions are made for the most appropriate supervision time allocation based on task-specific factors such as the levels of complexity and required quality. We find that students who have previous experience receive more supervisory attention, but only if their tasks during the internship are relatively complex. If the tasks they have are routine, theirprevious experience will lead to less supervisory attention. Also, we find that when interns are faced with a complex task, the amount of supervision they receive increases as the internship progresses.

We are now going to present a framework borrowed from cognitive science research. In section 2, a model will be developed based on this framework. In sections 3 and 4 the survey results will be presented and tested for the predictions of themodel. Section 5 contains discussions of the implications and limitations of this study, including recommendations for future research.

Section 1: The ACT Cognitive Architecture Framework

In his famous work Anderson (1983) presents a revised framework of logic-bound interactions between stimuli and responses that aims to explain how cognition works and what characteristics describe the human thought process. In an earlier article (Anderson, 1982), he presents the implications of this theory for skill acquisition and divides the improvement in cognitive skill in two separate categories: algorithmic improvement and strengthening. We are now going to describe in simple words the ACT framework, what the two terms mean and how they can be transferred to a basic microeconomic model.

The ACT Framework

The ACT framework is a very broady defined attempt to describe the mechanics of the brain. Anderson distinguished between three types of memory: declarative, procedural and working. Declarative memory is the one, where assossiations are stored – for instance thinking of ice cream can evoke assossiations of sensations such as cold, sweet, etc., but also of previous experiences, reminiscences, etc. In order to define procedural memory, it is best to first define what a production means in Anderson's works. A production is a procedural instruction that contains a condition and a respective action, for instance when a student is given the mathematical problem "2+2=?", the "?" is a trigger for the procedure of finding the result and the "+" sign is the trigger to use arithmetic sum as a method to find that result. Given this explanation it is rather obvious that procedural memory is the memory where all procedures are stored with their conditions and respective actions. However, it is important to note that different procedures have different times of evocation and only the fastest ones come into action. Evocation time of procedures improves with frequency of usage. Finally, working memory is the part of the two other types of memory which is in action when an cognitive activity is undertaken. Having defined the three types of memory, we are ready to move on to the terms from the skill acquisition framework.

Strenghtening

Anderson defines strengthening as a decrease in the time of evocation of certain procedures based on the frequency of usage and their results. For instance, the more geometric problems a student solves, the faster he can observe that a specific theorem is applicable to the given problem. Also, even when not used with great frequency, if a procedure always yields the required results, it will be strengthened. However, there's a backlash – procedures that give different results from what is expected or ones that are rarely used, tend to weaken and their evocation time becomes higher, so they are suppressed by other productions.

Algorithmic Improvement

In the same skill acquisition framework Anderson defines another way to improve a skill – by combining procedures. For instance when given the problem "2+2=?" at first a student will first recognize the question mark and then the plus sign and iterate through the two productions (first, "?" evokes looking for a result; second, "+" evokes the usage of an arithmetic sum). However, with many repetitions of this sequence of productions, the two will be combined to one which is triggered by both conditions and evokes both responses simultaneously ("?" AND "+" evoke looking for a result using an arithmetic sum). Combining procedures is what Anderson calls algorithmic improvement.

Task Types

Before we continue with the model, it is important to differentiate for the complexity of the task given. Relatively simple tasks will have a relatively smaller number of productions, which implies that algorithmic improvement will be impossible after a relatively smaller number of improvements. This, however, would not affect strengthening.

Section 2: The Model

First, let's start with the basic notations: the internship lasts t periods. We will denote the intern's skill with α_i for period i, where $0 \leq i \leq t, i \in \mathbb{Z}$. We assume α_0 as a given. The skill level of the boss is α_b. In each period each of them produces as many units as his skill.

The boss can spend a portion of his time in every period to teach the intern. These will be denoted by l_i for period i with $l_i = [l_{min}; l_{max}]$, where l_{min} is the minimum time he can spend per period and l_{max} is the maximum.

We assume the skill employed during the internship to be cognitive, rather than physical. If this is the case, we can apply Anderson's ACT architecture framework to investigate how the intern's skill is affected throughout the programme. We are first going to look into the two effects of skill increase separately and then combine the findings.

Strengthening

In his paper on skill acquisition Anderson first defines the time that a production takes to be carried out as T_j. He then specifies the strengthening effect as a decreasing function[1] $T_j = c + aP_j^{-1}$, where c, a and P_j are respectively – the time that is needed for a production to be evoked, processes in applying the production and the number of times the production has been applied until period j. P_j can be measured by the times the task has been completed since the start of the internship or in mathematical terms - $\sum_{i=1}^{j-1} \alpha_i$. If we assume that the time for the production to be carried out is fairly low – $c \approx 0$, then we obtain $T_j = \frac{a}{P} = \frac{a}{\sum_{i=1}^{j-1} \alpha_i}$. The units the intern produces in each period are inversely proportional to the time it takes to complete a cognitive production. To translate this to mathematical terms we assume $T_j \alpha_j = Q$, where Q is some constant. $\alpha_j = \frac{Q}{T_j} = \frac{Q \sum_{i=1}^{j-1} \alpha_i}{a} = \frac{Q}{a} \alpha_{j-1} + \frac{Q \sum_{i=1}^{j-2} \alpha_i}{a} = \frac{Q}{a} \alpha_{j-1} + \alpha_{j-1} = \alpha_{j-1} \left(1 + \frac{Q}{a}\right) = \alpha_{j-1} \delta$, where $\delta = 1 + \frac{Q}{a} > 1$ is a constant.

Algorithmic Improvement

Anderson defines the algorithmic improvement in skill mathematically as a reduction in the number of cognitive productions necessary for the accomplishment of the task. The formula is very similar to the skill increase due to strengthening, but the underlying effects are different. If by N_j we denote the number of cognitive productions necessary, we have $N_j = N_0 f^m P_j^{-1}$, where N_0 is a constant defined as the initial number of cognitive productions necessary for task comlpetion and f is some fraction that describes the decrease in cognitive production number from one algorithmic

[1] In Anderson's paper this form of the function is not the most developed one. He later includes the forgetting of productions (weakening) in his analysis. This effect, however, is assumed away for the model in this paper, because internships usually last only a few months and the weakening effect is negligible.

improvement to the next. m is also a constant in Anderson's model, but it contains several variables, among which the amount of working memory acivation. If we assume that tutoring by the boss decreases the amount of memory activation necessary for the accomplishment of a cognitive production number decrease, we get $\frac{dm}{dl} > 0$. Again we assume the number of steps necessary to be inversely proportional to the skill level $\alpha_j N_j = R$, where R is some constant. Rewriting for α_j, we get $\alpha_j = \frac{R}{N_j} = \frac{RP}{N_0 f^m} = \frac{R\sum_{i=1}^{j-1}\alpha_i}{N_0 f^m} = \frac{R}{N_0 f^m}\alpha_{j-1} + \alpha_{j-1} = \alpha_{j-1}(1 + \frac{R}{N_0 f^m})$, which is very similar to the increase pattern with observe with strengthening. However, since $\frac{da}{dl} > 0$ and f is a fraction, we can approximate $1 + \frac{R}{N_0 f^m} = 1 + l_j k$ (A formal proof for this relationship can be found in Appendix A). Now we get $\alpha_j = \alpha_{j-1} + l_j k \alpha_{j-1} = \alpha_{j-1} + l_j k_j$, where k_j is defined as $k_j = k\alpha_{j-1}$. For simplicity, $k_j$will be refered to as the intern's absorption rate of being taught[2].

The intern's skill increases in every period by a constant factor δ due to the strengthening effect and by a non-costant factor depending on the teaching he receives and his current absorption rate.

Then, the intern's inter-period skill level function can be given

$$\alpha_i = \alpha_{i-1}\delta + k_i l_i = \alpha_0 \delta^i + \sum_{j=1}^{i} k_j l_j \delta^{i-j}$$

Now, we can give the total production of the intern over t periods.

$$Y_i^t = \sum_{i=0}^{t} \alpha_i = \alpha_0 \frac{\delta^{t+1} - 1}{\delta - 1} + \sum_{i=1}^{t} (k_i l_i \frac{\delta^{t-i+1} - 1}{\delta - 1})$$

We can also derive the total production ot the boss

$$Y_b^t = \sum_{i=1}^{t} (1 - l_i)\alpha_b = t\alpha_b - \alpha_b \sum_{i=1}^{t} l_i$$

Summing the two together we get total production

$$\begin{aligned} Y = Y_i^t + Y_b^t &= \alpha_0 \frac{\delta^{t+1} - 1}{\delta - 1} + \sum_{i=1}^{t} (k_i l_i \frac{\delta^{t-i+1} - 1}{\delta - 1}) + t\alpha_b - \alpha_b \sum_{i=1}^{t} l_i \\ &= t\alpha_b + \alpha_0 \frac{\delta^{t+1} - 1}{\delta - 1} + \sum_{i=1}^{t} (l_i (k_i \frac{\delta^{t-i+1} - 1}{\delta - 1} - \alpha_b) \end{aligned}$$

The first terms of this sum are constant and don't depend on the interperiod choice of l_i. The last term, however, describes the effect of education on the intern's skill level and the time loss for the

[2] Actually, k is a constant that describes memory activation for certain cognitive procedures that lead to algorithmic improvements, but since we assume teaching to reduce the working memory necessary for these improvements, k can be considered a modifier on teaching, or, put otherwise, the absorption rate.

boss depending on how much time he chooses to devote to teaching in different periods. We are now going to examine two cases depending on how complex the task is.

Case 1: Routine tasks

In this case we assume that the tasks given to the intern are routine, such as filling in documents, sending e-mails, answering phone calls, etc. When the task at hand is relatively simple, we expect that the numer of productions necessary for completion will be relatively lower. This consideration implies that algorithmic improvement will be limited or, put othwerwise, teaching the intern will be inefficient after a certain level of skill is reached, because the intern has complete knowledge of the task he is given. He will still reduce the time it takes for him to accomplish each of the already algorithmically perfected productions with each consecutive trial, so we expect that strengthening will still apply.

Another important consideration for routine tasks is that we are more likely to expect that higher quality is appreciated, but not demanded. In terms of the model, this means that there will be a level of total production that the boss will want to be completed and he wouldn't put effort into increasing it further.

Both of these considerations will lead us to one result – we expect that the boss will put in teaching effort only until it has any effect or until he is certain that a level of production is achieved and not further. Also, for routine tasks we expect to find that the initial skill level of the intern, α_0 and the amount of teaching he receives are negatively related, because when a fixed amount of production is required, when the intern is already good at producing, he will receive less training. The same logic applies for k, the learning absorption rate. The more the intern can perceive, the less the boss will teach him.

Unfortunately, because k_i depends on $l_j \forall j < i$, we cannot make predictions on specific allocations in different periods. We will be able to do that in the next case, where the task is considered complex.

Case 2: Complex tasks

When the task given is rather complex, we expect that algorithmic improvements will be possible throughout the internship and so, teaching effort will be effective in every period. What's more, complex tasks are usually associated with higher quality expected, so in this case the boss will want to maximize total production instead of having it reach a threshold level. Now we can ignore the constant part of total production and focus on the effects of teaching:

$$\sum_{i=1}^{t}(l_i(k_i\frac{\delta^{t-i+1}-1}{\delta-1}-\alpha_b)$$

Our first observation is that when $k_i > \frac{\alpha_b}{\frac{\delta^{t-i+1}{}_{-1}}{\delta-1}}$ the boss will achieve maximum profit from teaching $l_i = l_{max}$. Note that this doesn't mean that he will not profit if he teaches the maximum amount when the inequality doesn't hold. We are now going to examine the relationship between the terms of the inequality and what it means for teaching effort allocation. In order to do that we will rewrite the constraint using the definition $k_j = k\alpha_{j-1}$:

$$\alpha_{i-1} > \frac{\frac{\alpha_b}{k}}{\frac{\delta^{t-i+1}-1}{\delta-1}} = \frac{\alpha_b(\delta-1)}{k(\delta^{t-i+1}-1)}$$

In order to understand better when maximum teaching will start or end, it is beneficial to take a look at the functions on both sides of the inequality and see when they intersect, as these points will be either the start or the end of a few periods with maximum teaching.

We are now going to examine the right hand side. Let's denote $g(i) = \frac{\alpha_b(\delta-1)}{k(\delta^{t-i+1}-1)}$. $g(i)$ is an inverted hyperbolic function over i. This means that

(1) $\lim_{i \to t+1} g(i) = +\infty$

And

(2) $\frac{d^2 g(i)}{di^2} > 0 \; \forall i > 0$

Before we continue, it is important to notice that because $g(i)$ approaches infinity for the last periods of the internship, it will not be profitable for the boss to expend time on teaching then. This result is very natural, as the boss benefits from his teaching in the periods after he teaches and when those periods are decreasing in number, it is more profitable for him to also decrease the amount he devotes to teaching[3].

As for the left hand side, if we assume that the intern received no teaching whatsoever ($l_j = 0 \forall j$), we obtain $\alpha_i = \alpha_{i-1}\delta + k_i l_i = \alpha_{i-1}\delta = \alpha_{i-2}\delta^2 = \alpha_0 \delta^i$, which is a power function over i. If we assume that he received maximum teaching in all periods instead ($l_j = l_{max} \forall j$), we obtain $\alpha_i = \alpha_{i-1}\delta + k_i l_i = \alpha_{i-1}(\delta + k l_{max}) = \alpha_{i-2}(\delta + k l_{max})^2 = \alpha_0(\delta + k l_{max})^i$, which is also a power function over i. Of course, generalizing for l_i will not give us any real results, but considering that α_i is limited between two power functions over i, we can assume that it's graph will resemble a power function over i. Any power function with a base bigger than 1 and a strictly positive exponent has a strictly positive second derivative as well, so using the assumption that α_i resembles a power function we can obtain

(3) $\frac{d^2 \alpha_i}{di^2} > 0 \; \forall i > 0$

Combining (2) and (3), we can deduce that α_i and $g(i)$ have at most 2 intersections. We are now going to examine the three scenarios based on the number of intersections. Let's denote the two intersection points (if they exist) with i^* and i^{**}, with $i^{**} > i^*$.

Case2.1: no intersection points

[3] The boss does receive some reimbursement for his teaching already in the same period, but it's smaller than what he could have produced himself. The real benefit he receives from the increase of skill in the following periods.

Using (1), we can observe that in order for α_i and $g(i)$ to have 0 intersection points, the following must hold

(4)$g(i) > \alpha_i \; \forall i$

This implies that in no period does the intern's skill pass the threshold and the boss's time is too valuable to waste for providing maximum teaching. Although, in this scenario the constraint for maximum teaching is never reached, or put differently, the quotient for each l_i is negative, we can still observe that it is increasing over i. This implies that although maximum teaching cannot be expected, we can still predict that some teaching will occur with increasing effort put in by the boss as i increases.

Case2.2: one intersection point

Using (1) and the assumption of one intersection point, we can see that α_o already satisfies the maximum teaching constraint[4] and the intersection point (i^*) depends on α_o and k as follows:

$$\frac{di^*}{dk} > 0 \text{ and } \frac{di^*}{d\alpha_0} > 0$$

This means that the larger α_o and k are, the more periods of maximum teaching there will be.

Case2.3: two intersection points

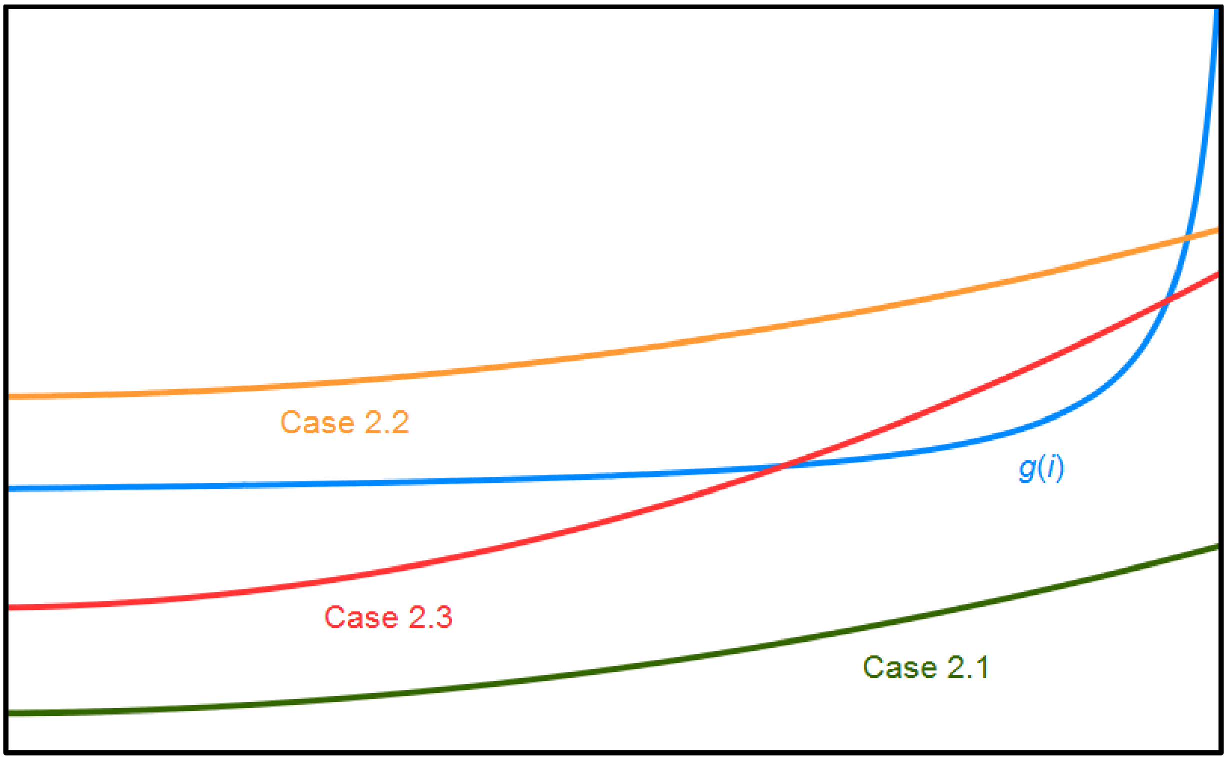
Using (1) and the assumption of two intersection points, we can observe that α_o doesn't satisfy the constraint, however $\alpha_i \forall i \in (i^*; i^{**})$ does. There will be maximum teaching for $i^{**} - i^*$periods. Also,

$$\frac{di^*}{dk} < 0 \text{ and } \frac{di^*}{d\alpha_0} < 0$$

$$\frac{di^{**}}{dk} > 0 \text{ and } \frac{di^{**}}{d\alpha_0} > 0$$

This gives us the same conslusion as in Case2.2: the larger α_o and k are, the more periods of maximum teaching there will be.

[4] If $i^*$is the point where maximum teaching begins, we get a contradiction with $\lim_{i \to t+1} g(i) = +\infty$, so i^* should be the point where maximum teaching ends. This implies that maximum teaching is beneficial already from the start at period 0.



Summary of case 2

With regards to initial skill level (α_0), we expect to find that interns who have higher initial skill levels will start receiving more coaching at the start of their internship as compared to other interns (case 2.2). Furthermore, the interns who have higher skill level will receive high training for longer if their absorption rates are higher.

We also expect to find that interns with relatively low initial skill will receive lower amounts of training at first and their training later on in the program increases or not depending on whether their absorption rate is high or low.

Summary of model predictions

Although the model presents its findings in terms of maximum teaching, the concept of a maximum portion of time the boss can spend on teaching is very static and unrealistic. For this reason, only the general relationships predicted by the model will be taken into account. Instead of maximum teaching we would expect to find relatively more teaching under the conditions predicted by the model. If the task at hand is relatively simple (routine), we expect to find that interns with higher initial skill and learning absorption rates receive lower amounts of training. However, if the task is more complex we expect to find the exact opposite relationship – more talanted interns (in terms of initial skill and absorption rate) will receive more attention.

In order to test the validity of the model, we will present four hypotheses on the general predicted relationships and another two for the specific findings for the allocation of teaching effort in the complex task scenario. They will then be tested against data collected by means of an online survey.

Hypotheses

The effects of intial skill level

H_1: When interns are given routine tasks, the more skilled they are at the start of the internship, the less amount and quality of supervision they will get.

H_2: When interns are given complex tasks, the more skilled they are at the start of the internship, the larger amount and quality of supervision they will get.

The effects of learning abilities

H_3: When interns are given routine tasks, the higher learning abilities they have, the less amount and quality of supervision they will get.

H_4: When interns are given complex tasks, the higher learning abilities they have, the greater amount and quality of supervision they will get.

Specific time allocation predictions for the complex task scenario

H_5: When interns are given complex tasks, they will receive relatively less training later on in the very last periods of the internship.

H_6: When the interns are given complex tasks, they will receive increasing amounts of training as the internship progresses up until thevery last period.

Section 3: Data and Methodology

Data

In order to test our four hypotheses we conduted an online survey among students who have participated in an internship program. The survey was completed by 64 people, who were asked to share the following information:

1) To what extent they considered the tasks given to them routine;
2) Have they had any previous experience that is relatable to their internship in terms of skills applied;
3) Their grade point average (GPA) converted to percent;
4) Subjective scores on different supervision-related factors for each quarter of their internship experience.

Because of the limited resources and scope of this paper, the measured variables are in some aspects different than the dependent variables of the model predictions. These limitations and advice for correcting them will be discussed in the conclusions section. We are now going to explain the question choices and how the information was converted to testable variables. The survey questions can be found in Appendix B

Question 1: Since our hypotheses control for the level of complexity of the task, we asked our subjects to rate how routine they perceived their tasks to be on the scale from 1 to 5 with 1 being not routine at all and 5 being extremely routine. We then encoded the responces with a dummy variable called *routine_dummy* which is 1 when the tasks were rated from 1 to 3 and 0 when the tasks were rated 4 or 5. Other encodings are discussed in Appendix C. The choice of this encoding is natural because as soon as the task is not completely routine, the number of production necessary to complete it increases dramatically.

Question 2: One of the predictors in our model is the initial skill level α_0. Since the resources available for conducting this survey are extremely limited and tasks across internships quite diverse, there is no other reliable way to measure initial skill level than to ask if the subjects had any previous relatable experience and to what extent it was related skill-wise. Respondents were asked to rate their previous professional experience from 1 do 5 with 1 being no experience whatsoever and 5 being having been given the exact same tasks before. It is interesting to note that not a single person responded with a 5 for this characteristic. Again we recoded this variable into a dummy called *experience_dummy* where 1 is no or non-relatable experience (values 1 and 2) and 0 is relatable experience (values 3 and 4).

Question 3: We asked our respondents to give their grade point average as a percent, which will be used to map their respective learning absorption rates. Although it is true that grades are a product of learning talent combined with effort, we have no feasible way of measuring learning absorption rates or controlling for the effort that respondents have put in their academic development, so this is the best approximation we can get. This variable is called *academic_average*.

Question 4: In order to make judgements about different periods in the internship we asked the respondents to report on each of four quarters of their internship. The factors analyzed are the amount and quality of specific task information, formal training and feedback, giving us a total of 6 factors. All factors are rated from 1 to 5 with 1 being low and 5 – high. In order to get a general score from the 6 factors, we multiplied the amount and quality factors for each of the three categories of supervision and then summed the three products together. The variables obtained are the general scores for each of the four periods – *period1score*, *period2score*, etc. The choice of construct for the general score will be discussed in Appendix D.

Methodology

In order to analyze the first four hypotheses for the general predictions of the model, we are going to subsequently regress the general scores for each period first on the variables *routine_dummy* and *experience_dummy* as well as an interaction term for hypotheses 1 and 2 and then on the variables *routine_dummy* and *academic_average* and an interaction term for hypotheses 3 and 4.

As for hypotheses 5 and 6, we are going to only select the cases with a complex task and analyze the data through means comparison for each of the periods.

Section 4: Results

First, the analysis of hypotheses 1 and 2 will be presented, namely the effect of previous experience on the subjective scores of periods 1 through 4, controlling for the routineity of the task. The results of the regressions are summarized in the following table.

		Period1score		Period2score		Period3score		Period4score	
1	R^2	,115		,119		,021		,037	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	20,435	,000	25,130	,000	34,000	,000	30,522	,000
	routine_dummy	8,638	,006	9,089	,005	3,854	,258	6,308	,129
2	R^2	,245		,419		,274		,085	
		B	Sig	B	Sig	B	Sig	B	sig
	C	16,000	,000	19,200	,000	28,600	,000	29,000	,000
	routine_dummy	17,409	,000	22,209	,000	16,036	,000	11,727	,053
	experience_dummy	7,846	,094	10,492	,015	9,554	,050	2,692	,683
	r_x_interaction	-17,203	,004	-26,007	,000	-24,190	,000	-11,104	,180

routine_dummy is coded as follows: 1 for complex tasks and 0 for routine tasks.

Experience_dummy is coded as follows: 1 for no or unrelatable experience and 0 for relatable previous experience.

In both models and across periods we can see that the coefficient for *routine_dummy* is positive, which means that as a whole the interns who had complex tasks received more training. When also controling for experience, we can see that the variance explained is higher, which coupled with the significant coefficients for periods 1 to 3, gives us a good reason to believe that previous experience is, indeed, a valid predictor of received training during the internship.

Now, to analyse the coefficients – for hypothesis 1, where the routine task is examined, we only have to take into account the coefficient for *experience_dummy* and it is positive for all 4 periods. This means that when the routine task is examined, interns with no or unrelatable experience received relatively more training than the ones who had such experience. Thus, we can confirm our first hypothesis.

As for the complex task scenario, in order to obtain the effect of having little or no previous experience on the supervision quality and time, we have to sum the coefficients of *experience_dummy* and the interaction effect (*r_x_interaction*). We can see that in all 4 periods, these sums are negative, which means that when interns are given complex tasks, they will receive more supervision if they had any experience in the past that resembled their current tasks, which confirms hypothesis 2.

Now, to analyse H_5 and H_6, namely the effects of the average academic grades, controlling for the routineity of the task. The results of the regression are summarized in the following table.

		Period1score		Period2score		Period3score		Period4score	
1	R^2	,166		,121		,030		,037	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	41,326	,000	29,541	,015	24,550	,057	29,571	,060
	routine_dummy	9,287	,003	9,226	,005	3,560	,301	6,278	,136
	grades	-,264	,059	-,056	,703	,119	,447	,012	,950
2	R^2	,190		,187		,229		,044	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	43,352	,000	31,298	,009	80,891	,000	33,262	,039
	routine_dummy	6,669	,064	7,913	,035	-87,691	,000	,795	,873
	grades	-,283	,051	-,079	,594	-,593	,012	-,039	,846
	r_grade_interaction	,062	,154	,066	,145	1,140	,000	,080	,189

Again, *routine_dummy* is coded 1 for complex task and 0 for routine task. The obtained results are very peculiar – we have a significant effect of grades on the general scores only for period 1 and 3 with the interaction term decreasing the explanatory power for period 1, but increasing it dramatically for period 3. The effects on periods 2 and 4 are highly insignificant. The fact that out of the 8 regressions we ran, only 2 return significant results gives us enough evidence to reject hypotheses 3 and 4. However, the highly significant results for the regression on *period3score* gives us some support for the findings of the model – the coefficients for the variables *grades* and the interaction term show exactly the relationship described by hypotheses. The sum of the two coefficients is positive, which means that when the task is complex, higher grades are associated with higher levels of teaching. The negative coefficient for the variable *grades* show us that when the task is routine, grades have a negative effect on the general supervision score. This partial support for the model gives us some reason to not disregard the effect of grades on the general scores. Considering that academic grades might not be the best predictors for training absorption, further recommendations will be given in the discussion section.

Now, we can continue with hypotheses 5 and 6. In order to compare scores for the complex scenario across periods, we will conduct a paired-samples T-test. The results obtained are summarized below.

	Mean	Std. Deviation	Sig.
period1score	29,0732	12,84794	-
period2score	34,2195	14,00091	-
period3score	37,8537	14,07224	-
period4score	36,8293	15,68583	-
Period1score-period2score	-5,14634	10,04878	,002
Period2score-period3score	-3,63415	12,46546	,069
Period3score-period4core	1,02439	17,43630	,709

By examining the last three means, namely of the three differences between consecutive periods, we can observe that the scores differ significantly between periods 1, 2 and 3. Also, the coefficients for both differences are negative, which implies that general scores grow over periods 1,2 and 3. Thus, hypothesis 6 is confirmed. As for the last periods difference, the observed mean is not significantly large, so the two periods' scores do not differ significantly. Even though that the coefficient is insignificant, it is still positive, which implies that period 4 scores are smaller than period 3 scores, which is in line with hypothesis 5. Our conclusion is that we do not have enough evidence to reject hypothesis 5.

Summary of hypotheses results

	H_1	H_2	H_3	H_4	H_5	H_6
Result	Confirmed	Confirmed	Rejected	Rejected	Inconclusive	Confirmed

Section 5: Dicussion

The results and implications for interns and supervisors

We found out that the most supervision effort goes to interns who have had previous experience and are dealing with complex tasks with that effort increasing over time. For interns, this means that if they want to maximize the received training, they should try to communicate as much as possible their previous experiences to their employer and take on as difficult tasks as possible. Another advice is that they should also enroll in longer internship programs, as the effects of time on the received supervision are cumulative. As for supervisors, if they want to create an internship program, which produces highly effective professionals, they should also comply with the findings of our model – by giving interns more complex tasks, trying to relate to their previous experience as much as possible and give them more training as the internship progresses.

The methods and implications for further research

We are going to continue the discussion with outlining the strengths and weaknesses of the developed model and make suggestions based on both. The first and most important strength of the model is that it is applicable across many different professional fields. The approach to consider the internship experience as increases in a single cognitive skill allows for many improvements on existing research. First, by applying the methods of our research activities within a single internship program could be compared based on what qualities the interns had before they started them. Second, our theory can be used as a bridge so experiences could be compared between different internship programs in order to escape from the continuous limited outlining of best practices within one field or another and expanding the toolkit of the internship ogranizer with the variety of methods other internship programs use. The model we have developed, however, also suffers from its generality – as supervisors are nearly impossible to compare between different fields and even internship programs within the same field, we suggest that this approach be tested on a single internship program with many participants under one supervisor. By comparing different student under the same supervisor this weakness can be removed as a factor in the analysis.

The analysis and implications for future research

In our analysis we used an online survey with self-reported values, which makes the reliableness of our findings questionable, especially considering the small sample. Apart from conducting this same survey in a more strictly regulated environment, another suggestion for improvements is to expand the measurements for both relatable previous experience and training absorption rates by conducting a longitudinal or panel study instead of a cross-sectional one. As far as relatable experience is conserned, this expansion can be achieved by giving interns different measurable tasks to complete at the start of the internship and map the initial skill levels based ontheir performance. As for the training absorption rates, the most important factor that is to be measured is the working memory activation levels. Maybe it would be possible to design a task that measures the working memory activation levels for interns at the beginning of the internship, but such a task design is not within the scope of analysis in this paper.

New perspective on the social issue

As mentioned in the introdution, there has been an ongoing public debate on whether itnernships are the path to teach sustainable working experience both for the benefitof the company and the intern himself. The answer to this question is both yes and no – as our findings show, interns receive different amounts of supervision depending on both their own experiences so far and the organizational structure of the internship (whether the tasks are routine or not). For instance, a routine internship would be of little use to someone who already has the skills needed to accomplish simple administrative tasks. However, interns with no initial skills would benefit from such a choice. On the other hand, when the task is complex these same interns with little initial skills will receive little to no supervision as they would be unable to apply themselves in a way that a complex task would require. Now, to apply these contentions to the debate – in defense of the thesis that internships are a valid way to increase intern's skills we presented the conditions in which interns do receive proper supervision and their tasks are not routine; however, our model also presents some support for the argument that interns are being abused – if a highly skilled intern ends up doing routine tasks, his potential is wasted. On top of that, we found another scenario – when a low-skilled intern is faced with a very complex task, he will waste both the company's and his own time and resources. In conclusion, the reality that's being debated on is not depicted accurately by neither side. Rather, both sides have their arguments supported by individual cases, and it is single-minded to claim that internships as a whole are very efficient or not, because of the multiple differences between individual internship programs.

Appendices

Appendix A

Formal proof of $1+\frac{R}{N_0 f^m}=1+l_j k$

W	Size of working memory needed for the ith algorithmic improvement defined as $W=GH^i$
G	Linear parameter for W
H	Exponential base parameter for W
Q	Parameter for total memory activation (not dicussed here)
L	Threshold level of memory activation for a given amount of information to be available in working memory

In our model we approximate $1+\frac{R}{N_0 f^m}=1+l_j k \Leftrightarrow f^m=\frac{R}{N_o k l_j} \Leftrightarrow m=\log_f \frac{R}{N_o k l_j}=$
$\log_f \frac{R}{N_o k} - \log_f l_j = \log_f \frac{R}{N_o k} - \frac{\log_H l_j}{\log_H f}$

From Anderson (1982) $m=\frac{\log Q}{\log H}-\frac{\log L}{\log H}-\frac{\log G}{\log H}=\log_H \frac{Q}{L}-\log_H G$

If we assume that teaching decreases the linear parameter for the necessary working memory for an algorithmic improvement, we can obtain that the decreased $G^*=GY$, where Y is some fraction. Let the changed m be denoted with m^*.

$$m^*=\log_H \frac{Q}{L}-\log_H G^*=\log_H \frac{Q}{L}-\log_H G-\log_H Y=m-\log_H Y$$

Because further improvements require more working memory, we can deduce from $W=GH^i$ that $H>1$. As a fraction, $f<1 \Rightarrow \log_H f<0 \Rightarrow -\frac{\log_H l_j}{\log_H f}=-x\log_H l_j=-\log_H l_j^x$, where $x=\frac{1}{\log_H f}$ is a positive constant. We can see now that all that is required for our assumption to be valid is that $-\frac{\log_H l_j}{\log_H f}=-\log_H Y \Leftrightarrow Y=l_j^x$, which is reasonable to assume, since Y, l_j and x are all fractions.

Appendix B

Survey questions:

1. How routine would you say your tasks were during the itnernship?
    a. I had no routine tasks whatsoever (1)
    b. A few of my tasks were routine (2)
    c. Some of my tasks were routine (3)
    d. Most of my tasks were routine (4)
    e. I only had routine tasks (5)
2. If you had had previous experience in the field of your internship, how closely were the set of skills related to your latest internship tasks?
    a. I had not had previous experience
    b. I had experience but it was almost entirely different
    c. It did have some common points, but otherwise quite different
    d. My previous experience was similar to the internship
    e. I had the exact same tasks as before
3. What was your average grade out of 100 in your last completed or ongoing education before the internship started?
4. Please rate the following factors of supervision for each quarter of your internship

The same checkbox table was given for each quarter.

	Very low	Low	Average	High	Very high
Amount of task-related information					
Quality of task-related information					
Amount of formal training					
Quality of formal training					
Amount of feedback					
Quality of feedback					

Appendix C

Robustness of *routine_dummy*

If we encode the perceived routineity as *routine_dummy1* with 1 for values 1 and 2 and 0 for values 3, 4 and 5, and keep other variables the same we obrain the following regression statistics when analysing the first two hypotheses.

		Period1score		Period2score		Period3score		Period4score	
1	R^2	,127		,158		,029		,016	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	22,930	,000	27,442	,000	34,930	,000	33,186	,000
	routine_dummy1	9,260	,004	10,701	,001	4,689	,177	4,195	,326
2	R^2	,234		,353		,223		,079	
		B	Sig	B	Sig	B	Sig	B	sig
	C	22,727	,000	28,182	,000	34,909	,000	33,636	,000
	routine_dummy1	16,773	,000	20,118	,000	15,091	,001	10,964	,071
	experience_dummy	,416	,902	-1,515	,638	,043	,990	-,922	,847
	R1_x_interaction	-14,370	,018	-17,876	,002	-19,861	,002	-12,860	,128

We can observe that by manipulating the routine_dummy variable, the model has lost its significance in its predictions for the variable experience_dummy. This effect occurs because the model predicts opposite results when controlling for routineity (H1 against H_2). A slight change in the mode of control could mean that results from the two categories are being considered together, leading to the insignificant results.

Appendix D

To construct the scores as the sum of products between the amount and quality of the three factors is a very intuitive choice. We will here show that other constructs for the score will obtain the same results.

		Period1score1		Period2score1		Period3score1		Period4score1	
1	R^2	,114		,107		,013		,054	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	15,217	,000	17,130	,000	19,870	,000	18,217	,000
	routine_dummy	2,880	,006	2,650	,008	,935	,375	2,319	,065
2	R^2	,256		,416		,234		,101	
		B	Sig	B	Sig	B	Sig	B	sig
	C	13,400	,000	15,100	,000	18,200	,000	17,600	,000
	routine_dummy	6,100	,000	6,855	,000	4,527	,002	4,082	,026
	experience_dummy	3,215	,040	3,592	,007	2,954	,056	1,092	,582
	r_x_interaction	-6,242	,002	-8,284	,000	-7,102	,000	-3,564	,153

All the regressions we run for the analysis of hypotheses 1 and 2 return the same relationships between coefficients and are significant for the same periods.

		Period1score1		Period2score1		Period3score1		Period4score1	
1	R^2	,167		,109		,021		,054	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	22,317	,000	18,304	,000	17,161	,000	17,609	,000
	routine_dummy	3,100	,003	2,686	,008	,851	,424	2,300	,071
	grades	-,090	,055	-,015	,743	,034	,482	,008	,894
2	R^2	,214		,120		,073		,098	
		B	Sig	B	Sig	B	Sig	B	Sig
	C	16,843	,001	15,800	,001	22,744	,000	23,819	,000
	routine_dummy	5,154	,001	3,626	,016	-1,244	,424	-,030	,987
	grades	-,021	,723	,017	,773	-,036	,555	-,071	,333
	r_grade_interaction	-3,144	,062	-1,438	,390	3,207	,072	3,566	,091

Hypotheses 3 and 4 are rejected for this construct as well.

	Mean	Std. Deviation	Sig.
period1score1	17,0625	4,12070	-
period2score1	18,8281	3,90992	-
period3score1	20,4688	4,01176	-
period4score1	19,7031	4,83269	-
Period1score1-period2score1	-1,76563	3,14083	,000
Period2score1-period3score1	-1,64063	3,52482	,000
Period3score1-period4score1	,76563	5,40647	,262

The same goes for the analysis of hypotheses 5 and 6.

Bibliography

Anderson, J. R. (1982). Acquisition of Cognitive Skill. *Psychological Review, 89*(4), 369-406.

Anderson, J. R. (1983). *The Architecture of Cognition.*

Asoka, K. (2014, May 7). Interns aren't just cheap labor to abuse: they're workers – and they deserve pay. *The Guardian*. Retrieved from http://www.theguardian.com/commentisfree/2014/may/07/unpaid-internships-unfair-cheap-labor-abuse

Benjamin, L. T. (2001). American Psychology's Struggles With Its Curriculum: Should a Thousand Flowers Bloom? *American Psychologist, 56*(9), 735-742.

Greenhouse, S. (2010, April 2). The Unpaid Intern, Legal or Not. *The New York Times*. Retrieved from http://www.nytimes.com/2010/04/03/business/03intern.html?_r=0

Korn, M. (2013, June 5). Internships Are Increasingly the Route to Winning a Job. *The Wall Street Journal*. Retrieved from http://online.wsj.com/news/articles/SB10001424127887324423904578525431344927240

Nash, J. M., Norcross, J. C., & Stevenson, J. F. (1986). Evaluation of Internship Training: Practices,. *Professional Psychology: Research and Practice, 17*(3), 280-282.

Rodolfa, E. R., Kaslow, N., Stewart, A. E., Keilin, W. G., & Baker, J. (2005). Internship Training: Do Models Really Matter? *Professional Psychology: Research and Practice, 36*(1), 25–31.

Santos, I. d., & Nunes, J. E. (2009). Length of internship influences performance on medical residency exam. *Revista da Associação Médica Brasileira, 55*(6), 744-748.