
Final Project
-The Learning Curve of Amazon-

Jesse Bruner
April 25th, 2018

Overview

I chose to study the learning curve of shopping on Amazon, and I studied this by determining the difference between experienced and inexperienced people and the time they spent performing a difficult purchase. I wanted to study the difficulty of Amazon as I am going into UX Design, which is a design field in tech where I would study users, find their pain points, and design a solution. In this case, I want to determine if Amazon is simple enough for an inexperienced user to perform more difficult tasks. Additionally, I want to know if experienced users, also, can perform more difficult tasks.

My hypothesis is that the experienced users will have an easier time (though, perhaps still a difficult one) finding the product compared to those who are not as experienced with shopping on Amazon. I also hypothesize that inexperienced users will either quickly figure out what to do, or take an extraordinarily long time figuring it out that they may give up.

The possible benefits of this study include 1) finding the difficult tasks involved with shopping on Amazon, 2) determining how long a user will take to complete a task knowing what their experience is, 3) collecting further data on how a shopper uses Amazon at different skill levels. All of these benefits are useful, as we can apply the knowledge learned to better understand how to further design and develop Amazon to make more conversions and sell more products. This study will accomplish those benefits by marking observations and times of the users to complete each task, which can be analyzed for further study.

For this class project, however, performing a full-scale study that would be very useful to Amazon for redesign would be far more complex than necessary. So instead, my focus will be on 1) time spent performing required task, 2) mark each major point in the process where the user got stuck.

Variables and Terms

Before I list my variables that I will measure for this study, I will first define some terms as to clear up any confusion.

- **Amazon** refers to Amazon.com, the largest online retailer on the market today.
- **User** refers to the subject being studied.

- **Filter** refers to the User Interface element located on the side of the screen which the user can use to “narrow down” their search.
- **Task** refers to a portion of the shopping experience where the user needs to perform a function. Finding the search bar, applying proper filters, and determining if a suspected product is the correct one are all examples of a task.
- **Purchase** will refer to the end goal, where the user put the product in the shopping cart and proceeded to checkout. By purchase, I DO NOT mean putting in the credit card information, applying shipping information, and buying the product “for real” but instead, “purchase” will refer to the end goal of the user.
- **Products** refers to every item sold on Amazon.
- **The Product** will refer to the item that the user must “purchase.”

Now, with those terms defined, I will specify which things need to be measured and how I will measure them. My variables are the *average time it takes to complete the purchase for each group* and *pain points in each group*. The time it takes to complete a purchase for each group is a continuous variable, as there is no limit to how long it can take someone to complete the purchase and we are counting in clear seconds. The pain points, however, are discrete. These are discrete as they are qualitative and do not need to be in any order.

The Groups being Studied

I am studying two groups. **Inexperienced** vs. **Experienced**.

An inexperienced user is one with no more than 2 real purchases on Amazon ever.

An experienced user is one with at least 5 real purchases.

The target population is anyone within South-East Michigan who is old enough and has a credit card/debit card to make a purchase on Amazon. My study does not factor in age, sex, other experience with online shopping, education, or any other descriptor that could be added to my population. I took a sample of people living in my local area who have credit cards and can make purchases on Amazon themselves. I did not restrict my study to analyze exclusively, say, students at WCC, but instead to anyone who was available to study. Since I am trying to understand the learning curve on Amazon, restricting my population may tamper with the overall results I want, which is to understand, on average, how all people generally perform. Essentially, by taking my samples from many different

places from any age, this will give me the most random sample to study. Inexperienced users that I studied represented the portion of the entire population that have not made a purchase or many purchases on Amazon where as my experienced sample is the contrary.

I intentionally added a 3-purchase difference between my experienced and inexperienced. This way, I know there is a significant difference in skill level between the two groups.

(A user with 2 purchases on Amazon in the past would make them the most “experienced” inexperienced sample user. Throughout those two purchases they would have learned the basics of shopping and possibly it’s core functions. They may still be clunky and unsure of where everything is. However, a user with 5 purchases has at least **150% more experience**, allowing for much improvement with the service. [$5 - 2 = 3$, which is the difference. $3/2 = 1.5$, or 150%])

It is also worth noting that if a user was unsure of how many times they have shopped on Amazon, but were confident they made more than 5 purchases, I counted them as experienced.

I also included people with no Amazon experience in the inexperienced group.

Both groups will be safe to study as the diversity in people, within the varied locations around, yet within, South-East Michigan, (ranging from Ann Arbor to Canton) the two groups will remain representative of the population and random enough to where it will be a fair comparison of the two groups. The total population I was dealing with is several millions in numbers, so for our purposes I will treat the population as a very large/infinite population. (Depending on what you consider “South-East Michigan” we’re talking anywhere from 3 million to 6 million possible people; and this population is constantly shifting and growing.) I assumed both groups (experienced and not experienced) may or may not have different standard deviations and that these standard deviations are unknown. (Case 3) My hypothesis, again, is that experienced users will, on average, perform the purchase faster than the inexperienced users.

Sampling

I did not use true random sampling. For this study, and for how many samples I needed to inevitably make it normal, I had to go up to strangers and ask them to participate in my study. I did this at WCC,

my church (in Canton), with a few friends, and at an Arby's in Ypsilanti. I did not keep track of how many samples I got from each location.

This was not truly random, as I had to walk up to someone and ask them to join my study. The randomness came from not knowing before hand how much experience they have with Amazon. This is pseudo-sampling. By mixing locations between WCC, Arby's, my church, etc., I was able to get a large variety of people which gave me the most random sample I could find for study. I had no other means of getting a sample, as I needed to be there in person to observe them. The other possibility would be screen sharing, and watching people remotely complete the task, however I decided the difficulty of this was much greater, and the people I could find online who would want to participate would (most likely) be more adept at online shopping and use of Amazon. Where as by finding people from many ages, (I varied from 17 up to 71 in ages studied) it would lessen the chance of finding mostly adept users. I wanted a wide variety of experience with internet and shopping, and finding people in person was the best way to randomize it.

I conducted my study, as said earlier, by walking up to someone and asking them to join. I did it by:

1. Finding someone who is alone and approaching them passively, keeping distance between me and them.
2. I asked them if they would like to participate in a study that I'm conducting for my stats class. I said it wouldn't take too much time, and all they would do is pretend to shop for an item on Amazon.
3. If they agreed, then I asked them about their history of Amazon and how many purchases they made. If they said no, then I said thank you and have a nice day.
4. After asking them, I began to setup the laptop in front of them. I ensured we were on a steady wifi connection to keep the results similar. As I setup, I was asking them questions about how their day is going and what they do for school or work.
5. I gave them my notebook which had the instructions on it. The instructions read: "You want a new laptop for gaming. You have a \$2500 budget and want a 4-year warranty. It must be 17 inches, with Full HD, 1TB of storage, 512gb of Solid State Drive, 16gb of DDR, and a powerful 8gb graphics card which is capable of supporting VR. You also want a Lenovo laptop because you trust their quality. Find any laptop which meets your desired specifications listed

above, and then place it in the cart and proceed to checkout. If you have any questions regarding what something is, you may ask me.”

6. Once they set the notebook to the side, I started the timer. The timer was unknown to them unless they asked if they are timed. If they asked, I would say yes, but I want to make it clear that I am not testing you but instead testing Amazon.
7. When they completed the purchase, I recorded the time in seconds, and noted the point during the purchase that was the hardest for the user.

Independent

The groups being studied (experienced vs. inexperienced) are both independent of each other. Whatever result one sample user achieves will have no impact on another sample user. Each sample user is independent of each other within each group as well. One inexperienced user (User A) could get the purchase complete in 7 minutes, and since the next inexperienced user (User B) I would study has no reference to what User A got, their results will be completely independent.

Both groups (Experienced and inexperienced) are independent. The times they will get will be completely different and not related in the same vein of why users within a group would be independent.

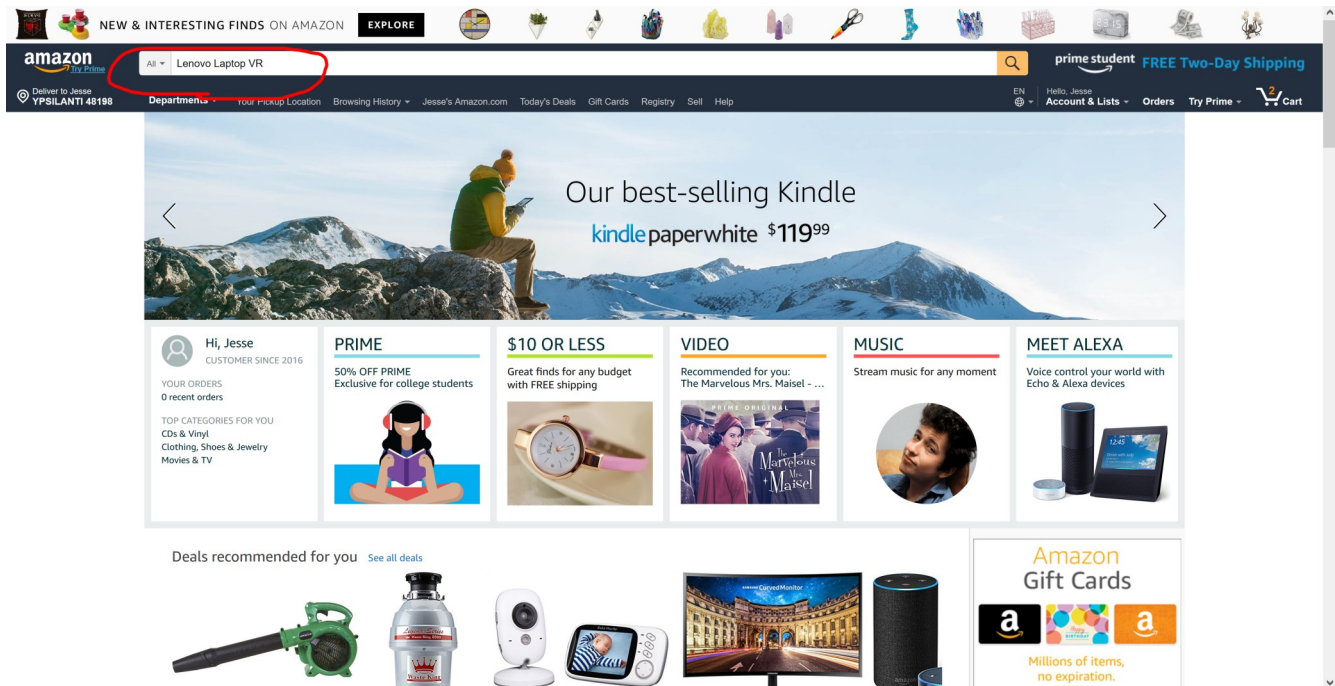
The Study

I gave participants a notebook where I wrote down instructions on what they had to do. Each user had the same notebook handed to them. They read what it said, and once they began to search is when I started the timer. What they had to do was find a Lenovo laptop that is 17 inches, Full HD, with 1TB of storage, 512gb of Solid State Drive, 16gb of DDR, and a powerful 8gb graphics card which is capable of supporting VR. They could hypothetically use any Laptop, however the only one (to my knowledge) which meets those requirements is the Lenovo Legion. The idea is by picking such a hard product to search for, most people would need to know how to search for something they are not familiar with. Especially if they want to meet all of those specs.

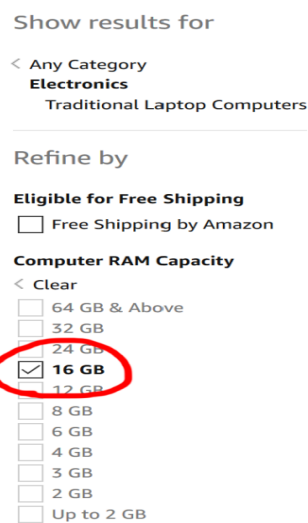
The proper, most efficient way of doing it would be to filter the keywords, and only search for the really important ones. (So search for something like Lenovo Laptop VR) and then put on a 1 TB storage filter, etc. until they find the Legion.

To complete the study efficiently, users would follow this track:

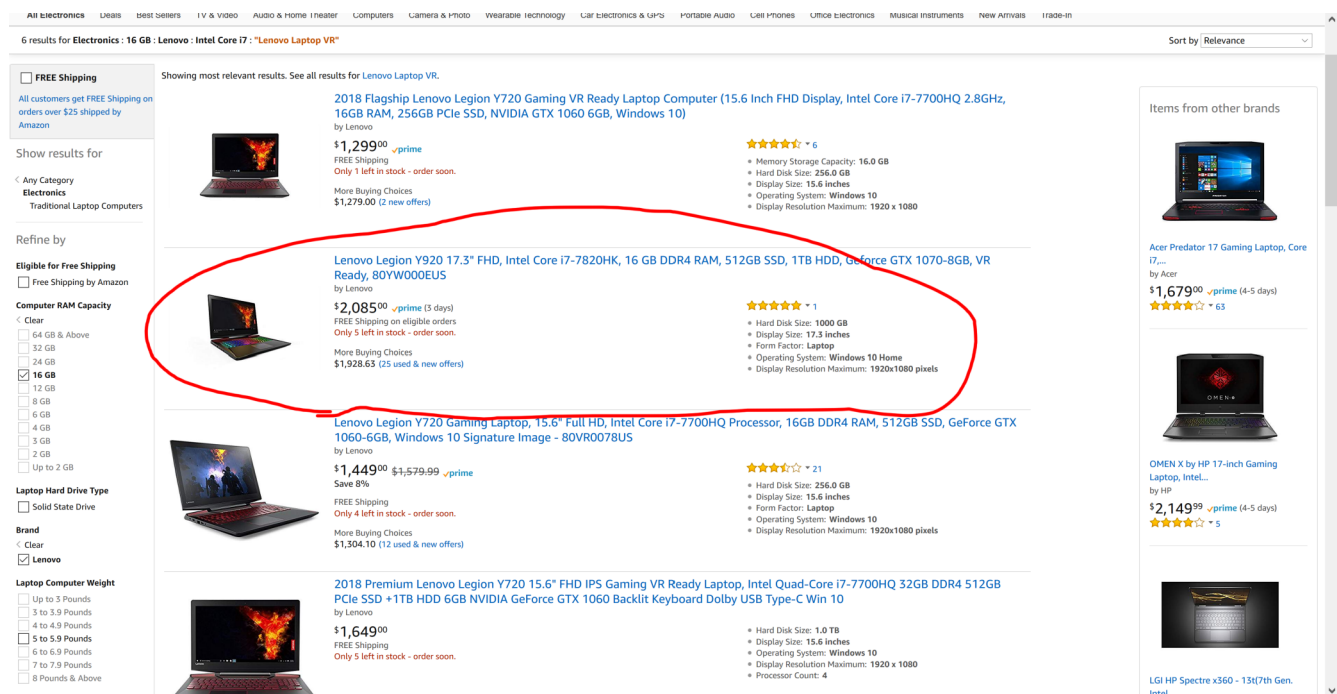
1) Search up something simple like “Lenovo Laptop VR” and hit search. This was the first major task.



2) Begin applying filters that fit the item description given to them.



3) Then scanning for the item to find it and purchase it.



From there, they would proceed to checkout.

This task is not terribly hard if you know how to do it, and most people could eventually figure it out. But what I wanted to know is how much harder is something like this for a new user to understand the learning curve on Amazon.

Challenges

For many, collecting the data would be the hardest part of the study. However, for me, that was not very challenging. I'm a social person, and talking to lone strangers in public places is not hard. And the task of finding people in public places is not hard, however, it was time consuming. I spent several hours each evening collecting data, starting Thursday, April 19th, through Sunday, April 22nd. The hardest part was keeping track of my observations. Trying to start the timer exactly at the same time was hard, as I knew it would be the easiest area to screw up the data. Another problem was when they started asking a lot of questions. Trying to decide what information to give them was tricky as I did not want to lead them so much that they would perform the purchase much faster than they normally would. Again, a point that would screw up data. So finding the points where data would be screwed up and trying to

lessen the chance of them drastically changing the results was the hardest part. With so many variables that could effect the outcome, I had to be very active in this area to prevent those from occurring.

Pain Points

I would like to make it clear, that pain points is NOT OFFICIALLY part of this study. I added this in addition because to further simulate a proper UX test, I would be testing more than just time to complete. I'd want to know where they spent the most time. BUT because it would be more complex to have two timers going; one to keep track of total time, and one to track each major task, I decided to NOT do this officially. However, I still decided to mark my observation on which task appeared to be the most difficult for the user. This is not quantitative data, and is only being added to make this study a little more useful for this hypothetical Amazon design test. But again, this is not part of my study officially and is added for depth/fun. On the table, I marked the pain points sections as !Pain Points!. The exclamation point is to make it clear that this is not officially part of my study because I am aware it does not qualify as meeting the project requirements.

There are 3 major pain point areas which I assigned to each user after observing their difficulties. I determined that the three big pints are:

- 1) Applying proper filters.
- 2) Scanning the page for products.
- 3) Applying the warranty.

Summarizing

To summarize the many points made out above:

-I wanted to study Amazon.com and how long it takes to complete a difficult task between experienced and inexperienced users. This would demonstrate learning curve. This data can be used to make Amazon easier to use so more people can find what they need. My hypothesis is that experienced users will perform the task faster, and that inexperienced users will be slower. I'm interested because I am going into UX design.

-My variables is the time it takes to complete the task and the pain points. I can measure time by timing users perform the purchase, and pain points can be marked by observation.

-I am studying inexperienced vs. experienced, defined as 2 or less purchases for inexperienced, and 5 or more for experienced. They represent portions of the entire South-East Michigan population who

have credit cards and could possibly make a purchase on Amazon. One portion being Amazon regulars, the other not.

- Results apply to all people using Amazon.com and I assume that their standard deviations are unknown and may or may not be equal. (Case 3)

- I will sample using a pseudo-sampling scheme where I would ask individuals in public places to participate in my study. I would then watch them perform the purchase and time them doing so. I would mark my observations and record the times.

- The groups are independent.

- I collected the data by observing people directly. The difficulty lied in ensuring that the test was unbiased and the same as every other test, and that all variables that would effect the outcome time would not interfere, or if they did, minimally. I tried to prevent this by being conscious of things that could mess up times.

Experienced

Sample #	Time in Seconds	!Pain Point!
1	363	Filters
2	162	Warranty
3	434	Scanning
4	265	Scanning
5	350	Filters
6	224	Filters
7	296	Warranty
8	412	Filters
9	86	Warranty
10	199	Warranty
11	136	Scanning
12	467	Scanning
13	468	Filters
14	165	Warranty
15	387	Warranty
16	243	Warranty
17	239	Scanning
18	332	Warranty
19	173	Filters
20	290	Warranty
21	124	Warranty
22	172	Filters
23	158	Scanning
24	501	Filters
25	295	Filters
26	371	Warranty
27	176	Scanning
28	57	Warranty
29	390	Filters
30	198	Filters

Average time to complete purchase: 271.1 seconds, or 4 minutes and 31.1 seconds. The most difficult pain point was Warranty, clocking in at 12 users. (40% of experienced users, 12/30 = 0.4)

- \bar{x} , or the Mean of experienced users is 271.1 seconds.
- The sum was a total of 8133 seconds.
- S_x , or sample standard deviation is 122.180858.
- σ_x , or population standard deviation is 120.127252.
- n , or number of samples, is 30.
- $\min X$ is 57.
- Q_1 is 172.
- Median is 371.
- $\max X$ is 501.

Inexperienced

Sample #	Time in Seconds	!Pain Point!
1	72	Warranty
2	898	Filters
3	663	Filters
4	642	Filters
5	652	Filters
6	105	Warranty
7	156	Scanning
8	185	Filters
9	342	Scanning
10	178	Filters
11	118	Warranty
12	340	Filters
13	689	Filters
14	514	Filters
15	357	Scanning
16	321	Scanning
17	352	Filters
18	222	Warranty
19	210	Warranty
20	502	Filters
21	267	Scanning
22	345	Scanning
23	230	Scanning
24	281	Warranty
25	615	Filters
26	290	Scanning
27	498	Filters
28	365	Filters
29	409	Filters
30	188	Warranty

The average time to complete purchase: 366.866 seconds, or 6 minutes and 6.866 seconds. The most difficult pain point was filters, clocking in at 15 users. (50% of inexperienced users, 15/30 = 0.5)

- \bar{x} , or the Mean of inexperienced users is 366.86666 seconds.
- The sum was a total of 11006 seconds.
- S_x , or sample standard deviation is 203.8848108
- σ_x , or population standard deviation is 200.4579313
- n , or number of samples, is 30.
- $\min X$ is 72.
- Q_1 is 210.
- Median is 341.
- $\max X$ is 898.

Analysis

alpha = .05

H₀ : u₁ = u₂

H₁ : u₁ < u₂

Mu 1 = Inexperienced

Mu 2 = Experienced

-	u1	u2
n	30	30
\bar{x}	366.86666	271.1
S _x	203.8848108	122.180858
ConInterval	(290.73, 443)	(225.48,316.72)

Using 2-SampTTest:

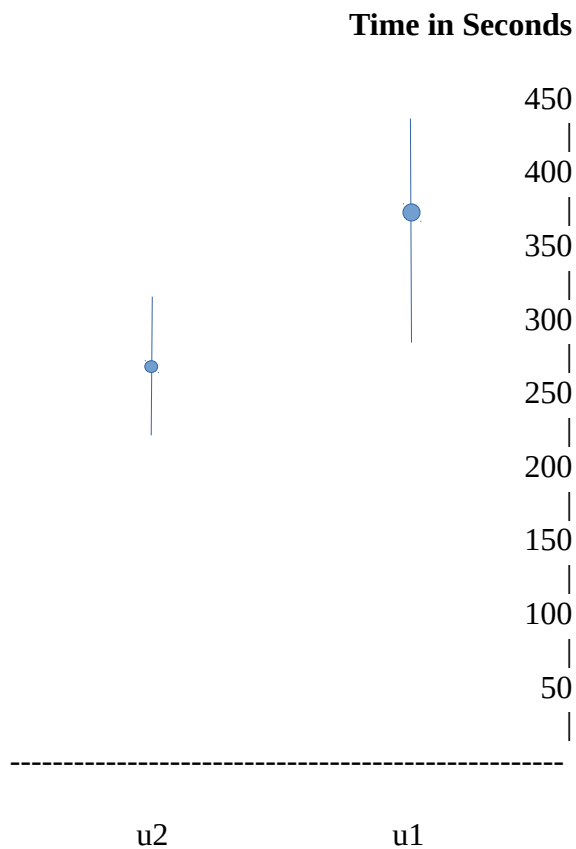
u₁ > u₂

t = 2.20679

p = .0161009

df = 47.44

Point difference: x₁-x₂ = 95.7666



Conclusions

Given that $P = .0161009$, and $\alpha = .05$, we can reject the null, meaning we know that there is a statistical difference between the two groups with 95% confidence.

With this study, we can say that it is probable that the two groups are not equal, and that Experienced users do in fact perform better and faster than inexperienced. This means there is a learning curve on Amazon, and a rather significant one.

I would predict that future tests regarding performing tasks will see a significant difference between the two groups, and that the more experienced you are the faster you will be.

I recommend further study of exactly where users have trouble, and quantitatively studying pain points. With my study, all I can concur is that there is a rather large (95.76 second) difference between the two groups.

With this knowledge, we can conclude that further testing should be done with changes to the design so newer users can learn how to find what they want easier. People who find what they want faster and with less frustration will be happier customers, which will hopefully turn into more sales. This will benefit the users, as more people will improve in proficiency, and it will help Amazon improve sales.

Good design is key in this situation, and as I concluded, there is room for growth. The goal would be to redesign the site so there is no significant difference between the two groups. And naturally, we would want to see the time of the inexperienced drop.