

User Behaviour on Google Search Engine

Bartomeu Riutord Fe

Stamford International University
Bangkok, Thailand

Abstract. The objectives of this study are (1) to analyse the relationship between the user's knowledge about Google search engine and the position of the websites that they visit on it, and (2) to learn how the user access to Pay Per Click links according to their knowledge about them. Research methodology: The sample consisted of 120 Spanish people from Palma de Mallorca who have been evaluated through a questionnaire to define their knowledge about Google search engine and then with an analysis of the position of the websites visited on 4 different searches using it. Research findings were as follows: (1) Users with low Google search engine knowledge tend to click on links in top positions in the organic search results (Pearson $r=0,706$) (2) and on Pay Per Click links (Pearson $r=-0,358$) with more incidence than users with high knowledge.

Keywords: search engine, SEO, SEM; Google, user.

Introduction

Our society is changing very fast. The way people live is getting hardly influenced by Internet and specifically by the search engines that we find on it. On the other hand, every business and organization wants a presence on the web and they use the search engines to bring traffic to them. Rehman and Naeem (2013) define the search engine as the process based on different strategies with the main goal to organize the information from the Internet and give it to the users. We consider them useful and efficient if they give us the correct information and within a short period of time. Google's mission statement from the outset was "to organize the world's information and make it universally accessible and useful" (Tung & Wu, 2013) and they start working with the Google web search, which is the basis of the company and the most used search engine on the World Wide Web (WWW), handling more than three billion searches each day (Schuster 2010). The information about search engines user's activities is very difficult to find, because it is stored at the search engine servers and is not public. There are different ways to analyse those activities. One of them is using an eye tracker like "Tobii T120" how Marcos & González-Care (2010) had used in their study borrowed by Alt 64 Company. Their research was about the behaviour of the user when using the search engine and recording the data in a processor, but it is a very expensive and complex way to work.

On the other hand Yamin (2013) had studied more than 1000 users through an interface framework to study the Google search engine. The results he got were from a huge data but very general. He was only able to draw conclusions about the dimensions of the user search behaviour. After analysing those precedents, we decided to investigate Google search engine user's behaviour in a manual way, which would involve an interview and analysing their searches, face to face.

The focal point of the research is to analyse the relationship between the user's knowledge about Google search engine and the position of the websites that they visit on it. We can divide the position of the websites in 2 different groups. The websites which link are located in the Pay Per Click (PPC) areas and the websites that have their link in the organic search results area. There are 3 different areas for Pay Per Click links and between 1 and 3 positions on each area. On the other hand, we can find 10 different positions in each page for the websites that have their link on the organic searches area, and the number of pages we can find only depend on the search (figure 1).

We suspect that according to their knowledge, the users will differ on the links they will click. The user with better knowledge should use few links on PPC position and links in positions closer to the bottom of the page on the organic search areas. If we can prove it, and understand in what proportion they are acting, it would be very helpful for webmasters and they could act according to the target group of people of their websites. The website administrator has the important duty to attract users from search engines, and because of this they use two different techniques: Search Engine Marketing (SEM) and Search Engine Optimization (SEO) (Jain, Iyengar & Arora, 2013). Using those techniques they will receive more visitors from search engines, because they are improving the position on it and also the visibility from the users, as we can see described on Figure 1.

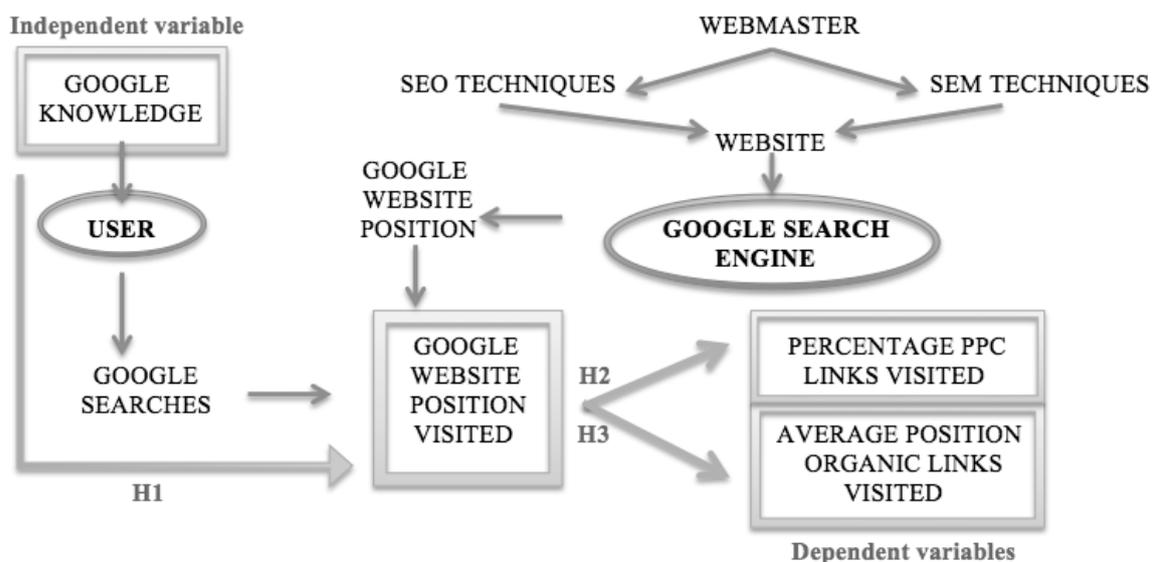


Figure 1: Conceptual Framework

Each technique has a different goal. When a search engine returns its results, it gives us two types of links: organic and Pay Per Click (PPC). Organic search results are also

called natural search results, and improving their position on it is what SEO is all about. SEM is to improve the relation between how much you pay for click to Google and how much money you get from each click from the users (Orense & Rojas, 2010). So we can summarize that SEO is about free positioning on the search engine, SEM about paying positioning.

Speaking about the Google search engine's user; the connection between them and Google can be studied according to how they act on it. Marcos and González-Caro (2010) or Yamin and Ramayah (2011) had studied the relationship between their knowledge about the search engine and search satisfaction of them, concluding that users with higher knowledge got better results on the search engine.

- Research Objectives

- 1) To analyse the relationship between the users knowledge about Google search engine and the position of the websites that they visit on it.
- 2) To learn how the user access to Pay Per Click links according to their knowledge about them.

- Research Hypotheses

H1: Users with different Google search engine knowledge are different when using it.

H2: Users with low Google search engine knowledge visit more Pay Per Click links on Google search engine in percentage than users with high knowledge.

H3: Users with low Google search engine knowledge visit more websites in top positions on the Organic results of Google search engine than users with low knowledge.

Materials and methods

- Population and Sample Size

The population of this research project will be Spanish people who are living in Mallorca. We can consider them representative of the Google search engine users who are using Spanish in their searches around the world and in a lower relation of the all the Google search engine users in the planet. We will not contemplate the gender of the study because according to Loringo et al. (2008) who did three different studies using eye tracking to study the online search, gender was considered but without influence in the behaviours. The behaviour on Google search engine will be studied from people chosen by using a simple random choice and following the study of Agresti (2010), where he says that a sample of 20-30 people for each group is enough to reach saturation in our results, we will study 120 people, because we will divide them in 4 different groups, according to their knowledge about the search engine. Therefore, we will analyse 30 people from each group.

- Data Collecting Procedure

In order to collect the data to do the study, we will use structured survey questionnaires for face-to-face interviews, and in that way understand what relationship has the interviewee with Google search engine.

After that, we will proceed to take the real data from them. For that reason we will ask them to do 4 different searches using Google search engine with the aim to prove that the search behaviour is influenced by the user knowledge like Shih et al. (2013) have shown in their study about user satisfaction.

The conditions of the research will be a limited time of two minutes for each search, and they may not use the name of any website, program or software. There are four main kinds of searches: informational, navigational, transactional or multimedia tasks. Doing informational searches, users try to gain information, with the navigational, they want to go to some specific website, with the transactional, the user wants to do some action and finally using multimedia searches they will be focus in try to watch some specific image or video (Marcos & González-Caro, 2010). We will propose 2 informational searches because 50% of the searches are like this and 1 navigational and 1 transactional search because 20% of the searches are in this purpose.

Only a minority of the searches are multimedia, or with other aims like use Google search engine like spellcheck or a search bar, and for that reason we will leave out of account this kind of searches.

The Google's searches that we are going to propose will be:

1- How high (m) is the 18th tallest skyscraper in the world? (Informational search)

We formulate this first question, because it is an informational search which no one knows and requires a lot of research through Google search engine because it needs actual data to answer it correctly.

2- How many Spanish male professional tennis players are included in the top 100 players in the world? (Informational search)

Like the first question, in Spain, even tennis fans are unaware of how many Spanish players are in the top 100. This is because this information is something that changes each month, as players receive points from each tournament they take part in; therefore their ranking moves up and down. This informational search will also need a deep research, the same as the first question.

3- Watch a specific football match (live) (Navigational search)

This is one of the most common action or search in Google search engine in Spain. It is a navigational search because the user will be trying to find a website where he/she can watch the match online.

4- Book without prepayment a hotel in New York for tomorrow on the 8th Avenue for 3 people and for less than 100 Dollar (Transactional search)

For this transactional search, the user needs to book a hotel in New York for the next day. The search is very complicated, because it has different conditions. First of all the hotel need to be in one of the most important streets in Manhattan, for three people and for less than 100 Dollars, which is almost impossible, but in this way, we can see how he/she use the Google search engine with this purpose. This question is also very useful for us, because it's easy to get PPC links on the results of the searches about booking hotels, and it will help us to study the percentage of incidence on this kind of links by the users. For each search, we are going to write all the links that the users will visit, differentiating them if they are PPC or Organic links and their position as we can see on Figure 2.

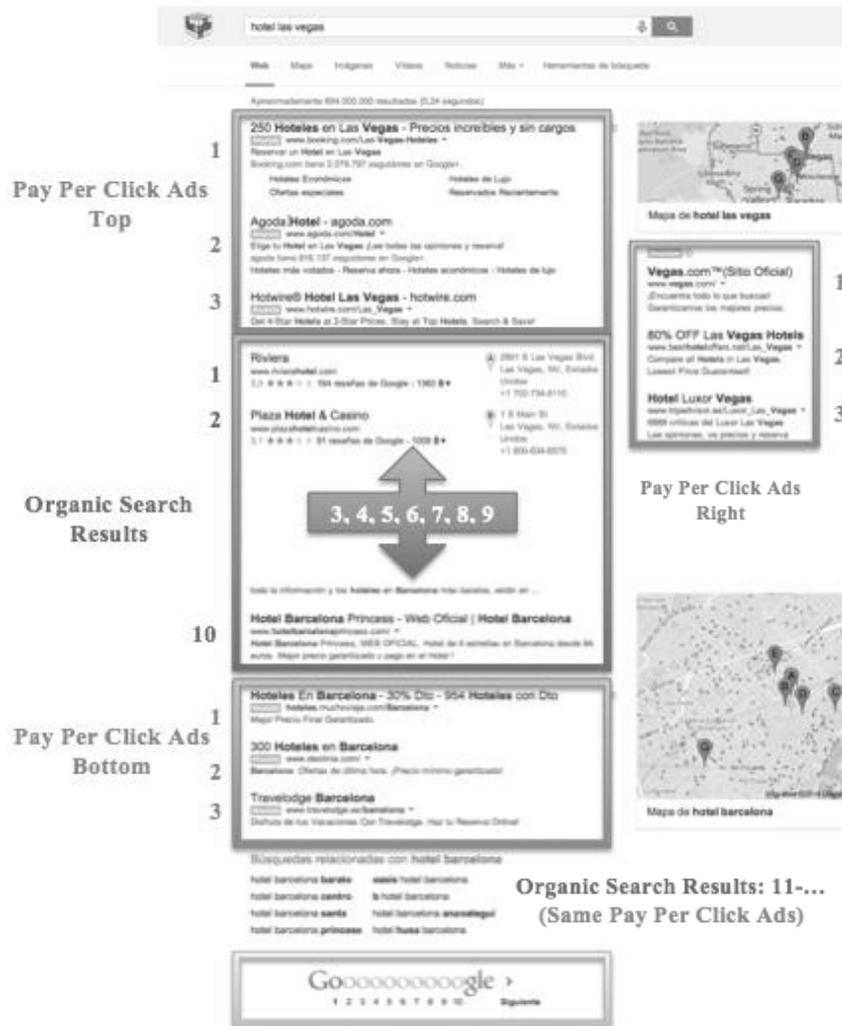


Figure 2. Links position

- Quantitative Data Analysis

We will do quantitative data analysis because the data collected will be numeric in form and we will use statistical techniques to draw conclusions about participant behaviour.

The analysis will be made using correlation analysis with a Data Analysis Program in order to test the relationship between the percentage of PPC and Organic links visited, each group of Google search engine knowledge, and the average position of the websites visited on Organic searches from each group. The analysis of the data collected will be studied through Pearson Correlation and Linear Regression to determine how hard is the relation between variables.

Results

- Test Hypothesis 2

H2: Users with low Google search engine knowledge visit more Pay Per Click links on Google search engine in percentage than users with high knowledge.

For the second Hypothesis, we have two variables. On one hand we have the test mark, obtained from the questionnaire, and the values of the marks are between 0 and 40, which is the higher punctuation. On the other hand, we have the percentage of Pay Per Click links visited on Google search engine. Theoretically, we have divided the two variables of this Hypothesis in dependent and independent. In this case, my independent variable is the test punctuation, like we can see in the conceptual framework of the study, and my dependent variable is the percentage of PPC links visited.

To test this second hypothesis, the first step that we did was to study the Pearson Correlation between variables. For that reason firstly we studied its graphic relation, where we could see that there is a negative correlation, therefore, my prediction is fulfilled, because how higher is the knowledge about Google, lower is percentage of Pay Per Click Links visited. We can also observe how the points are diverse, but it is because a big percentage of users didn't click any time on the PPC links, and it means that his percentage of PPC links visited is 0%.

After the graphic relation, we had studied the numeric relation through Pearson bivariate correlation. The results of the correlation are the following:

Table 1. Hypothesis 2 Correlation Coefficient

		Test Punctuation (0-40)	PPC (%)
Test Punctuation (0-40)	Pearson	1	-,358**
	Correlation		
	Sig. (2-tailed)		,000
	Sum of squares and Cross-products	15281,967	-4145,800
	Covariance	128,420	-34,839
PPC (%)	N	120	120
	Pearson	-,358**	1
	Correlation		
	Sig. (2-tailed)	,000	
	Sum of squares and Cross-products	-4145,800	8776,800
	Covariance	-34,839	73,755
	N	120	120

** . Correlation is significant at the 0,01 level (2-tailed).

We can see in the Table 1 how the correlation is negative and $r=-0,358$. This value means that the relation between variables is weak-medium, because we can consider weak relation with the total value of 0,25 and medium of 0,5. Anyway, the relation exists and it is significant between Pay Per Click percentage of visits and the test punctuation about Google Knowledge.

Based on Pearson Correlation at 95% confidence interval, the result of the test indicates that significant value of 0,00, which is less than 0,05. It means then, that the null hypothesis should be rejected. The figure interprets that there is significant relationship between the result in the test and the percentage of PPC links visited.

To find the influence of the percentage of PPC links visited and the knowledge about Google, we used a linear regression between variables. In this case we also use the test punctuation like independent variable and percentage of PPC websites visited as previous with the dependent variable.

On the other hand, as we can see on the Table 2 obtained from the linear regression, the value of R is 0,358 and means that 36% of PPC websites visited that are influenced by the punctuation of the test we did in the interview, equals 36%. It is not a very hard relation, but enough to say a linkage exists.

Table 2. Hypothesis 2 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,358 ^a	,128	,121	8,05284

a. Predictors: (Constant), Test Punctuation (0-40)

- Test Hypothesis 3

H3: Users with low Google search engine knowledge visit more websites in top positions on the Organic results of Google search engine than users with low knowledge.

For this third Hypothesis, we also have two different variables; on one hand the independent one, the same in the test for the Hypothesis 2: the test punctuation about Google search engine knowledge, and the dependent one is the average of the organic links visited during the four searches analysed.

Firstly, to study this Hypothesis 3, we looked at the Pearson Correlation between both variables. The dependent one is going to the Y Axis, the independent to the X Axis, and we had a positive correlation like we had predicted. It means that experimented users use lower links on the Google search engine, and for that reason the average of the position of the websites visited is lower than non experimented users with less knowledge about the search engine.

Secondly, we got the following results with Pearson Correlation:

Table 3. Hypothesis 3 Correlation Coefficient

		Test Punctuation (0-40)	PPC (%)
Average Searches	Pearson Correlation	1	-,706**
	Sig. (2-tailed)		,000
	Sum of squares and Cross- products	230,210	1324,891
	Covariance	1,935	11,134
	N	120	120
Test Punctuation (0-40)	Pearson Correlation	,706**	1
	Sig. (2-tailed)	,000	
	Sum of squares and		

Cross- products	1324,891	15281,967
Covariance	11,134	128,420
N	120	120

** . Correlation is significant at the 0,01 level (2-tailed).

Based on Pearson Correlation at 95% confidence interval, the result of the test indicates a significant value of 0,000, which is less than 0,05 and it means that the null hypothesis should be rejected. The figure interprets that there is significant relationship between the result in the test and the average of Organic links visited.

On the other hand, based on Pearson Correlation Coefficient, the relation between variable is strong and positive, because is: 0,706, like we can see on the table 3. We consider the relation medium-high, because from 0 to 1, we can say that 0,5 is medium and 0,75 high.

Secondly, using linear regression with the same variables, we will be able to know how influenced are both variables, the table 4 show us how the influence of a 49,9%, on the average of the Organic links visited and the punctuation on the test about Google knowledge, which is a high result.

Table 4. Hypothesis 3 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,706 ^a	,499	,495	,98869

a. Predictors: (Constant), Test Punctuation (0-40)

- Test Hypothesis 1

H1: Users with different Google search engine knowledge are different when using it.

This Hypothesis is like a summary of the second and third Hypothesis. The results we took from the Hypothesis 2 is that the correlation is significant (0,00 lower than 0,05) and the relation between variables is weak-medium ($r=-0,358$). On the other hand, from the Hypothesis 3, we got the results that the correlation is also significant (0,00 lower than 0,05) and the relation between variables is medium-high ($r=0,706$).

So, we can conclude that the users with different Google search engine knowledge act differently when they are using it.

Discussion

H2 and H3 are not enough to disprove H1, because there are other aspects where users can act differently using the search engine. If the result would be negative for H2 and H3 We couldn't say that H1 is not true, but in my case both results are positive, and we can also prove H1.

Saying that users act in a different way is not very difficult and complex. It is obvious that with different experience and knowledge about some application, program, website... they will act different, because they know its performance better. In my case, it is easy to know that there are differences using the search engine, but it is difficult to prove it and the most important, to take advantage from these differences and to improve the SEO and SEM techniques.

Now, referring to the H2 and speaking about Pay Per click links, say that the result of the correlation is significant and the relation between variables is just weak or medium. But if we divide the results of the four different groups, we can see clear differences between them, even if the relation between variables is weak. 25/30 users from group 1 were using PPC links, 22/30 from group 2 were using them, and only 15/30 from group 3 and 7/30 for group 4. In my opinion, from the four different searches that we have proposed to the users, just the last one, the transactional search: (“Book without prepayment an hotel in New York for tomorrow at the 8th Avenue for 3 people and less than 100 Dollar”), is the unique proposition to tend to do searches where there are a lot of PPC links, and it is not enough to analyse the incidence to click on them. Just 59 from 120 people interviewed visited websites with PPC, and for that reason are very difficult to get meaningful conclusions. One solution could be to increase the number of searches with higher percentage of PPC or do a specific study about this kind of links.

The differences between users have sense. The users who don't know how PPC links are working, and why they are at the top of the websites and in an square with different colour, tend to click on them, but those who know that the webmasters of those websites have to pay for each click on its links, already know that it has a high cost and they need to get an average of money for each click higher than the money that they pay to Google for each click, so usually the services that they offer in those websites are exchange for money transferences, so many expert and usual users try to skip them, depending on the kind of searches that they are doing. For example if they only need information or some applications that should be free, they will not use PPC links, but if they want to buy something they could use them, even this is not very common.

Finally, speaking about the H3 about Organic searches, the result of the correlation is also significant and in this case, the relation between variables is high with $r=0,706$. There was a big difference in the average of the position of the links visited by the users. Using the correlation equation, and using the values of the lowest punctuation in the test in the higher, we can predict how a user with a mark of 1/40 should have an average of 1,57, when a user with 40/40 an average of 4,96. This difference is very high, because it is almost 4 positions on the Google page.

The explication about the difference of the average of the websites visited is that a user with experience and high knowledge knows how to select better a website according to the information visible in the preview of the website in Google. On the other hand he/she also knows that in the top positions are websites with good SEO techniques, and it doesn't mean that they have better quality of contents. Everyday Google is getting better, and through their algorithm, they try to ensure the best websites are at the top, with the most reliable information and quality, but webmasters have some tricks to get these first positions, and some users know it, and they can even identify and skip them.

Conclusions

A better understanding of the behaviour of the users of Google will help the webmasters to improve their techniques, and for this reason this research used a quantitative study to explore it and in order to solve the objectives of the study.

We can conclude that a relation exists between the knowledge of Google search engine and the links of the websites visited by the users according to their position on the result page of the search engine.

On one hand, exist a weak-medium negative relation (Pearson's $r = -0,358$) between the knowledge about the search engine and the percentage of Pay per Click websites visited by the users. It means that how higher the knowledge about the search engine, lower is the incidence from them on PPC links.

On the other hand, the relation between their knowledge and average of the position of the links visited from the organic results is medium-high (Pearson's $r = 0,706$) and positive. It means that users with good knowledge about the search engine tend to click on links with higher positions on the Organic results, and it means that those links are closer to the bottom of the page.

For that reason, we consider that the experimented users, and therefore with better knowledge, can identify the websites with better information and benefits according to their searches, and usually exclude PPC links and websites in top positions of the organic results.

References

- Agresti, A. (2010). *Analysis of ordinal categorical data* (Vol. 656). John Wiley & Sons.
- Jain, R., Iyengar, S., & Arora, A. (2013). Overview of Popular Graph Databases. In Proceedings. *Fourth IEEE International Conference on Computing*, 18
- Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., & Gay, G. (2008). Eye Tracking and Online Search: Lessons Learned and Challenges Ahead. *Journal of the American society for information science and technology*, 59(7), 1041-1052. Doi: 10.1002/asi
- Marcos, M., & González-Caro, M. (2010). Comportamiento de los usuarios en la página de resultados de los buscadores. Un estudio basado en eye tracking. *El profesional de la información*, 19(4), 348-358. DOI: 10.3145/epi.2010.jul.03
- Rehman, K., & Naeem, M. (2013). The foremost guidelines for achieving higher ranking in search results through Search Engine Optimization. *International Journal of Advanced Science and Technology*, 52, 101-109
- Schuster, H. (2010). The use of Google services and tools in academic institutions - A critical review. *International Review of Law, Computers & Technology*, 24(1), 17-24. Doi: 10.1080/13600860903570129
- Shih, B. Y., Chen, C. Y., & Chen, Z. S. (2013). An empirical study of an internet marketing strategy for search engine optimization. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 23(6), 528-540.
- Tung, C., & Wu, H. (2013). An Internet-Connected World: Google's Platform Strategies to Network Industry. In *Workshop Proceedings of the 9th International Conference on Intelligent Environments*, 17, 201
- Yamin, F. (2013). Interfacing Google search engine to capture user web search behavior. *International Journal of Electronic Commerce Studies*, 4, 47-62. Doi: 10.7903/ijecs.1111
- Yamin, F., & Ramayah, T. (2011). The impact of user knowledge on web search satisfaction. *American Journal of Economics and Business Administration*, 3(1), 139