Eindhoven University of Technology

MASTER

Targeting website content real-time

the effects of real-time targeting website content on user experience and behavior

Swelsen, K.J.M.

Award date:
2013

Link to publication

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, 21-02-2013

Targeting website content real-time

The effects of real-time targeting website content on user experience and behavior

by Kevin Swelsen

identity number 0665422

in partial fulfilment of the requirements for the degree of

Master of Science
in Human Technology Interaction

Supervisors:
Dr. Ir. M.C. (Martijn) Willemsen (TU/e)
Prof. Dr. C.C.P. (Chris) Snijders (TU/e)
M.P. (Mark) Graus, MSc (Adversitement B.V.)

Abstract

Content providing websites are aimed at keeping a visitor engaged. Facilitating those visitors with content they are interested in is a way of keeping them on a website. In recent years websites have developed techniques and algorithms to change the website to facilitate those users in real-time. Different users could have different interests and by tracking the behavior of those users on a website, that website can make sure the user is presented with only the content that user is interested in. What was missing in this field of research is a study on the effect of real-time changing a website on user experience. Therefore, the main goal of this study was to study the effect of directly and real-time targeting of content to specific user segments on both behavior and user experience.

This project was done in collaboration with Hardware.Info, a large online IT-platform. Using their log data it was found that two distinct types of sessions exist on the website, with users in both segments each showing different behavior in those sessions. One segment were tech-savvy users interested in hardware like motherboards and processors, the other segment were users who were interested in gadgets like mobile phones and tablets. An algorithm based on limited information from the first five visited pages of a user in a session could then predict the segment of a user in a session in most cases correctly.

An experiment was designed to test the effects of facilitating content to both segments of users. The sidebar on the Hardware.Info website was adjusted to either showing tech-savvy content, gadget content or neutral (both tech-savvy and gadget) content. Several hypotheses were defined to test the effect of the manipulations on both the behavior of the user as well as the user experience.

The results of the experiment once again showed that users on the Hardware.Info website can indeed be divided up into two segments. There was a direct positive effect between tech-savvy users who were presented with tech-savvy content and the usage of the sidebar. Results also showed that the effectiveness of the sidebar increases when users find the sidebar more appealing and accurate. A higher perceived sidebar effectiveness led to an increase in the usage of the sidebar. Tech-savvy users presented with tech-savvy content found the sidebar on average the most accurate. No differences were found among gadget users.

The direct change in behavior as a result of targeting content is a replication of earlier studies and this study provided evidence that the change in behavior, as a result of directly targeting content towards a user, can at least partly be explained by a change in user experience.

Preface

This thesis presents my graduation project in the group Human Technology Interaction at the Eindhoven University of Technology. This project could not have been completed without the assistance and help of several people. I would like to use this chapter to thank those people.

First of all I would like to thank Martijn Willemsen. The weekly meetings in his office led to sometimes heated but always useful discussions. These meetings always led to more work than previously thought but at the same time really helped my project forward.

Second, I also would like to thank Mark Graus. Without his support and help this project would not have been possible. Whenever there was a problem that I could not handle myself, he was there to help and give support. His help in the data analyses turned out to be of vital importance.

Third, I would like to thank Bart Knijnenburg for helping me with the confirmatory factor analyses and the structural equation modeling.

Last but certainly not least, I would like to thank Hardware.Info, and especially Koen Crijns, Eric van Ballegoie, Karel Souffriau and Frank Everaardt. Not only did they allow me to use their data and run my experiment on their website, but they were always open to my sometimes crazy ideas, and they were always helpful in coming up with solutions.

This was a very challenging but interesting project! I am very happy I got to work with you all! Thank you!

Kevin

Index

1. Introduction

People use the internet more and more these days (CBS, 2012). Nowadays there is a website about virtually every topic. Where some websites have the goal to sell products, other websites are more about informing people and providing content. A lot of content providers have grown over recent years covering a large variety of topics. This also means these websites have started to gain interest from a larger and more varied public who are not all looking for the same content. Visitors to such a website might differ in the kind of information they are looking for or they might have different reasons why to visit a website. For these websites it is important to keep users satisfied. Satisfied users will visit more pages which could be beneficial for the website in terms of advertising. To keep users satisfied, the website has to provide the right content for those users. Facilitating users of a website by directing them to the content of their liking can be done in several different ways and a lot of research has been done in this area.

1.1 Recommender Systems

One way of providing users with the content of their liking is by giving users recommendations on certain items. For example, when looking at a product in an online store that store can give users a set of other products that the user is likely to be interested in as well. When a user is looking at a certain book in an online bookstore the store could give the user other books the user might be interested in as well based on the book the user is looking at (e.g. Herlocker, Konstan & Riedl, 2000; Ma, Yang, Lyu & King, 2008; Knijnenburg, Willemsen & Kobsa, 2011a).

Knijnenburg et al. (2011a) reviewed literature that tried to fully capture how a certain user experiences a particular website or (recommender) system. Hassenzahl & Tractinski (2006) defined user experience as a consequence of a user's internal state (e.g. mood, motivation, and needs), the characteristics of the designed system (e.g. complexity or usability) and the context of the interaction (e.g. the setting). Mahlke (2005) reached almost the same conclusion but added the note that it has not been answered which design characteristics support a positive experience and therefore these models are only usable for evaluating user experiences. Van Schaik & Ling (2008) also experimentally tried to define a model using user experience. In their experiments they tried to focus on the effect of different design principles.

Knijnenburg et al.(2011a) then took a slightly different perspective when trying to measure user experience themselves. In their model they argue that objective system aspects influence subjective system aspects which in turn may influence user experience and therefore the interaction with the system. Figure 1 shows an overview of their model. The advantage of this kind of model over previous research is that by using this model different kinds of constructs can be distinguished that together measure user experience. The model by Knijnenburg et al. (2011a) takes the aesthetics of the system into account like van Schaik & Ling (2008) did as well, but also the concepts as defined by Hassenzahl & Tractinski (2006).

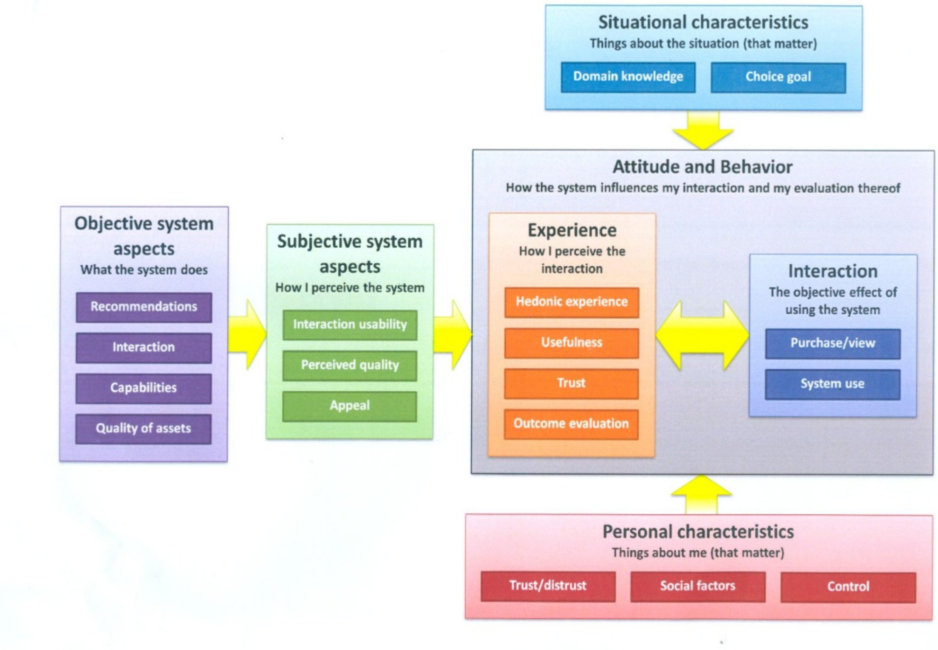

Figure 1: User experience framework (Knijnenburg et al. 2011a)

1.2 Early adaptive hypermedia systems

Another way of facilitating users can be to change (morph) part of the looks and/or the content of the website to match the interests or behavior of a specific user. A website makes a model of the user and then adapts itself to match that user model. These kinds of web interfaces are also called adaptive hypermedia.

Already in the nineties researchers saw the possibilities in adapting a system to specific groups of users. Brusilovsky (1996) extensively reviewed some of the existing methods and techniques in early adaptive hypermedia systems. One interesting distinction made in early research on adaptive hypermedia was between adapting the content of a website and adapting the navigation support on the website.

The idea behind adaptive navigation support, according to Brusilovsky, is to help users to find their paths in hyperspace by adapting the way of presenting links to goals and knowledge. One way a website can do this according to Brusilovsky (1996) is by restricting the navigation space by hiding links that are not relevant to the user. Brusilovsky and Pesin (1998) developed an example of such a system. A screenshot of this system is shown in figure two. Even though the system is in Russian, the screenshot still demonstrates that some links are removed, presumably because they are not of interest to the user of that system at that moment.

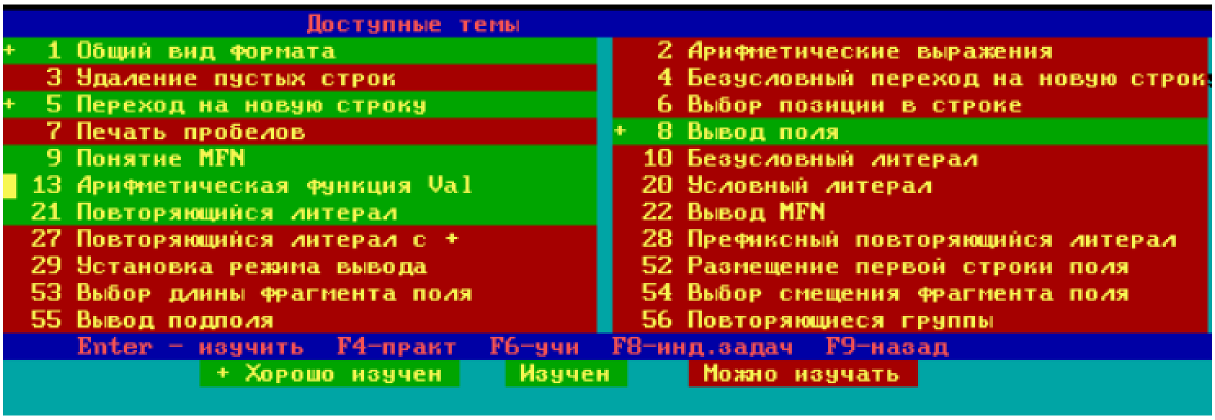

Figure 2: Hidings links (Brusilovski & Pesin, 1998)

1.3 Learning Environments & Adaptive Hypermedia

A lot of research on the topic of adaptive hypermedia has been centered on learning environments. As Fan (2004) states clearly, with adaptive hypermedia it is possible to deliver content to a student based on that student' learning requirements, learning style and experience. De Bra, Aerts, Berden, De Lange, Rousseau, Santic, Smits & Stash (2003) for example showed a system called 'AHA!' , an adaptive hypermedia platform build to give users in online courses extra guidance by conditional extra explanations and hiding of certain links.

The AHA! system gets a notice every time a user visits a page, and uses that information to construct a user model. The system is based on concepts and attributes. Every page is associated with a concept, and visiting a page increases, for example, the knowledge attribute for the concept corresponding to that page, for that particular user. The system also shows the suitability of certain links. The AHA! system determines whether a page is suitable for a certain user given its model of that user. Based on that model the system marks links with different colors. A black link means for example that the page is not suitable for the reader.

The AHA! system is a good example of adaptive hypermedia centered on learning environments. The user model is very simplistic, and is updated real-time. However, no study to date investigated whether the user experience of students using such a system increased.

1.4 User model construction

The main problem researchers faced when building an adaptive hypermedia system is the construction of a user model. A user model can be constructed based on information acquired from surveys and tests prior to actually using the system. This is called explicit modeling. However, like the AHA! system described in the previous chapter, systems can also update a user model based on the interactions a user makes with the system. This is called implicit user modeling.

1.4.1 Explicit user modeling

Research has shown that building a model based on the cognitive style of a user can be successful (Bajraktarevic, Hall & Fullick, 2003). Bajraktarevic and his colleagues (2003) for example determined the cognitive style of a user based on an extensive questionnaire the user had to fill out prior to using the adaptive system. A cognitive style refers to a person's habitual, prevalent, or preferred mode of perceiving, memorizing, learning, judging, decision-making and problem solving (Tappin-Bernard & Habieb-Mammar, 2005). Information about the cognitive style of a user can then be used to adapt a system or website specifically towards users or segments of users.

There is one large obvious drawback to explicit user modeling. In modern-day large web based systems with a lot of different users it is very impractical to give all these users a survey prior to using that system. For a user it is very time consuming to fill out a survey prior to entering a website, and therefore this method is undesirable.

1.4.2 Implicit user modeling

In order to overcome the problem of having to ask every single user to fill out a survey before using a system or website, researchers in recent years have tried to focus on automatically creating a user model. This is called implicit user modeling. Friaz-Martinez, Chen & Liu (2007) for example tried to automatically identify a user's cognitive style in a digital library setting. They were able to automatically categorize a reasonable amount of users, but acknowledged it was not that easy. Their main problem was that some of the algorithms they implemented to detect the cognitive style of the user were so complex that they were useless when trying to detect that cognitive style real-time. The algorithms simply took too long to generate a result because they were computationally to heavy.

A lot of research has gone into improving algorithms to analyze navigation sessions or forecast user navigation (e.g. Borges & Levene, 2007; Lu, Dunham & Meng, 2005; Hassan, Junejo & Karim, 2009; Dembczynski, Kotlowski & Sydow, 2007). These kinds of algorithms have to scale well, have to be able to handle large amounts of data and they have to be able to work with only a very limited amount of data.

When these kinds of algorithms are implemented, they have been proven to work. In research done by Hauser, Urban and Liberali (2008), purchase intentions went up with 20% on a telecom website compared to that with the same website without morphing. This research focused on predicting the cognitive style of a user based on their interaction with the website and did so real-time.

What is missing in all of the studies done on implicit user modeling is analyzing the effect of the categorization on the experience of the user. All of the algorithms on implicit user modeling gather input that is coming from real-time interactions the user has with the website or system. After a number of interactions, the system has enough information and is able to categorize the user and change its looks or its content to facilitate that category or segment. If a system or website is changing real-time because of a change in the user model, this means that users actually see the interface change. This could even mean users see content that one click later is all of a sudden gone.

1.5 Online advertising

Showing content specifically targeted towards specific users has also been a major topic in online advertising. So-called behavioral targeting has gotten much attention in recent research in the field of online advertising. In systems using behavioral targeting, information about a person's behavior on the internet is being collected and then used to segmentize that person based on that behavior. Online advertisements specifically designed for a certain user segment can then be shown only to the segment it was designed for. Microsoft, Yahoo, Google, Facebook and many others use behavioral targeting to show the right advertisement to the right user.

In research by Yan, Liu, Wang, Zhang, Jiang and Chen (2009) the effects of behavioral targeting were studied. The study used the data of one month from a commercial search engine. One of their conclusions was that as a result of behavioral targeting the Click-Through-Rate (CTR) of an advertisement can improve as much as 670% by segmenting users and showing advertisements specifically designed towards those segments. Another interesting finding in this study was that segmenting users based on the behavior of that user of only the past day worked significantly better compared to segmenting users based on behavior of seven days.

1.6 Summary

Research so far on web morphing or adaptive hypermedia can be split up in two distinct directions. A lot of mainly older research on this topic focused on developing a user model based on a set of tasks a user had to do or filling out an extensive questionnaire prior to entering the website or system. This has an obvious major drawback: it cannot be asked from a user to fill out an extensive test before entering every website that user visits. Therefore, more recently, research has focused on automatically identifying different types of users. Automatically identifying different types of users implies that users see the website change. It takes a while, depending on the algorithm used, before the system has enough information to segment a user, and when the system is done it will change the website's look-and-feel or content based on its model. Even though studies has proven that automatically discovering preferences and cognitive styles of a user has increased for example purchase intentions and click-through-rates, this does not imply a user likes this kind of adaptation.

This study tries to combine some of the research as described in this chapter. As found by Yan et al. (2009) and Hauser et al. (2008) users change their behavior when content on a website is targeted (real-time) towards those users or the segments of those users. This study will try to replicate these findings while at the same time investigate the effects of this targeting of content on the user experience of the system. Therefore, the main research question of this paper is as follows:

"Will targeting the content of a website real-time to a specific user segment result in a higher user experience and a difference in behavior on that website?"

The current project is the result of a collaboration between the Eindhoven University of Technology (TU/e), Adversitement and Hardware.Info. Hardware.Info is an online platform with news and reviews of computer hardware and electronics, a forum where users can post questions they have and a section where users can look up prices of different kinds of electronics and be redirected to a shop that sells that specific component or system. Hardware.Info has roughly 550.000 unique visitors a month, which makes it a really interesting partner for this project.

2. Preliminary data analysis

As stated in the previous chapter, no research to date investigated the effect on user experience when targeting the content of a website real-time to specific user segments. To fully explore this question, this project is a collaboration with Hardware.Info, a popular online IT-platform.

Hardware.info has had a shift in content from purely components towards more end user products, and had the question whether this lead to a shift in audience and if so, how to cater to this new audience. Among their visitors they believed to have very tech-savvy users: users who are for example looking for specific information about certain components but also gadget-oriented users: users who are looking for gadgets and/or pre-built computer systems. In order to answer the main research question, the first task was to find out if this hypothesis was true. This would be an advantage because, as explained in the previous chapter, most algorithms that create user models based on real-time data are not only very complex but need to scale well, handle large amounts of data and generate a result in a reasonable amount of time. If this hypothesis were to be true, a possible algorithm that builds a user model can stay relatively simple. With a simple algorithm it will be easier to investigate the main research question.

This stage of the research has two different goals. First it is necessary to find out if the two hypothesized user segments actually exist. If that is indeed the case and a high percentage of users can be categorized, then the second step will be to predict in which category a random user falls based on very limited information: as described in chapter one, eventually the website needs to be morphed within a only a few interactions the user makes with the website because visits to a website are typically not that long.

2.1 Categorization

To test if the hypothesized user segments exist, Hardware.Info provided access to their website log data. The data acquired from Hardware.Info were the server logs of 1 month, consisting of roughly 28 million http-requests. From this, the following relevant variables were used:

- IP-address
- The URL of the http-request
- A timestamp
- HTTP Referrer

Because this data was in plain text, a number of steps were taken to clean up the data and recode it into a format which would make the analyses easier. Eventually, the data was put into an SQL-database with the structure shown in figure 3.

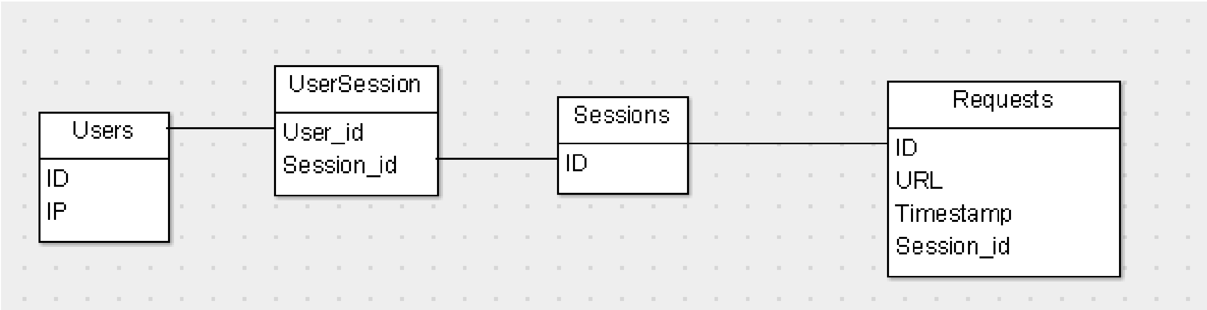

Figure 3: Structure Hardware.Info data

This structure implies a separation between a user, a session and a request. A request is the http-request that was in the original data. A session consists of all requests within a certain time period, in the case of this study 9000 seconds. The original plan was to keep this time period at 900 seconds but due to an error in one of the conversion scripts it became 9000 seconds. The difference between a 900 and 9000 second interval when constructing sessions turned out to be minimal. A session can be used to distinguish between visits. One particular user could visit the website several times over the period the data was acquired but maybe for different purposes. That is also why in the database structure a different entity is necessary for users. By structuring the data like this, roughly 1.6 million sessions and roughly 450000 users could be distinguished. Unfortunately, it was only possible to distinguish users based on their IP-address. Different users can have the same IP-address when they are behind the same router, for example when people share an apartment or people at an office.

Not all the http-requests in the database were of interest when analyzing the behavior of users. Among the requests were requests coming from search engine spiders, requests to advertisements, graphs and parts of the discussion board Hardware.Info has (graphs, smiley's of the discussion board and advertisements were all separate requests: a request to a certain page with an advertisement would result in two requests: one for the page itself and one for the advertisement). These requests were not relevant for this study and therefore removed from the dataset. Out of all sessions (originally 1.6 million), 1.4 million sessions still had requests left. 177415 Sessions did not have requests that were still in the database after removing all requests to advertisements, graphs, smiley's etc.

Most of the remaining requests could be linked to a page type. A page type here is defined as being for example a news page or a review page: a certain particular area of the website. 15.4 million requests could be linked to such a category. This was important because with a page type a page request can be linked to a product group. Hardware.Info provided us with 116 product groups (for example processors, motherboards, laptops, etc.). Every request from several page types could be linked to those product groups. For example, a review about a certain motherboard (the page type here is reviews) could be linked to the product group motherboards using an ID nested in the URL belonging to the page request. All requests from seven different page types could be linked with a product group. To test whether the initial hypothesis from Hardware.Info could hold, the 116 product groups were divided into two groups: tech-savvy and gadget. A processor for example is a tech-savvy product group, a mobile phone is a gadget product group. An overview of all product groups with its categorization can be found in Appendix I. Fifty-nine product groups (linking to 4.1 million requests) were defined as tech-savvy, the remaining fifty-seven product groups (linking to 3.3 million page-requests) were defined as gadget groups.

In summary, the following steps were taken to categorize requests:

i) Start: 28 million http-requests.
ii) Remove requests from spiders, requests to advertisements, smiley's, graphs etc. from the dataset
iii) Most pages that are left were linked to a page type (for example review, news, forum etc.).
iv) Based on the page type, a lot of requests were linked to a product group (for example processors or tablets).
v) Each product group was categorized as tech-savvy or gadget.
vi) Because each product group is now categorized, every request linked with a product group is categorized as well.

2.1.1 Sessions

Out of all sessions, roughly 630.000 sessions did not have any requests to tech-savvy or gadget pages. This is possible because not all requests were linked to a product group. For a possible categorization on a session level users would have to make a lot of requests from one category. For example, if a user would make a lot of tech-savvy requests and only very few or none gadget requests in a single session then we should be able to categorize that session as tech-savvy. For this analysis two things are important. First of all, the length of the session in terms of the amount of pages a user visited within a session (the length of the click path) is important. If a session only consists out of two or three requests, it will be almost impossible to confidently categorize that session into tech-savvy or gadget. But also more importantly, such a session is not interesting in terms of this research: only long sessions (with many requests) are interesting and useful. Users will not get to see any kind of possible manipulation when they only visit a few pages. Also important is the ratio of tech-savvy and gadget pages a session consists of. If a session consists out of an equal amount of tech-savvy and gadget pages, it is impossible to categorize that session into tech-savvy or gadget based on the information acquired.

In a lot of sessions, users visited only a limited amount of pages. On average users visited 9.44 pages per session (Mdn=3). In only 293864 sessions users visited more than 10 pages. Figure four shows the frequency of the amount of clicks per session.

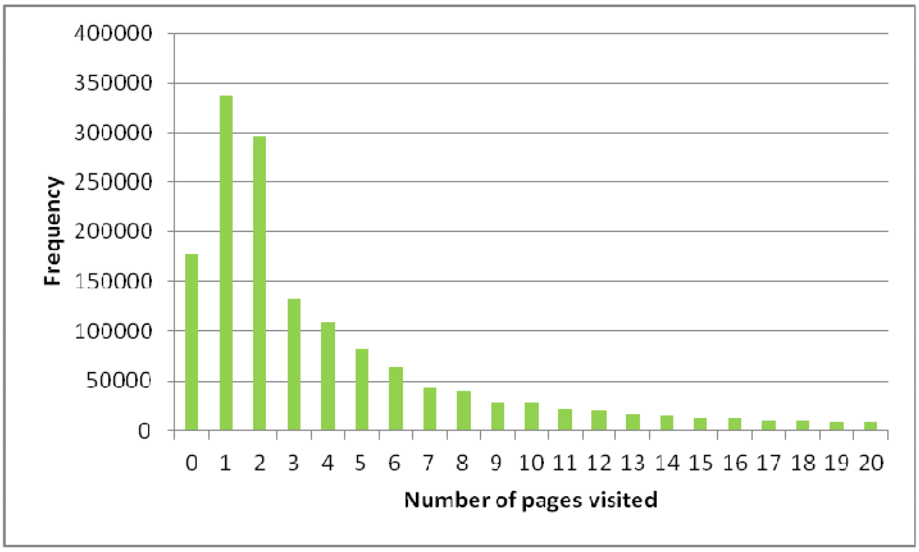

Figure 4: Number of clicks per session

It was decided that a session had to consist of at least 10 clicks for it to be useful in terms of categorizing. It was hypothesized that after 10 clicks a user spend enough time on the website and possible patterns in the behavior of that user should be visible. Like stated earlier, with the information that is presented so far it would not be possible to categorize a session that has as many tech-savvy clicks as gadget clicks. A ratio has to be defined that says how many more requests from one category compared to the other category has to be made in order for a categorization to be sufficient. Table one shows the amount of possible gadget and tech-savvy sessions with different ratios. A ratio of five means that in order to be categorized as tech-savvy, a session had to consist of at least five times as many tech-savvy requests compared to gadget requests.

	Ratio = 5	Ratio = 4	Ratio = 3	Ratio = 2	Ratio = 1.5
Tech-savvy sessions	37.90% (N=111393)	39.31% (N=115524)	41.28% (N=121317)	44.60% (N=131057)	48,91% (N=143727)
Gadget sessions	23.36% (N=68634)	24.26% (N=71285)	25.68% (N=75468)	27.82% (N=81761)	30,61% (N=89953)
Not categorized	38.74% (N=113837)	36.43% (N=107055)	33.04% (N=97079)	27.58% (N=81046)	20,48% (N=60184)

Table 1: Categorization sessions

Eventually the ratio was set at two. On the one hand the difference in amount of tech-savvy clicks and gadget clicks had to be as large as possible but on the other hand as many sessions as possible had to be categorized. Setting the ratio at two was seen as a good trade-off between those discrepancies.

In conclusion, out of 293864 sessions that were of potential interest for this research given the length of the session, 212818 could be categorized as tech-savvy or gadget. The validity of this categorization is investigated by comparing behavior of users of both segments.

First, of all sessions, very few users visited both gadget and tech-savvy pages. This is demonstrated by the figure below. Zero switches means that a user did not visit any tech-savvy or gadget pages in a session, one switch means the user only saw pages belonging to one category: either tech-savvy or gadget pages next to possible uncategorized pages. As the figure demonstrates, of all users, only 13% visited pages from both categories (two switches) in one single session. This is the main reason why this categorization works this well.

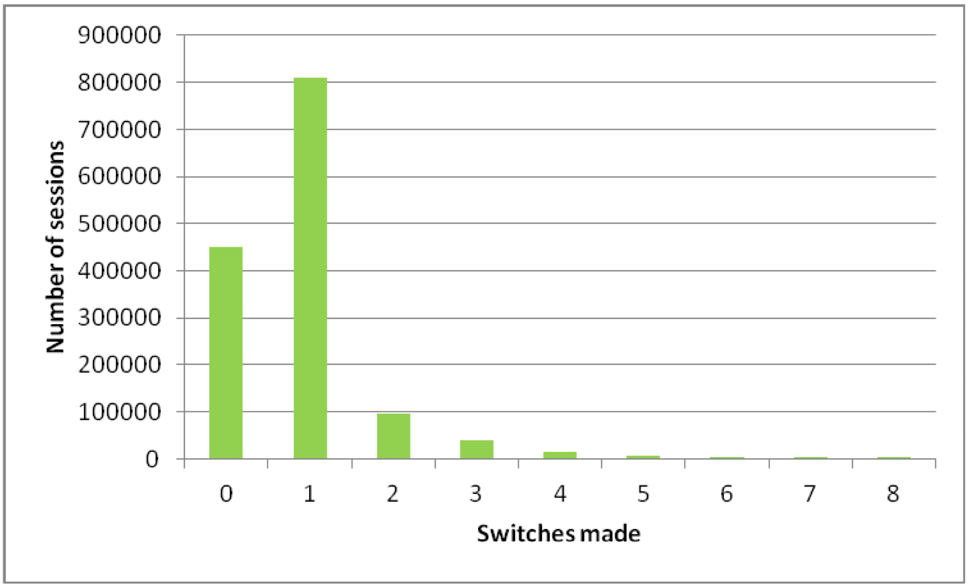

Figure 5: Number of switches between segments within a session

In sessions categorized as gadget, users visited less pages compared to sessions categorized as tech-savvy (t(212816) = -13.727, $p < 0.001$). This is demonstrated in figure six.

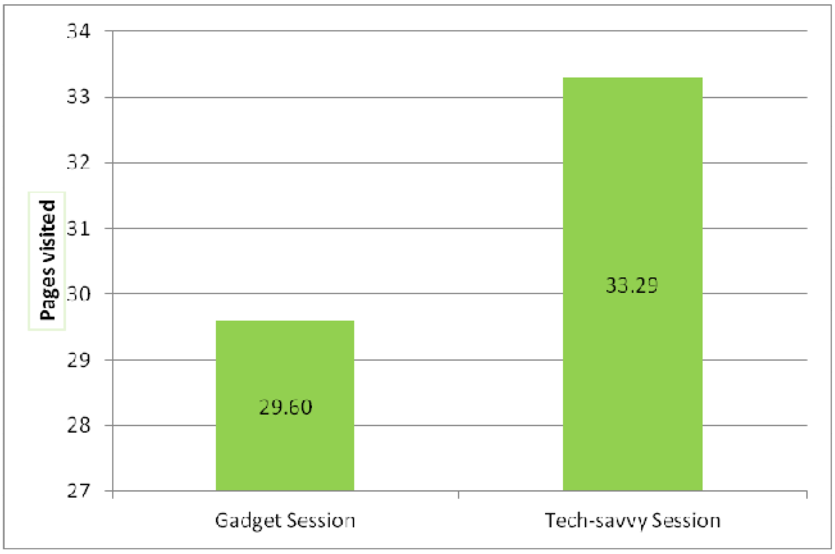

Figure 6: Comparison of number of pages visited between segments

In sessions categorized as gadget, users would visit the main page significantly less often compared to sessions categorized as tech-savvy (t(212813) = -35.3, $p < 0.001$).

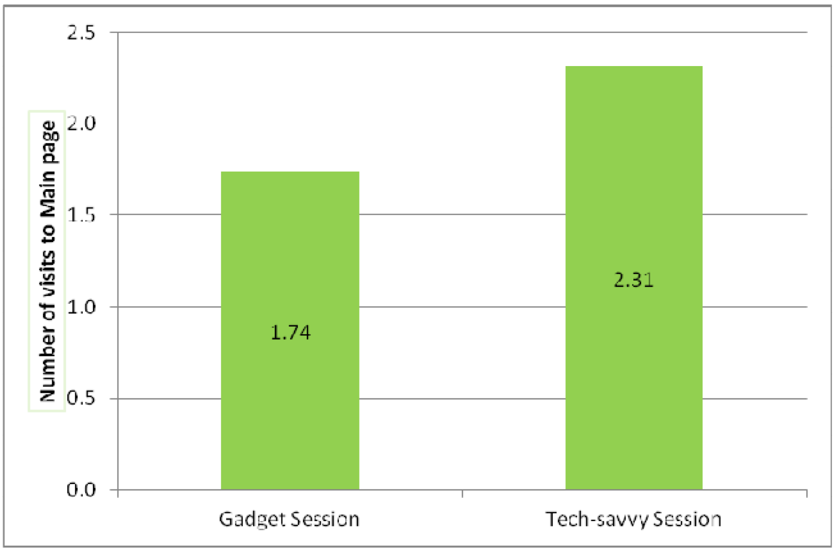

Figure 7: Comparison of number of visits to the main page between segments

In sessions categorized as gadget users also visited significantly less product groups compared to tech-savvy sessions ($t(212816) = -51.83$, $p < 0.001$). Again, this even holds when controlling for the fact that users in a tech-savvy session visit more pages compared to users in a gadget session ($F(1, 212815) = 2656.82$, $p < 0.001$).

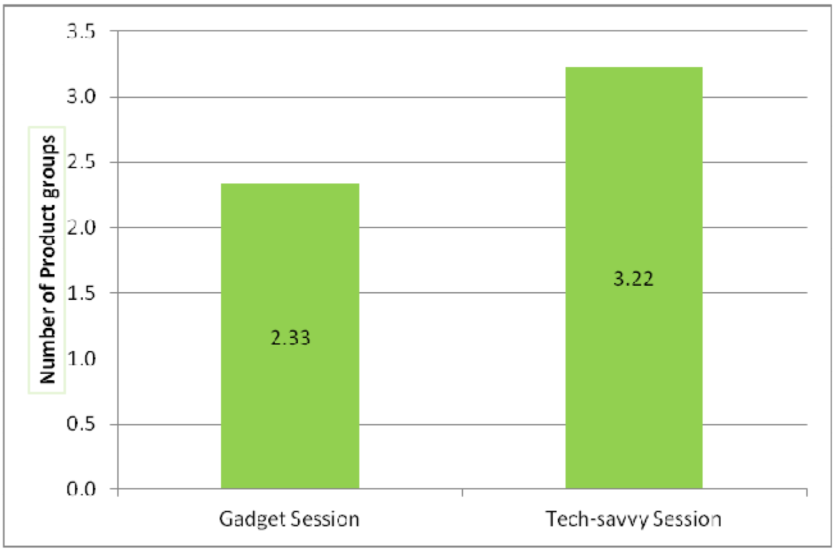

Figure 8: Comparison of number of product groups seen between segments

As a last example, in sessions categorized as gadget, users visited more review pages compared to users in sessions categorized as tech-savvy ($t(212813) = 10.83$, $p < 0.001$). This effect is enlarged when controlling for the fact that users in a tech-savvy session visit more pages compared to users in a gadget session ($F(1,21812) = 154.179$, $p < 0.001$).

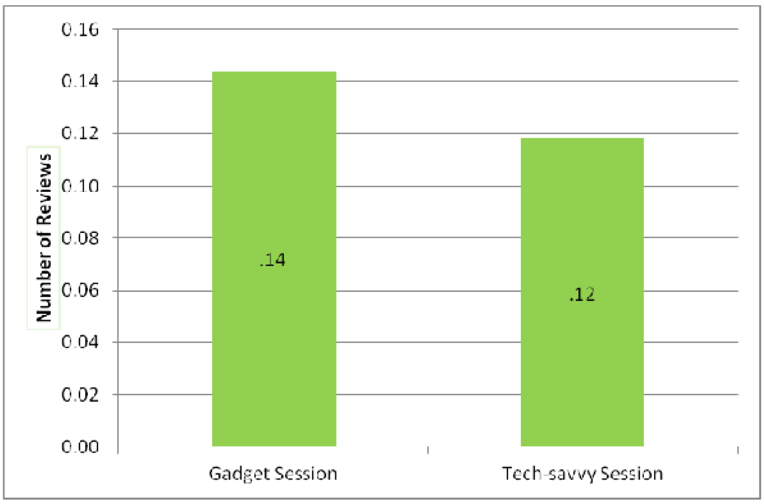

Figure 9: Comparison of review pages visited between segments

2.1.2 Users

The analyses above are about sessions, not about users yet. 456233 Users were differentiated in the data. A user made on average 3.2 (Mdn = 1) sessions. Part of the distribution of the amount of sessions users made is shown in figure 10. Of all users, 67.4% (307595 users) only made one session, only 5% of all users, 21268, made more than 10 sessions.

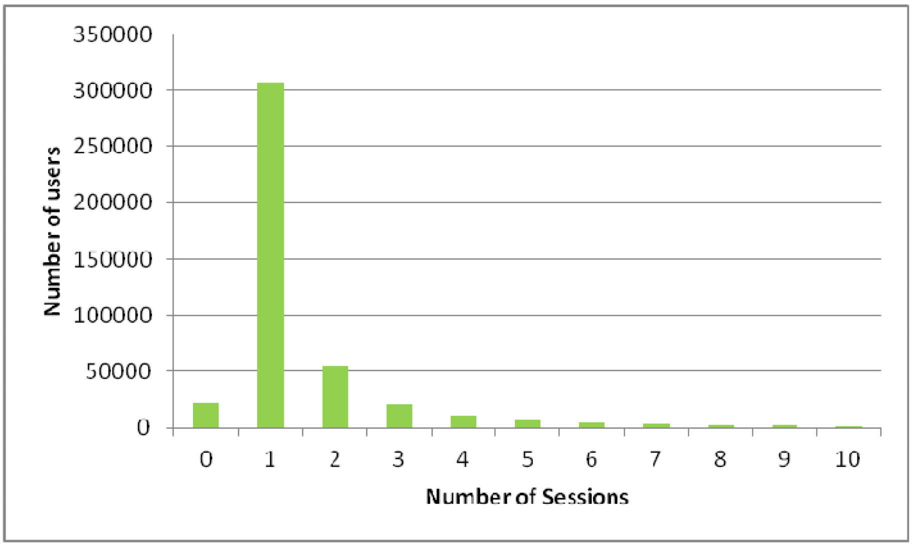

Figure 10: Number of sessions by users

Unfortunately, on average, there were only a very limited amount of users who consistently made long sessions. Ideally, users could be categorized in a similar way as sessions were categorized earlier in this chapter. This turned out to be very difficult since only very few users returned to Hardware.Info on a frequent basis. Only 21268 users made more than 10 visits to the website, 34525 made more than five sessions. For a similar kind of categorization as done with sessions, users would have to have a number of sessions that were categorized as tech-savvy or gadget, and most of the sessions had to be categorized the same.

Similar criteria were tested to see how many users could potentially be categorized as either tech-savvy or gadget. For only those users that made more than 10 sessions on the website, several ratios were tested. A ratio of 2 here means that of all sessions at least twice as many had to be from one category compared to the other category. With a ratio of two, only 11458 users could be categorized. That is only 2.5% of all users. With a ratio of 3, 13705 users (3%) could be categorized.

	Ratio = 3 > 10 sessions	Ratio = 2 > 10 sessions	Ratio = 1.5 > 10 sessions
Tech-savvy users	32.24% (N=6856)	38.46% (N=8180)	45.16% (N=9605)
Gadget users	13.23% (N=2814)	15.41% (N=3278)	19.28% (N=4100)
Not categorized	54.53% (N=11598)	46.13% (N=9810)	35.56% (N=7563)

Table 2: Categorization users with more than 10 sessions

The same kind of analysis can be done when setting the criteria of the amount of sessions to only five. Only 34525 users made more than 5 sessions.

	Ratio = 3 > 5 sessions	Ratio = 2 > 5 sessions	Ratio = 1.5 > 5 sessions
Tech-savvy users	31.46% (N=10861)	35.77% (N=12346)	40.81% (N=14090)
Gadget users	15.60% (N=5384)	17.13% (N=5914)	20.11% (N=6942)
Not categorized	52.95% (N=18280)	47.11% (N=16265)	39.08% (N=13493)

Table 3: Categorization users with more than 5 sessions

Even with these criteria, only 21032 users (4.6% of the total amount of users) can be categorized. This is still a very low and undesirable number of users.

2.1.3 Summary

This section on categorizing users and sessions showed that users visit the Hardware.Info website with a clear view on the content they are looking for. For example users want to read a certain news-items, or to read a review about a certain product. It is clear that when a user visits the website and starts reading tech-savvy items, it is very likely that user remains reading tech-savvy items for the remainder of that visit. These goals and purposes of users differ however between visits. The section above showed that only very few users make sessions on the website where they visit the same category of pages every time.

One reason why users stick to visiting pages of one segment within a session could be the way the Hardware.Info website is designed. When reading a certain news-item for example, Hardware.Info has a sidebar showing other relevant news items which usually belong to the same segment. One could argue that this could be a reason why a user stays in the same segment within a visit. Finally, it is clear that a categorization is not possible on a user-level, only three to five percent of all users can be categorized with the data available. Of all sessions, about 13 to 15 percent can be categorized. More than 72% of all long sessions can be categorized as either tech-savvy or gadget. This is however in line with the study done by Yan et al. (2009) who found that segmenting users based on behavior of one day was easier and more feasible compared to segmenting users based on an entire week.

2.2 Prediction model

The categorization of the last chapter is an important step in answering the main research question of this paper. However, categorizing visits is only part of the problem at hand. Not only is it necessary to be able to categorize a large amount of sessions, it is also necessary to predict in which category a visit falls based on only very limited amounts of information. Such a prediction model can then be used in an experimental setting. This section will discuss possible models to predict a visit on the Hardware.Info website.

In a possible prediction model, the dependent variable (the segment of the session) can have three different outcomes. A visit can be categorized as either tech-savvy, gadget or unknown. Therefore a multinomial logistic regression was used. This kind of model allows for more than two discrete outcomes.

The dataset as described in the previous section of this chapter held many different possible variables that could be used as independent variables in the regression analysis. Examples of these variables are for example the segment (gadget or tech-savvy) of the first couple of pages a users visited, the number of product groups the user visited pages of, the number of times a user switched between segments within a visit, the number of pages of a certain segment the user visited and which page types a user visited.

Many possible prediction models were tested to see which model could best predict in which segment a session would fall. For all models it was decided that only data of the first five clicks could be used in the model. This meant that in a possible experiment users would see a manipulation based on the segment of the session they are in after those five clicks. If, for example, 10 clicks would be used, 20% users less would see the manipulation. The simplest model is that with only taking into account the amount of gadget and tech-savvy pages the user visited in the first five clicks of that session. Results are shown in table four.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	73.99%	19.09%	6.93%
	Tech-savvy session	4.32%	88.82%	6.86%
	Not categorized session	19.94%	53.08%	26.98%

Table 4: Results prediction model 1

This model already does pretty well. Of all sessions that were categorized as tech-savvy, the model predicted 88.8% of those sessions correctly as tech-savvy. Also, the model predicted 74% of all gadget sessions correctly. The biggest flaw in this model is for sessions that are not categorized. The model only predicts 27% of all sessions that were not categorized, correctly.

Another model that was build used the number of tech-savvy and gadget pages within the first five clicks, and also the first page type (forum, news page, review etc.) the user visited. The model slightly improved, especially for those sessions that were not categorized.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	72.18%	20.41%	7.41%
	Tech-savvy session	4.24%	88.91%	6.85%
	Not categorized session	16.03%	38.49%	45.48%

Table 5: Results prediction model 2

In a third model, another variable was added, the amount of product groups the user saw within the first five pages of the session.

		Predicted		
		Gadget user	Tech-savvy user	Not categorized user
Categorized	Gadget user	70.94%	19.51%	9.56%
	Tech-savvy user	3.53%	87.96%	8.51%
	Not categorized user	12.48%	35.06%	52.46%

Table 6: Results prediction model 3

This model does even better than the previous one. More elaborate models were tested as well, with many other variables (the segment of the first five pages, the page type of the other pages visited within the first five clicks, the amount of switches made, and the amount of segments seen) . The improvement was in some cases significant, but the difficulty of the model rose as well with each variable added. It was decided that the model as shown in table six, with the amount of tech-savvy pages and the amount of gadget pages within the first five pages of a session, together with the amount of product groups seen in the first five pages of the session and the first page type a user landed on in a session was a good trade-off between the performance of the model and the difficulty of the model. For an overview of all tested models see appendix II.

These tests showed that simple models with only a limited amount of variables were able to predict a large amount of visits correctly. Even though more advanced models with more variables did significantly better in predicting visits, they did not make a big difference in terms of prediction power. An easier model has the advantage of being easier to implement and easier to understand. This model can be used in an experimental setting to show only the content of interest to the user. In the next chapter the details of the experiment that will answer the research question will be discussed.

3. Experimental design

The previous chapter showed and proved that two distinct segments of users visit the Hardware.Info website for different purposes. One segment of users visits the website because they are interested in tech-savvy content like for example a motherboard or a processor, the other segment because they are more interested in news and reviews about gadgets like for example a mobile phone or a tablet. The difference between both groups was evident, especially on a session-level.

Based on these findings and based on previous literature in this field, the main research question for this paper was defined as: *"Will targeting the content of a website real-time to a specific user segment result in a higher user experience and a difference in behavior on that website?"*

3.1 Hypotheses

This research question, together with the data analyses as described in chapter two led to several hypotheses. These hypotheses combined will help in answering the main research question. Two kinds of hypotheses can be defined. On the one hand there are several behavioral hypotheses. When targeting content on the Hardware.Info website to the specific segments as were in found chapter two, we expect a difference in behavior of the Hardware.Info user. On the other hand targeting content to those specific segments will also lead to a difference in user experience.

The more satisfied a user is with the content of a website, the longer that user will stay on that website. For the website itself it also important a user stays as long as possible on that website. The more pages a user visits, the more advertisements that user will be exposed to. Therefore,

- *H1: When presenting a user with content targeted towards the segment of that user, that user will visit more pages.*

During a visit, users visit pages that are linked with product groups. A news item on a new hard disk is linked to the product group hard disks. When users are confronted with only the content of a certain segment, more articles will be presumably be of interest to that user. More articles will result to a larger variety of product groups. Therefore,

- *H2: When presenting a user with content targeted towards the segment of that user, that user will visit pages linked to a larger variety of product groups.*

When users are presented more frequently with only the content of a certain segment, users have less opportunity to visit more than one segment. Therefore,

- *H3: When presenting a user with content targeted towards the segment of that user, that user will make less switches between segments*

3.2 Manipulation

Together with Hardware.Info several possible manipulations were discussed. Hardware.Info insisted on a manipulation that would not disrupt their users too much or would interfere with the main functionality of the website. On the other hand, the manipulation had to be large enough so users would actually see that something had changed.

Eventually it was decided to alter the sidebar Hardware.Info has on its website. Normally this sidebar would have links to relating items or products when users are reading news-items or review-items, but also to the most popular items on the website at that given moment.

Figure 11: Sidebar Hardware.Info

It was decided this would be replaced with a table showing the latest three reviews and the latest three news-items the user has not read yet.

An example of how this manipulation would look like is shown in figure 12. This kind of manipulation is very much in line with the concept of adaptive navigation support as introduced by Brusilovsky (1996). An example of adaptive navigation support was further explained in chapter one.

Figure 12: Manipulation

Based on the findings of chapter two and based on previous literature in this field, a 3x3 design was developed to answer the main research question. This design is shown in table seven.

The experiment will be done online in collaboration with Hardware.Info. Users will visit the website, their behavior is being monitored and used in the prediction model as described in chapter 2.2. After five clicks the model is able to predict in which segment the user falls, in that particular session. As soon as the prediction model has identified the session, the user in that session will be placed in one of the conditions as shown in the above table. At that moment the Hardware.Info website will change and the manipulation corresponding to the condition the user was put in, will be shown.

		Shown Segment		
		Tech-savvy Manipulation	Gadget Manipulation	Neutral Manipulation
Predicted Segment	Tech-savvy session	Congruent	Incongruent	Neutral
	Gadget session	Incongruent	Congruent	Neutral
	Not Categorized Session	Incongruent	Incongruent	Neutral

Table 7: Research design

As table seven shows, users visiting the Hardware.Info website will either be placed in a congruent condition or an incongruent condition. Finally, there is also a neutral condition which will be a combination between tech-savvy and gadget content and will serve as a baseline. A congruent condition is a condition where a user in a user segment is facilitated by a manipulation matching that segment. For incongruent conditions, where a user in a user segment is facilitated by a manipulation not matching that segment, opposite effects are expected.

Giving the manipulation, several other hypotheses specific to this sidebar can be defined as well. As the data analyses in chapter two showed, users often go back to the main page to look for content of their liking. When the sidebar shows more articles specifically tuned towards the interests of the user, that user does not have to go the main page anymore to look for interesting content. Therefore,

- *H4: When presenting a user with sidebar content targeted towards the segment of that user, that user will less often go back to the main page after reading a news item or article*

When the sidebar is showing more relevant content, users should also use the sidebar more often, therefore:

- H5: *When presenting a user with sidebar content targeted towards the segment of that user, that user will make more use of the sidebar.*

When the sidebar is showing more relevant content, users could visit the Hardware.Info website more often because they think there is more relevant content. This will result in more sessions.

- H6: When presenting a user with sidebar content targeted towards the segment of that user, that user will visit Hardware.Info more often.

The previous hypotheses were all about a change in behavior of the Hardware.Info users. Given the main research question of this study, there are also several hypotheses defined about the change in user experience. The different constructs that are hypothesized to change as a result of the experiment are derived from earlier research in the field of user experience (Knijnenburg et al. , 2011a). Relevant studies on user experience are discussed more in depth in chapter one.

When users are presented with content in the sidebar of their liking, the targeted content will be experienced as more relevant and of a higher quality. Therefore,

- *H7: The perceived accuracy (quality) of the Hardware.Info sidebar will go up*

When users are presented with content in the sidebar of their liking, the targeted content will be experienced as less varied. When the sidebar is only content relating to one specific segment, there will be less variation in the offered content. Therefore,

- *H8: When users are presented with content in the sidebar of their liking, the perceived variety of the sidebar will go down*

As a result of both the increase in accuracy and the decrease in variety, the effectiveness of the sidebar should go up. Users should experience the sidebar being of more use now that it is showing more relevant content. Therefore,

- *H9: When users are presented with content in the sidebar of their liking, the perceived sidebar effectiveness should go up as a result of an increase in perceived accuracy of the sidebar and an increase in perceived appeal of the sidebar*

Also, as a result of the increase in perceived accuracy and a decrease in variety the perceived appeal of the sidebar should go up. The sidebar should look more appealing when it is presenting content that is more in line with the segment of the user. Therefore,

- *H10: When users are presented with content in the sidebar of their liking, the perceived sidebar appeal should go up as a result of an increase in perceived accuracy of the sidebar and a decrease in perceived variety of the sidebar*

As a result of the increase in perceived effectiveness and the increase in perceived appeal, the behavior of the users should change as well. Not only is a direct effect between the different conditions and the behavior of users expected (see hypothesis one through six) but also an effect on the behavior of users because of a change in user experience. Therefore,

- *When users are presented with content in the sidebar of their liking, users will visit more pages on the website (H11), see pages linked to a larger variety of product groups (H12), make less switches between segments (H13), make less visits to the main page (H14) and finally use the sidebar more often (H15), all as a result of the change in user experience.*

Finally, It is also hypothesized that giving users in a session categorized as gadget, tech-savvy items will be experienced as worse compared to giving users in a session categorized as tech-savvy, gadget items (H16). This is also the main reason why it was decided to also have incongruent conditions.

When joining the hypotheses as described earlier, together with the behavioral measures as well as the measured user experience constructs in the questionnaire, the graph from figure 13 can be drawn.

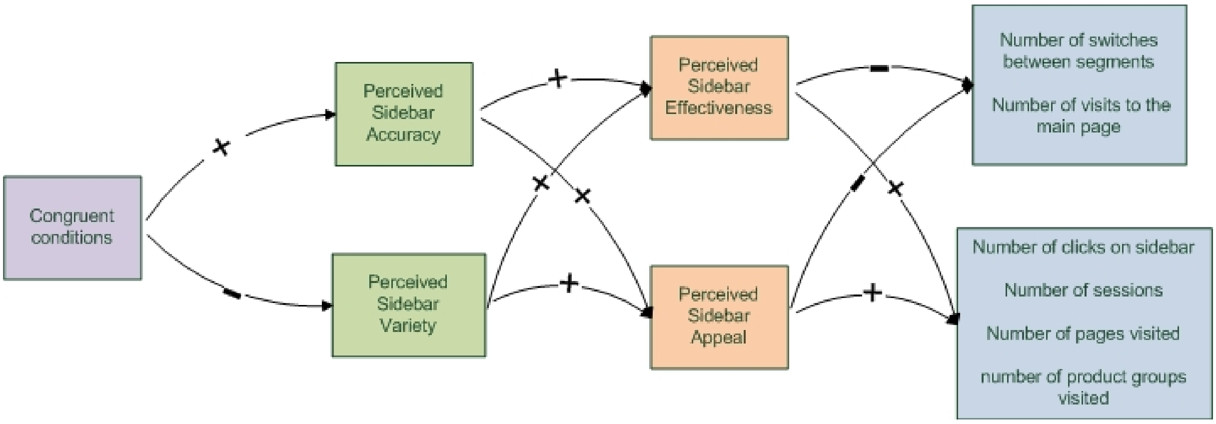

Figure 13: Hypotheses

3.3 Technical setup & Questionnaire

The behavior of users on the website will be measured using a pixel server. A pixel server is a tiny image of 1x1 pixels that is invisible on the website. By adding this image, a separate request to the source of that image is made every time a page on the website gets loaded. This request can then be used to store data about the behavior of that user. The following variables will be measured during the experiment.

- URL of the page
- Timestamp
- Predicted Segment
- Shown Segments
- Product group of the page
- Product linked to the page
- The type of page (i.e. review-page or a news-page)
- Whether a user has opened the questionnaire

The modifications to the sidebar are loaded using a JavaScript call. A JavaScript-script saves into a cookie which pages and how many pages a user visited and uses this data as input to the prediction model as described in chapter two.

When the user made five clicks the manipulation will be presented. When the manipulation is presented a link will be visible as well, guiding the user to the questionnaire.

Figure 14: Notification questionnaire

The questionnaire will consist of questions about the demographic information about a user, and several concepts relating to user experience. These constructs were taken from earlier research on the topic of user experience (Knijnenburg et al., 2011a). The following constructs were measured:

- Perceived sidebar quality/accuracy
- Perceived sidebar variety
- Perceived sidebar effectiveness
- User interaction satisfaction
- Strength of preference for one segment

For the full questionnaire see appendix III. Figure 15 shows a part of one of the pages of the questionnaire.

Figure 15: Questionnaire

An incentive was available for those who fully answered the questionnaire. Participants in the survey could either win a SSD or a mobile phone. These prices were selected because they also fall into the two segments. A SSD is a typical tech-savvy incentive whereas a mobile phone is a typical 'gadget' incentive. At the end of the survey participants had to leave their email address behind but also had to pick one of the two prices. The selection of which prize a user wants to receive can be an additional test of the prediction model. Also, the stated strength of preference each user indicates can be another measure of the segment of that user.

4. Results

In this chapter the results of the experiment as described in the previous chapter will be presented. The experiment ran from December 6, 2012 until December 20, 2012 live on the Dutch Hardware.Info website. During those two weeks roughly 2.2 million page requests were made. Each page request was saved into a database with a corresponding unique user-id and a unique session-id. Users were identified based on cookies. A new session-id was created every time a user made a page request when the previous page request was more than 15 minutes prior to that request.

The results will be split up in two distinct sections. First the behavioral results will be presented. Given the manipulation in the experiment we expect a difference in behavior between users in different experimental conditions. Second, the results of the analyses on the questionnaire data will be presented. Given the manipulation in the experiment we expect a difference in experience between users in different experimental conditions.

4.1 Behavioral results

In total 340854 users visited the website while the experiment was ongoing. Users which did not make at least five clicks (and therefore did not see the manipulation) or users who visited an unrealistic high number of pages or made an unrealistic amount of sessions were left out of the analyses. For most dependent variables the top 1 % of all data was removed. Those kinds of 'users' would greatly bias the results and should therefore be removed. 70604 users were eventually used in the behavioral analysis.

Users were categorized based on the first five clicks they made during their visit or visits. Figure 16 shows the number of participants of each category. 25623 users (36.3%) were categorized as being gadget users, 37549 users (53.2%) were categorized as tech-savvy users. For the remaining 7435 (10.5%) users the prediction algorithm was not able to assign them to a segment. Compared to the pre-experimental data analysis slightly more users were categorized as gadget or tech-savvy. One reason for this could be because the experiment was done using cookies as identification where in the pre-experimental data analyses IP-addresses were used.

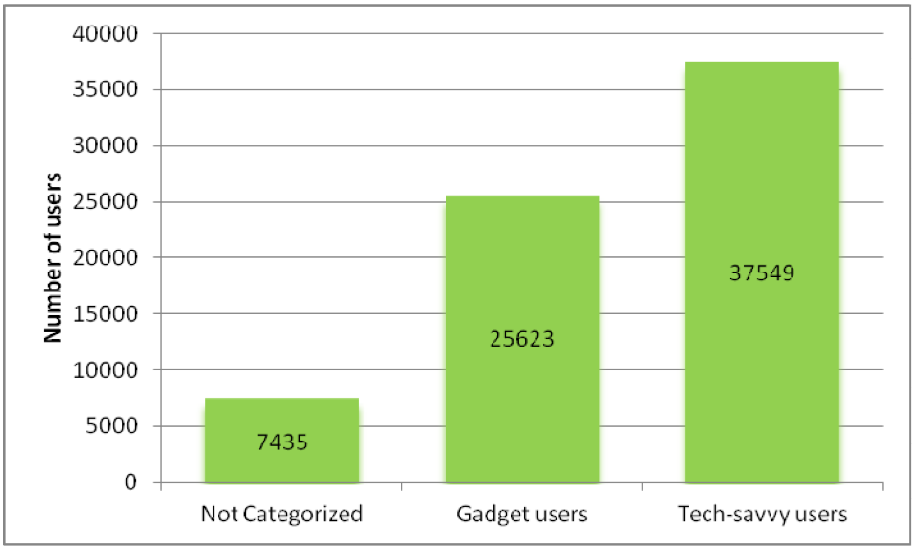

Figure 16: Number of users per segment

Each user, regardless of the categorization, was randomly assigned to a manipulation. That manipulation was either a gadget manipulation where only gadget items would appear in the sidebar, a tech-savvy manipulation where only tech-savvy items would appear in the sidebar or a neutral condition where both tech-savvy and gadget items would appear. Table 8 shows the amount of participants per condition.

		Manipulation		
		Gadget	Tech-savvy	Neutral
Predicted Segment	Tech-savvy	12516	12527	12506
	Gadget	8506	8493	8624
	Not Categorized	2475	2507	2453

Table 8: Number of users per condition

Several dependent variables were derived from the dataset and used to test the hypotheses as described in chapter three. These variables are:

- The number of sessions users made
- The number of visits to the main page
- The number of product groups visited
- Amount of switches between segments
- The usage of the sidebar
- The number of pages visited

In the rest of this chapter, the results of the experiment will be presented based on those different variables.

All these variables are count variables. It is common to analyze count variables using so-called poisson regressions. However, it turned out all variables measured in the experiment where heavily over-dispersed. Dispersion is a measure of the extent to which data are spread around an average. It is called over-dispersion when the variance of a variable is higher than the mean, where it should be equal. Therefore, Negative binomial regression models were used to analyze the effects between all the conditions. Negative binomial regression modeling is a special kind of regression for modeling count variables which can handle over-dispersion better than poisson regression models. (Cameron & Trivedi, 1998).

To be able to state whether the hypotheses as defined in chapter three were either true or false, effect coding was used in all the negative binomial regression models. With effect coding it is easier to isolate the effect of a certain manipulation or interaction compared to dummy coding (Alkharusi, 2012). For the regression models to be interpretable, the users that were not categorized were left out of the analyses.

4.1.1 Number of sessions

Users made on average 4.29 sessions (SD=6.125, Mdn=2). Part of the distribution of the amount of sessions is shown in figure 17. A large amount of users (39.5%) made only 1 session. 90% of all users made less than 10 sessions. These numbers are slightly higher than we found in the pre-experimental data analysis which again can be caused by a different way of identifying users.

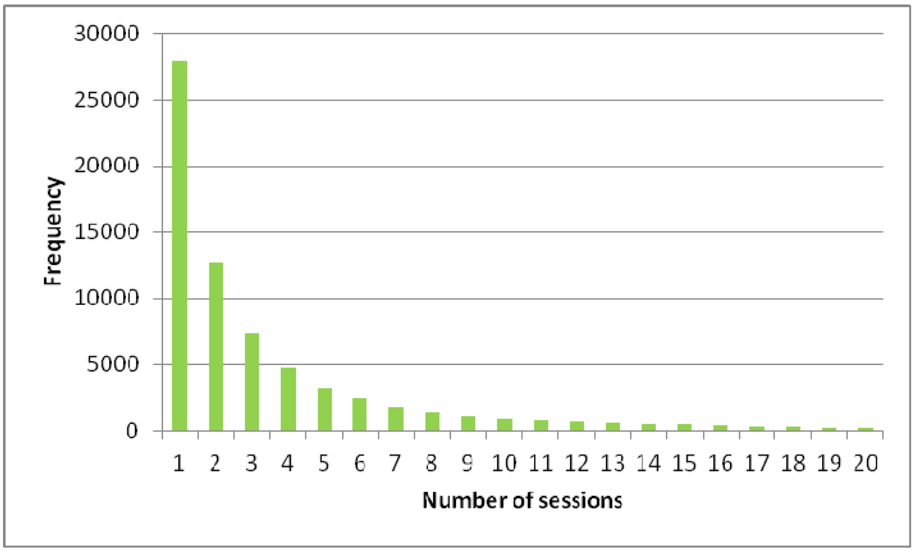

Figure 17: Number of sessions

It was hypothesized in chapter three that users in congruent would make more sessions compared to users in incongruent conditions (H6). Also, it was hypothesized that showing users who were categorized as being gadget, tech-savvy content would be worse in terms of behavior and user experience compared to tech-savvy users presented with gadget content (H16).

Parameter	B	Std. Error	Sig.	Exp(B)
(Intercept)	1.483	.0099	0.000	4.404
Predicted Segment = Gadget	-.140	.0157	.000	.869
Predicted Segment = Tech-savvy	0			1
Users who experienced a neutral manipulation	-.007	.0140	.625	.993
Users in a incongruent condition	-.010	.0140	.460	.990
Users in a congruent condition	0			1
Gadget users in a neutral condition	.013	.0221	.552	1.013
Gadget users in an incongruent condition	-.007	.0222	.740	.993
Gadget users in a congruent condition	0			1

Table 9: Number of sessions - regression table

Users that were categorized differently also behaved differently in respect to the amount of sessions they made. Users categorized as gadget made 13.1% less sessions on the Hardware.Info website compared to tech-savvy users (Exp(B) = 0.869, $p < 0.001$). There was no difference found between congruent and incongruent conditions. Therefore, H6 is not true. Also the effect of incongruency is the same for gadget users and tech-savvy users. Therefore, H16 is not true for this particular type of behavior. An overview of the different means per condition is presented in figure 18.

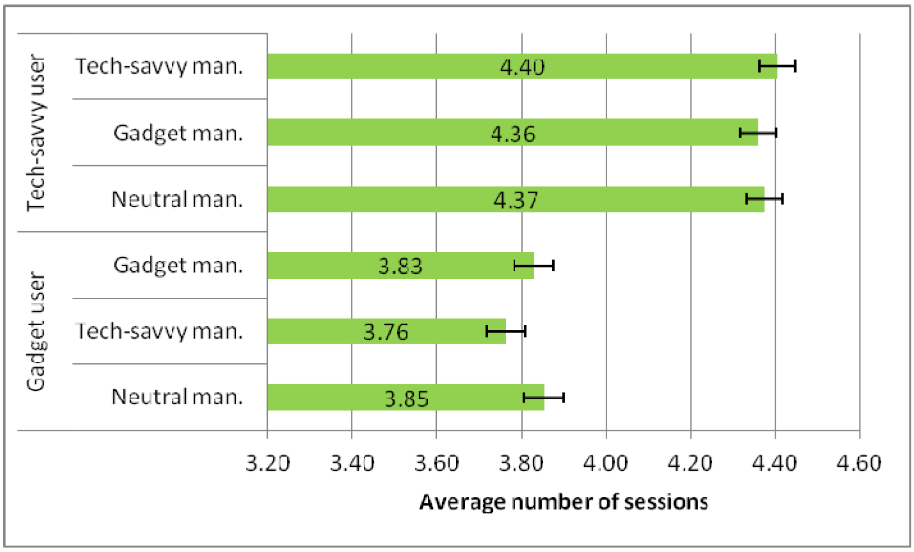

Figure 18: Number of sessions - means per condition

4.1.2 Number of visits to the main page

A large portion of all users never visited the main page on the website. This can occur when users follow a link from a different website or from an RSS-feed. Part of the distribution of main page visits is shown in figure 19. On average users visited the main page 4.78 times during the weeks of the experiment (SD=11.60, Mdn=1).

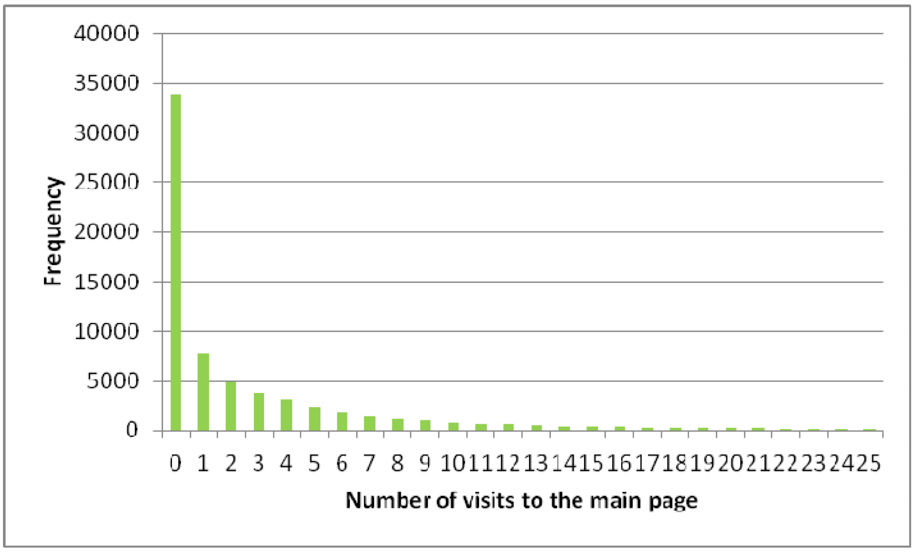

Figure 19: Visits to main page

Hypotheses four states that users in congruent would visit the main page less often compared to users in incongruent conditions (H4). Also, it was hypothesized that showing users who were categorized as being gadget, tech-savvy content would be worse in terms of behavior and user experience compared to tech-savvy users presented with gadget content (H16).

Parameter	B	Std. Error	Sig.	Exp(B)
(Intercept)	1.600	.0098	0.000	4.955
Predicted Segment = Gadget	-.234	.0156	0.000	.792
Predicted Segment = Tech-savvy	0			1
Users who experienced a neutral manipulation	.020	.0138	.141	1.021
Users in a incongruent condition	.026	.0138	.059	1.027
Users in a congruent condition	0			1
Gadget users in a neutral condition	.005	.0220	.825	1.005
Gadget users in an incongruent condition	-.043	.0221	.049	.957
Gadget users in a congruent condition	0			1

Table 10: Visits to main page - regression table

As previously seen with the amount of sessions as well, there is a large effect between the different segments. Users categorized as gadget made significantly less visits to the main page compared to tech-savvy users. Gadget users visited the main page 20.8% less compared to tech-savvy users (Exp(B) = 0.792, $p < 0.001$). As table 10 shows, there is a marginal significant difference between users in a congruent and users in an incongruent condition (Exp(B) = 1.027, $p < 0.06$), also, gadget users in an incongruent condition made less visits to the main page compared to tech-savvy users in an incongruent condition (Exp(B) = 0.957, $p < 0.05$). These effects combined result in tech-savvy users congruent users visiting the main page a little less compared to gadget users. This relation is not significant, but a trend is visible. All other differences between conditions are not significant either. An overview of the different means per condition is presented in figure 20.

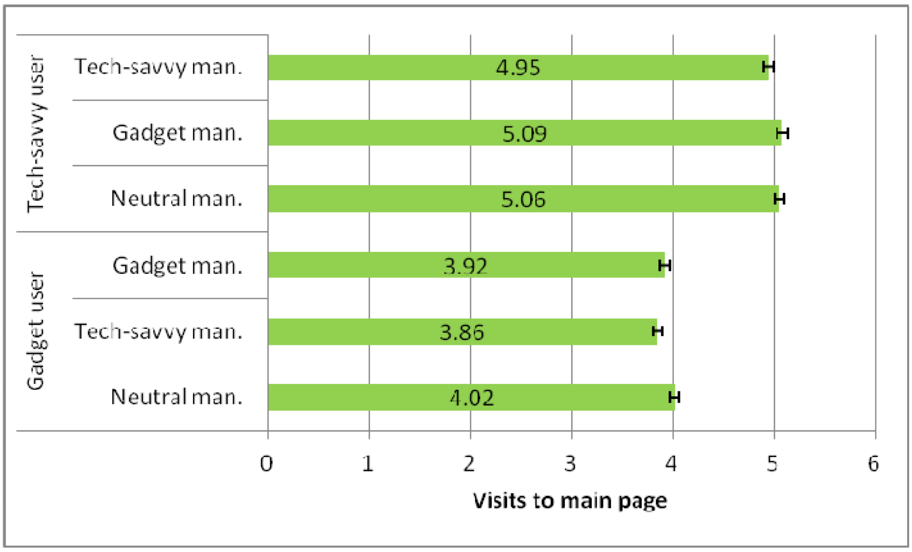

Figure 20: Visits to main page - means per condition

4.1.3 Number of product groups

Figure 21 shows part of the distribution of the amount of different product groups users visited pages of during their time on the Hardware.Info website. Most Hardware.Info users are clearly interested in only a very small range of products. This trend was clear also during the timeframe of the experiment. Users visited only a small range of different products. Each product on the Hardware.Info website is linked to a product group. A product group is for example processors, but also motherboards, tablets or hard drives. In total there are 116 different kinds of product groups. On average users saw 3.17 product groups during the two weeks of the experiment (SD=3.46, Mdn=2). A lot of visitors, 27092 (38.4%) only saw one product group. Only 5% of all visitors saw more than 10.

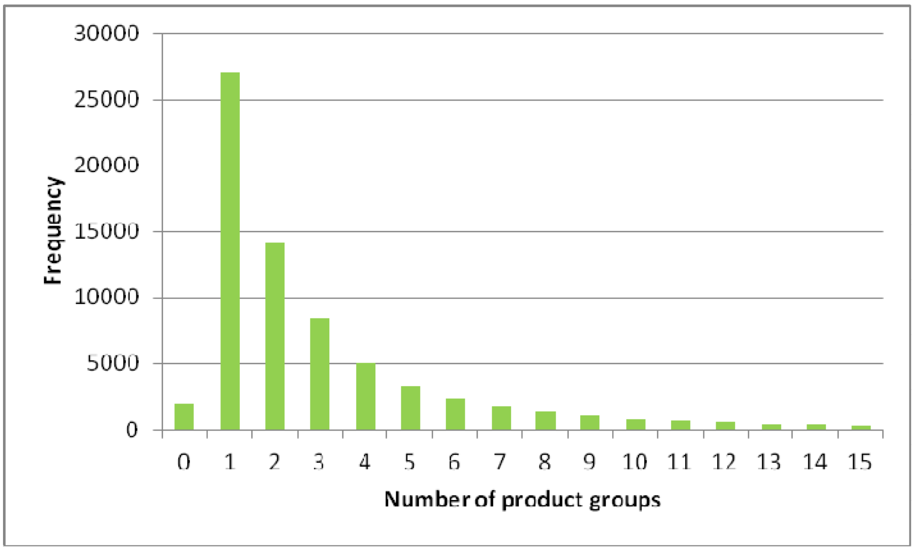

Figure 21: Product groups visited

In chapter three it was hypothesized that users in congruent would visit more product groups compared to users in incongruent conditions (H2). Also, it was hypothesized that showing users who were categorized as being gadget, tech-savvy content would be worse in terms of behavior and user experience compared to tech-savvy users presented with gadget content (H16).

Parameter	B	Std. Error	Sig.	Exp(B)
(Intercept)	1.119	.0103	0.000	3.062
Predicted Segment = Gadget	-.114	.0163	.000	.892
Predicted Segment = Tech-savvy	0			1
Users who experienced a neutral manipulation	.010	.0146	.503	1.010
Users in a incongruent condition	.007	.0146	.649	1.007
Users in a congruent condition	0			1
Gadget users in a neutral condition	-.022	.0230	.342	.978
Gadget users in an incongruent condition	-.031	.0231	.182	.970
Gadget users in a congruent condition	0			1

Table 11: Product groups visited - regression table

Users categorized as gadget visited significantly less product groups compared to tech-savvy users (Exp(B) = 0.892, $p < 0.001$). Users categorized as gadget visited 10.8% less product groups compared to users in the tech-savvy condition.

As table 11 shows, there is no significant difference between congruent and incongruent conditions. Therefore, hypothesis two is not true. No significant differences can be found in the interaction between segment and manipulation, the effect of incongruency is the same for gadget users and tech-savvy users. Therefore, H16 is not true for this particular type of behavior. An overview of the different means per condition is presented in figure 22.

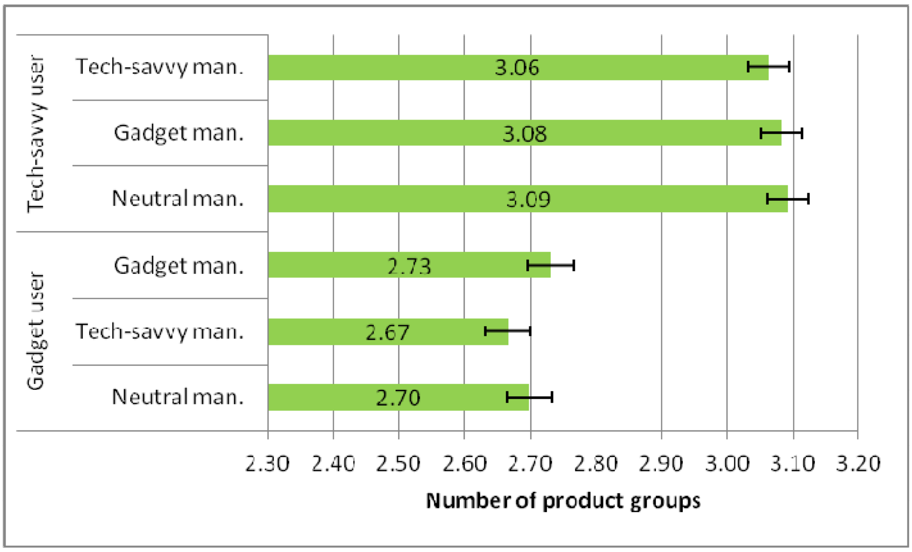

Figure 22: Product groups visited - means per condition

4.1.4 Amount of switches between segments

Product groups were key in the initial distinguishing between gadget and tech-savvy users. Certain product groups like for example motherboards and processors were marked as tech-savvy, others like tablets or mobile phones were marked as gadget. One of the interesting findings during the pre-experimental data analysis was that users almost never switched between those segments. Most users only visited tech-savvy product groups or gadget product groups. This trend was also visible during the time of the experiment as figure 23 shows. On average a user switched 0.49 (SD=1.50, Mdn=0.00) times during all their visits on Hardware.Info. Only 19.6% of all users switched once or more and only 1.7% switched more than five times.

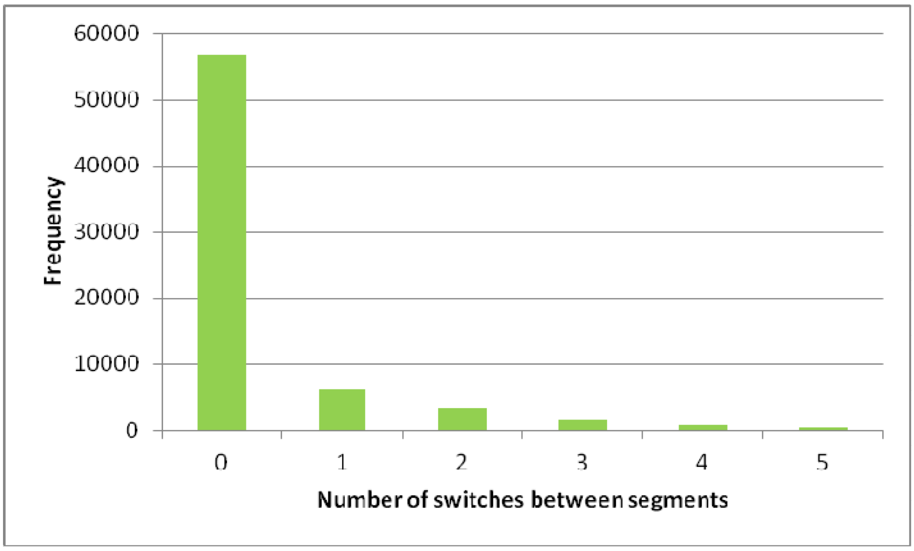

Figure 23: Switches made

In chapter three it was hypothesized that users in congruent conditions (users who were presented with a manipulation that fitted with the segment of that user) would make fewer switches between segments compared to users in incongruent conditions (H3). Also, it was hypothesized that showing users who were categorized as being gadget, tech-savvy content would be worse in terms of behavior and user experience compared to tech-savvy users presented with gadget content (H16).

Parameter	B	Std. Error	Sig.	Exp(B)
(Intercept)	-.951	.0169	0.000	.387
Predicted Segment = Gadget	-.007	.0266	.803	.993
Predicted Segment = Tech-savvy	0			1
Users who experienced a neutral manipulation	.046	.0237	.054	1.047
Users in a incongruent condition	.102	.0235	.000	1.107
Users in a congruent condition	0			1
Gadget users in a neutral condition	-.193	.0381	.000	.825
Gadget users in an incongruent condition	-.193	.0378	.000	.824
Gadget users in a congruent condition	0			1

Table 12: Switches made - regression table

For the amount of switches, there is no significant difference between gadget users and tech-savvy users. Given the pre-experimental data analysis, this was not expected. Users in an incongruent condition made more switches between segments compared to users in the congruent condition (Exp(B) = 1.107, $p < 0.001$). But, table 12 also shows that gadget users in an incongruent condition make 17.6% less switches compared to tech-savvy incongruent users (Exp(B)=0.824, $p < 0.001$). The effect of congruency together with the interaction effect leads to the result that tech-savvy users in an incongruent condition switch more than users in the tech-savvy congruent condition. It also means that gadget users in an congruent condition switch less than gadget users in a congruent condition. This is demonstrated by figure 24.

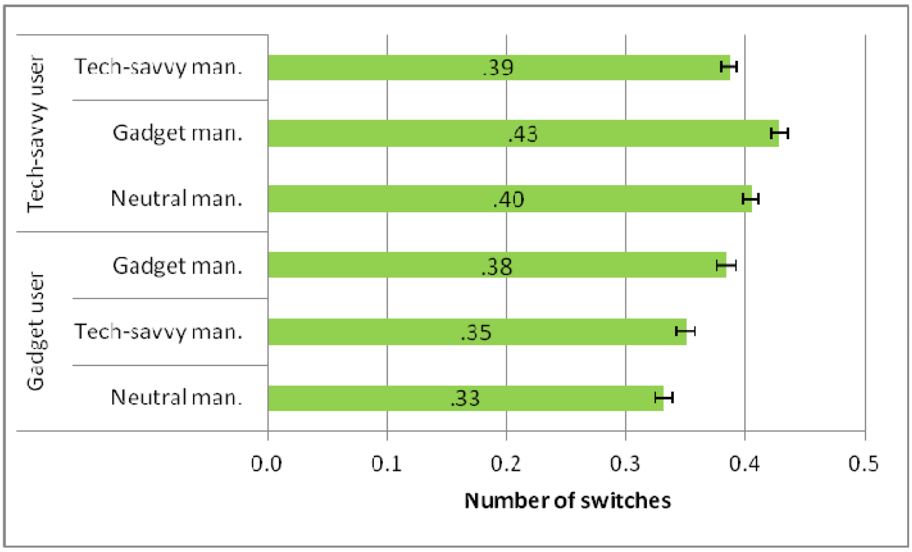

Figure 24: Switches made - means per condition

4.1.5 Clicks on the sidebar

During the pre-experimental data analysis, the sidebar on Hardware.Info was not often used. This trend did not change during the weeks of the experiment as table 13 shows. On average a user clicked on the sidebar 0.04 times (SD=0.244, Mdn=0.00) during the experiment. Because of the large difference in frequency between zero clicks on the sidebar and one click on the sidebar, a table is presented instead of a graph.

Number Of Sidebar Clicks	Frequency	Percentage
0	68649	97.23
1	1632	2.31
2	222	0.31
3	59	0.08
4	19	0.03
5	13	0.02
6	5	0.01
7	6	0.01
8	1	0.00
9	1	0.00

Table 13: Clicks on sidebar

In chapter three it was hypothesized that users in congruent conditions (users who were presented with a manipulation that fitted with the segment of that user) would use the sidebar more often compared to users in incongruent conditions (H5). Also, it was hypothesized that showing users who were categorized as being gadget, tech-savvy content would be worse in terms of behavior and user experience compared to tech-savvy users presented with gadget content (H16).

Parameter	**B**	**Std. Error**	**Sig.**	**Exp(B)**
(Intercept)	-3.074	.0425	0.000	.046
Predicted Segment = Gadget	-.433	.0765	.000	.649
Predicted Segment = Tech-savvy	0			1
Users who experienced a neutral manipulation	-.250	.0641	.000	.779
Users in a incongruent condition	-.466	.0681	.000	.627
Users in a congruent condition	0			1
Gadget users in a neutral condition	.057	.1138	.619	1.058
Gadget users in an incongruent condition	.351	.1149	.002	1.421
Gadget users in a congruent condition	0			1

Table 14: Clicks on sidebar - regression table

Users categorized as gadget used significantly less often the sidebar compared to tech-savvy users ($Exp(B) = 0.649$, $p < 0.001$). Users categorized as gadget used the sidebar 35.1% less compared to tech-savvy users.

Users in incongruent conditions use the sidebar 37.3% less compared to users in congruent conditions ($Exp(B) = 0.627$, $p < 0.001$). This means hypothesis five is true. Table 14 also shows that gadget users in an incongruent condition use the sidebar 42.1% more compared to tech-users in an incongruent condition ($Exp(B) = 1.421$, $p < 0.003$). Together with the main effect of gadget users compared to tech-savvy users, that means that the effect of incongruency only affects tech-savvy users. That is also demonstrated by figure 25 below. The difference between gadget congruent users and gadget incongruent users is very small, but there is a very large effect between tech-savvy congruent users and tech-savvy incongruent users.

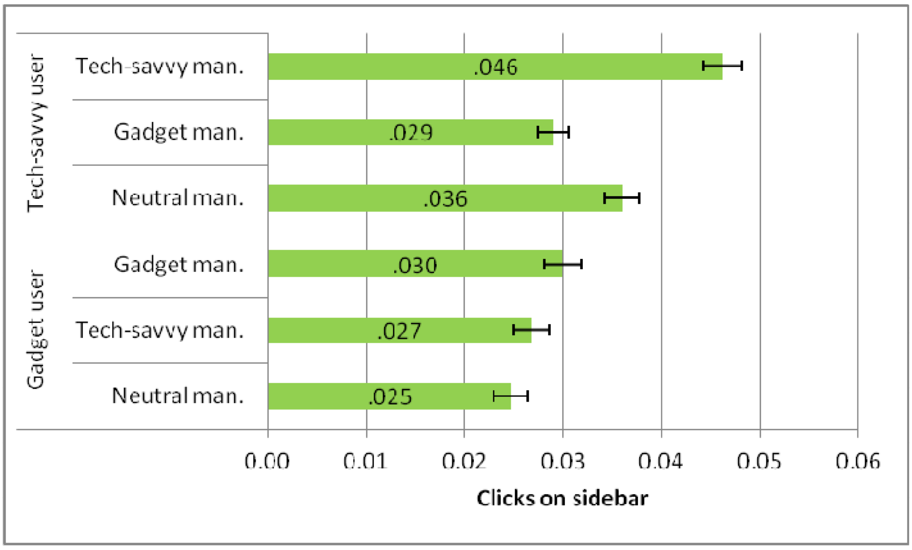

Figure 25: Clicks on sidebar - means per condition

4.1.6 Number of clicks

Users visited on average 21.76 (SD=27.434, Mdn=12) pages on the Hardware.Info website during the two weeks of the experiment. Figure 26 shows part of the distribution of the amount of clicks. Users who made 5 clicks or less were removed from the manipulation because they did not see the manipulation. As the graph shows, there are many users who only visited six (12% of all users used in the analyses) or seven (10.1% of all users used in the dataset) pages on the website. 91.2% of all users visited 50 pages or less, 97.5% visited 100 pages or less. These numbers clearly show the long tail this distribution has.

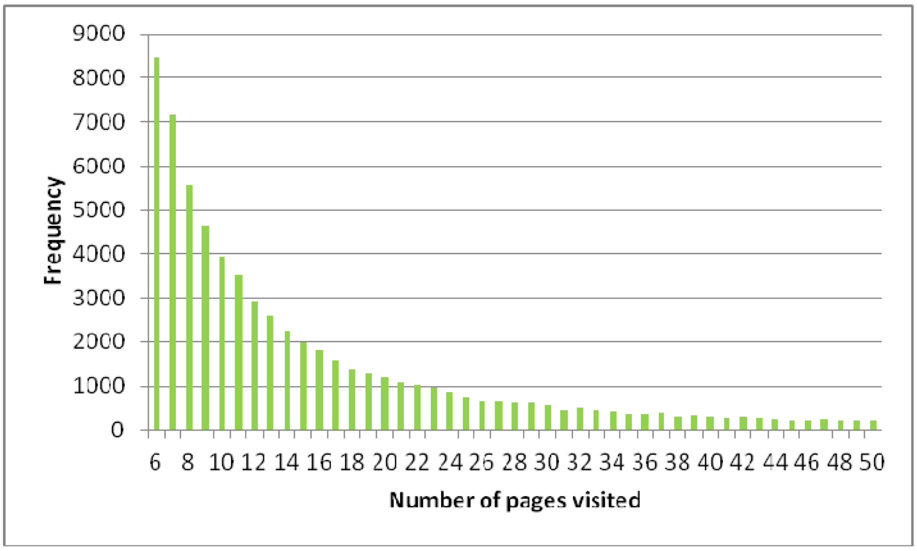

Figure 26: Number of pages visited

In chapter three it was hypothesized that users in congruent conditions (users who were presented with a manipulation that fitted with the segment of that user) visit more pages compared to users in incongruent conditions (H1). Also, it was hypothesized that showing users who were categorized as being gadget, tech-savvy content would be worse in terms of behavior and user experience compared to tech-savvy users presented with gadget content (H16).

Parameter	B	Std. Error	Sig.	Exp(B)
(Intercept)	3.114	.0091	0.000	22.521
Predicted Segment = Gadget	-.130	.0144	.000	.878
Predicted Segment = Tech-savvy	0			1
Users who experienced a neutral manipulation	.007	.0129	.605	1.007
Users in a incongruent condition	.004	.0129	.732	1.004
Users in a congruent condition	0			1
Gadget users in a neutral condition	-.036	.0203	.074	.964
Gadget users in an incongruent condition	-.031	.0204	.126	.969
Gadget users in a congruent condition	0			1

Table 15: Number of pages visited - regression table

Gadget users made 12.2% less clicks compared to tech-savvy users (Exp(B) = 0.878, $p < 0.001$). No significant difference can be found between levels of congruency. Users in congruent conditions visited the same number of pages as users in incongruent conditions. Given that result it can be concluded that H1 is not true.

No significant differences can be found in the interaction between segment and manipulation, the effect of incongruency is the same for gadget users and tech-savvy users. Therefore, Hypothesis 16 can be is not true for this particular type of behavior. An overview of the different means per condition is presented in figure 27.

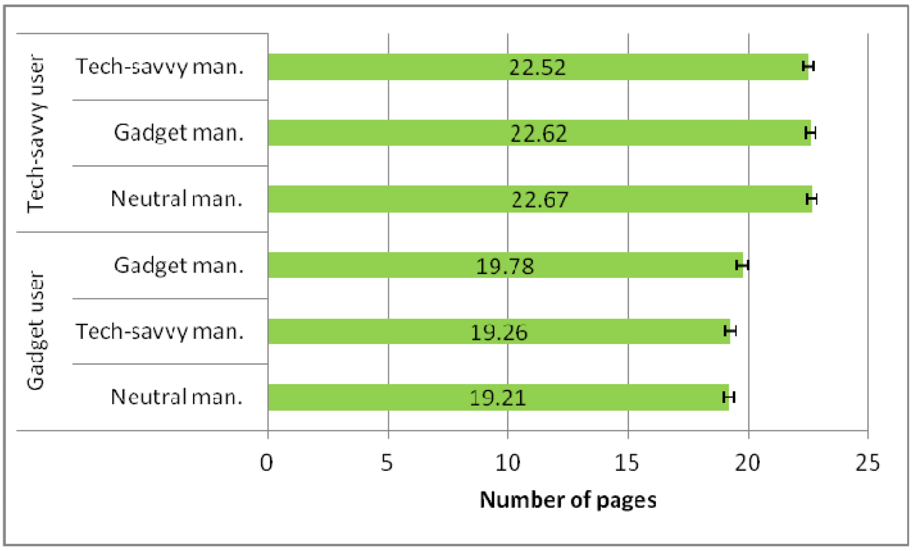

Figure 27: Number of pages visited - means per condition

4.2 Questionnaire data

Of all Hardware.Info users, 5818 users started with the survey. Several measures were taken to verify that users filled out the survey completely and in a serious matter. Users that needed less than 10 seconds to complete any of the pages of the survey were removed from the analyses. Also, users that gave the same score to all questions, regardless whether these questions were formulated as positive or negative, were not used in the analyses. This led to 2752 complete surveys.

Figure 28 shows the number of participants per category. Of all users who correctly filled out the survey, 443 (16%) were users that were not categorized. 791 (29%) of all users were categorized as a gadget user, 1518 (55%) were categorized as tech-savvy. Compared to the users in the behavioral dataset (see figure 16), 7% less gadget users filled out the questionnaire completely. The tech-savvy category is almost equal in percentage of participants.

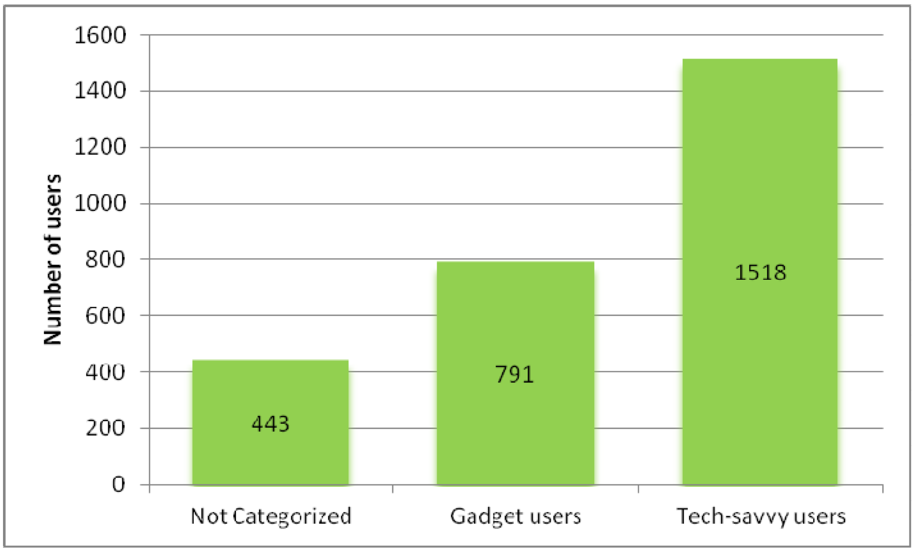

Figure 28: Number of users per segment

Each user was randomly assigned to a manipulation. That manipulation was either a gadget manipulation where only gadget items would appear in the sidebar, a tech-savvy manipulation where only tech-savvy items would appear in the sidebar, or a neutral condition. Table 16 shows the amount of participants per condition.

		Manipulation		
		Gadget	Tech-savvy	Neutral
Predicted Segment	Tech-savvy	491	522	505
	Gadget	267	274	250
	Not Categorized	128	166	149

Table 16: Number of participants per condition

The first questions users got were demographic questions. The average age of users was 31.4 years old (SD= 15,4) and most users were male (99% male, 1% female). Most users stated they were single (52%), 20% stated they were living together with a partner, 28% of all users were married and 1% was divorced. Users were also asked about their degree. 40% had a vocational education, 36% of all users stated they had an undergraduate degree and 11% stated they a graduate degree.

Users could also state their own preference for a certain segment. Of all gadget users, 45% stated they were very interested in gadget content, and another 32% stated they were at least a little interested. Interestingly, of all tech-savvy users, 35% stated they were very interested in gadget content, and another 35% stated they were a little interested. Figure 29 shows the distribution of all gadget and tech-savvy users and their interest in gadget content. Even though gadget users were significantly more interested in gadget content compared to tech-savvy users ($p < 0.001$), this distribution shows that a large number of tech-savvy users indicated they were interested in gadget content even though their behavior does not portrait this.

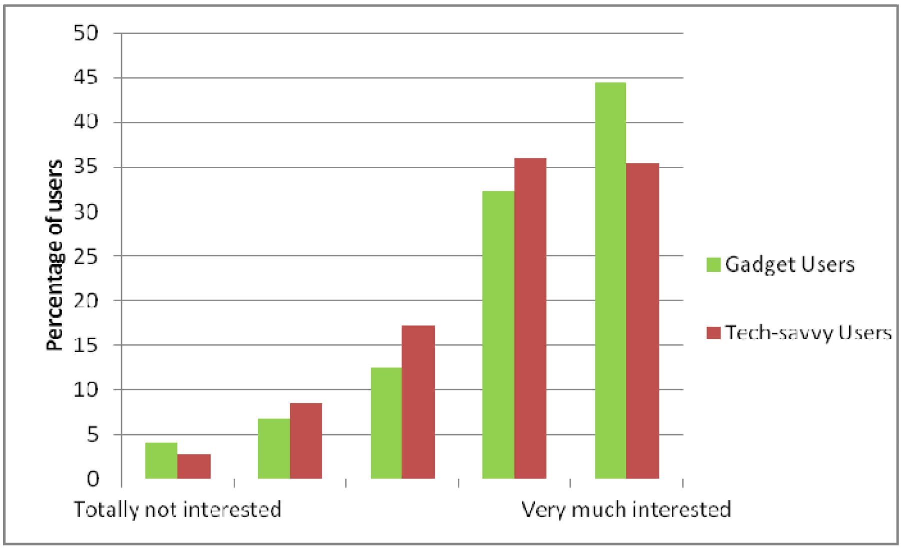

Figure 29: Interest in gadget content

Of all gadget users, 54% stated they were very much interested in tech-savvy content. Another 27% stated they were a little bit interested. Of all tech-savvy users, 65% stated they were very much interested in tech-savvy content, another 22% stated they were at least a little bit interested. Figure 30 shows the distribution of all gadget and tech-savvy users and their interest in tech-savvy content. Tech-savvy users were significantly more interested in tech-savvy content compared to gadget users ($p < 0.001$), but even though their behavior does not provide evidence for it, this figure shows that a lot of gadget users at least think they are very interested in tech-savvy content as well.

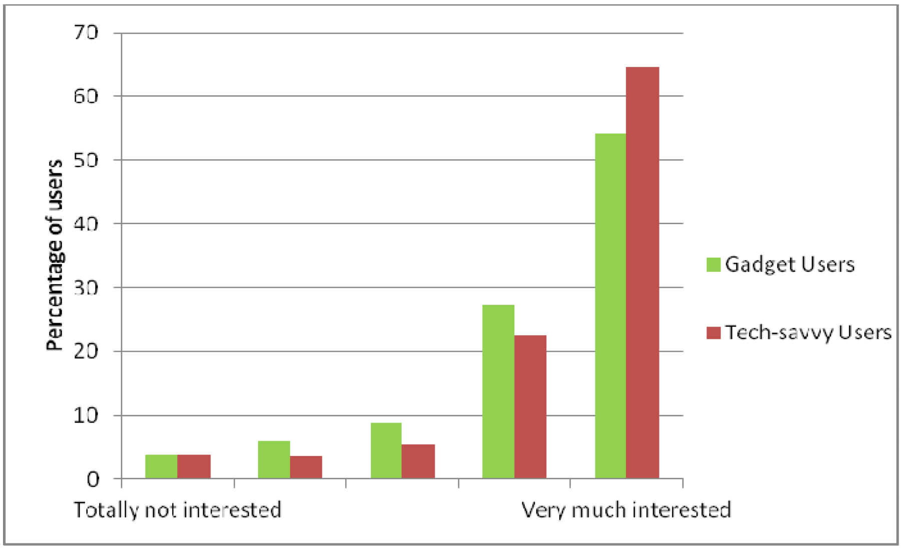

Figure 30: Interest in tech-savvy content

At the end of the survey users could pick which prize they wanted to win in the lottery. There was a tech-savvy prize (a solid-state drive roughly worth 100 euro), and a gadget prize (a smartphone roughly worth 200 euro) users could choose between. 592 (22%) of all users picked the mobile phone, 2143 users (78%) picked the SSD. 17 users did not want to win a prize. Even though the SSD was by far the most popular prize, gadget users chose significantly more often the mobile phone compared to tech-savvy users (chi-square(4, N = 2792) = 25.16, $p < 0.01$). This is yet another confirmation that the prediction algorithm worked.

All other questions in the questionnaire were related to concepts of user experience. These constructs were taken from earlier research on the topic of user experience (Knijnenburg et al., 2011a). The following constructs were measured:

- Perceived sidebar accuracy
- Perceived sidebar variety
- Perceived sidebar effectiveness
- Perceived sidebar appeal
- Strength of preference for one segment
- Satisfaction with the website
- Satisfaction with the sidebar

For a full list of all questions in the survey see appendix III. Questions were asked in Dutch. The result of the factor analysis is shown in the table below. Two questions on perceived sidebar accuracy, two questions on perceived sidebar appeal, one question on perceived sidebar variety and two questions on perceived sidebar effectiveness were excluded from the analysis because of bad loadings.

	Perceived Sidebar Accuracy	Perceived Sidebar Appeal	Perceived Sidebar Variety	Perceived Sidebar Effectiveness
I liked the items in the sidebar	0.863			
The items in the sidebar fitted my interests	0.855			
The items in the sidebar were relevant	0.812			
The sidebar is designed well		0.782		
The sidebar does not look nice		-0.603		
The sidebar is nicely build		0.761		
The sidebar looks pleasing		0.855		
The items in the sidebar were very diverse			0.751	
The items in the sidebar were all on the same topic			-0.771	
The items in the sidebar covered may differen product groups			0.827	
Many items in the sidebar differed from other items in the sidebar			0.738	
The items in the sidebar differed from eachother on many aspects			0.768	
The items in the sidebar all looked alike			-0.724	
If I have to decide which article I'm going to read next, I make better choices using the sidebar				0.792
If I have to decide which article I'm going to read next, the sidebar makes me aware of the choices I have				0.787
If I have to decide which article I'm going to read next, I can find better options using the sidebar				0.822
I save time by using the sidebar				0.839
The sidebar does not help me when navigating through the website				-0.616

Table 17: Results factor analysis

The Confirmatory factor analysis (CFA) indicated a fit with RMSEA = 0.057, CFI = 0.98, TLI = 0.97 and a chi-square(129)=1263.38,p< 0.01. According to guidelines set up by Hu and Bentler (1999) this is a decent fit.

After recoding the reversed items of the survey, the factors also showed to be reliable with Cronbach's alphas of 0.837 for perceived sidebar accuracy, 0.795 for perceived sidebar appeal, 0.858 for perceived sidebar variety and 0.854 for perceived sidebar effectiveness. A Cronbach alpha of above 0.7 is acceptable, a Cronbach alpha of above 0.8 is considered good. (George & Mallery, 2003 p. 231) The reliability of the extracted factors can also be demonstrated by the Average Extracted Variance (AVE). The AVE for perceived sidebar accuracy = 0.681, 0.572 for perceived sidebar appeal, 0.580 for perceived sidebar variance and 0.602 for perceived sidebar efficiency. An AVE of above 0.5 is required (Anderson, Babin, Black & Hair, 2010).

Several hypotheses were defined about the effect of the manipulations on the user experience. It was hypothesized that when showing a user content targeted towards the segment of that user, the perceived accuracy of the sidebar should go up (H7) and the perceived variety of the sidebar should go down (H8). As a result of this increase in perceived accuracy and the decrease in perceived variety, the perceived sidebar effectiveness should go up (H10) and the perceived sidebar appeal should go up (H11). As a result of these changes in the perceived user experience, users in congruent conditions are also expected to behave differently because of this change. It is hypothesized that due to the change in user experience users will visit more pages (H11), see a larger variety of product groups (H12), make less switches between segments (H13), make less visits to the main page (H14) and use the sidebar more often (H15).

4.2.1 Structural Equation Model (SEM)

A structural equation model was used to answer the hypotheses about the subjective experiences of the participants in the experiment. The factors from the CFA were modeled together with the behavioral outcomes as discussed in chapter 4.1. The resulting model as shown in figure 31 had a good fit: CFI=0.980, TLI=0.978, RMSEA=0.038 and a chi-square(290)= 44929.702, $p < 0.001$. An arrow indicates a directional relationship between two variables, accompanied by the regression coefficient, the standard error (in parentheses) and the significance of the relation. For clarity reasons questionnaire items are not added to the model, neither are non-significant relations.

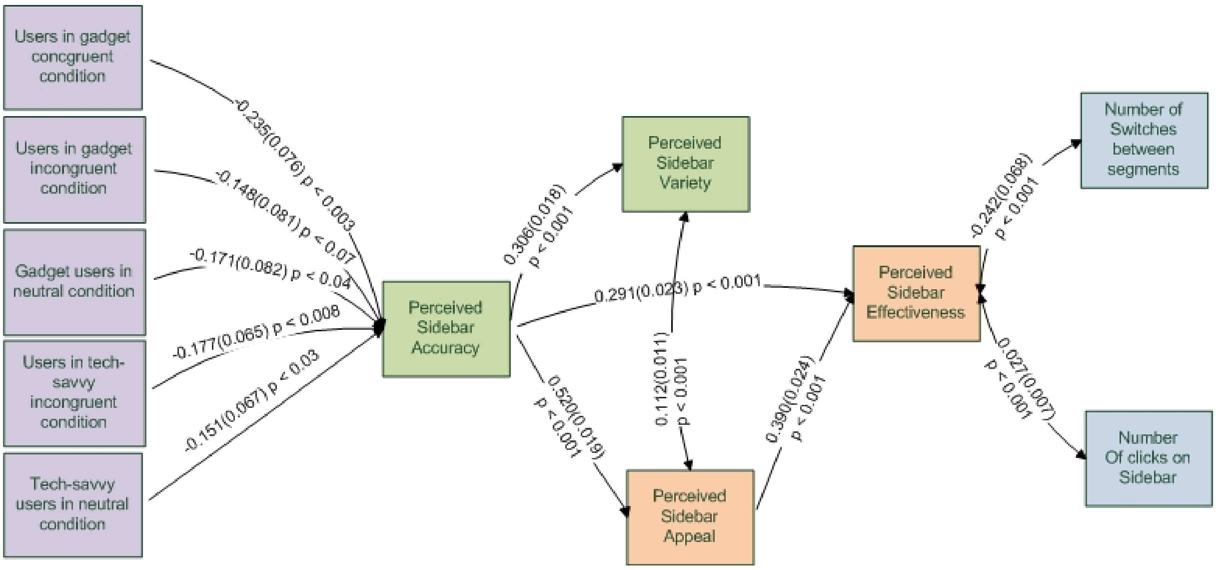

Figure 31: Structural Equation Model

All conditions in figure 31 are compared against tech-savvy users in a congruent condition. As figure 31 shows, all other conditions experienced the sidebar to be significantly less accurate compared to tech-savvy users in a congruent condition. There were no significant differences between the other conditions.

Users in the tech-savvy congruent condition found the website also more varied as a result of the increase in accuracy. This relation was not expected. As a result of the increase in accuracy, they also found the sidebar more appealing and more effective. Part of the variance in perceived effectiveness can be explained by the increase in perceived sidebar appeal. There was also a significant relationship between perceived sidebar appeal and perceived sidebar variety. Last, there is a significant relationship between the perceived sidebar effectiveness and two behavioral measures:

the amount of switches users made and the usage of the sidebar. The satisfaction questions about the website and sidebar did not improve the model, neither did the strength of preference users had.

The results showed that several hypotheses were confirmed. Indeed we found that users in tech-savvy congruent conditions find the sidebar more accurate. Therefore hypothesis seven is true at least for tech-savvy users. We also hypothesized a direct effect between congruent users and variety (H8). This relationship was not found. Perceived sidebar effectiveness was hypothesized to increase as a result of an increase in perceived sidebar accuracy and a decrease in perceived sidebar variety (H9). The perceived sidebar effectiveness did increase because of an increase in accuracy, but not because of a difference in perceived variety. Instead, an increase in perceived appeal also led to an increase in perceived effectiveness. This was not hypothesized. Hypothesis 10 stated that perceived appeal should go up as a result of a decrease in perceived sidebar variety and an increase in perceived accuracy. The relationship between appeal and accuracy was found, but the increase in variety also led to an increase in perceived appeal, or the other way around. Due to the change in behavior users made less switches and used the sidebar more often. Therefore, evidence in support for hypothesis 13 and 15 were found. No evidence was found for hypothesis 11, 12 and 14. These changes in behavior are in line with the behavioral results as described in chapter 4.1.

These results are also partly in line with research done by Knijnenburg and his colleagues (2011a). We found indeed a relationship between the objective system aspects (the different conditions) on one of the tested subjective system aspects (perceived sidebar accuracy). However, no direct effect between the conditions and variety was found, even though perceived sidebar variety is a subjective system aspect as well. The subjective system aspects should, according to Knijnenburg et al. (2011a), relate directly to the experience. Again, this was partly found. Variety only had an effect on one of the experience measures in the experiment. No relation was found between perceived sidebar variety and perceived sidebar effectiveness. As expected, there was a direct relation between accuracy and both sidebar effectiveness and sidebar appeal. Last, the change in experience should lead to a change in behavior according to Knijnenburg et al. (2011a). Again, these results are only partly found. While there was a relation between the experienced effectiveness of the sidebar and the number of switches between segments and the use of the sidebar, there was no relation between the perceived appeal of the sidebar and any of the behavioral measures.

5. Discussion

In a study by Yan and his colleagues in 2009 it was found that click-through-rates on advertisements can rise with 670% when those advertisements are directly targeted towards specific groups of users. Hauser et al. did a study in 2008 where purchase intentions went up with 20% when a telecom website changed its appearance based on the user that was browsing through the website. None of these studies however investigated the effect on user experience when targeting content towards users. Therefore, the main goal of this project was to study the effect of directly and real-time targeting of content to specific user segments on both behavior and user experience.

This project was done in collaboration with Hardware.Info, a large online IT-platform. Using the server logs of one month evidence was found that Hardware.Info has two distinct types of users. On the one hand tech-savvy users especially interested in for example motherboards or processors and on the other hand gadget users who are especially interested in tablets or mobile phones. Using a relatively simple categorization method roughly 85% of all sessions with at least 10 pages visited could be categorized in one of those segments. A prediction algorithm was able to predict many sessions correctly even with the information of only the first five pages visited.

For two weeks the Hardware.Info sidebar would change based on the user using the website. Gadget users in a congruent condition would see only gadget items presented in the sidebar. In an incongruent condition they would see only tech-savvy items and finally there was a neutral condition where those gadget users would see both gadget and tech-savvy items presented in the sidebar. The same kind of conditions also existed for tech-savvy users.

Several hypotheses were defined about the behavior of users. It was expected, based on previous research, that showing a user content targeted towards the segment of that user would result in a change of behavior by that user. It was also hypothesized that for gadget users it would be worse to get only tech-savvy items presented in the sidebar compared to tech-savvy users who would only get gadget items presented in the sidebar.

There were also several hypotheses defined about the user experience. It was expected that the user experience would change when showing a user content targeted towards the segment of that user.

In this chapter the main findings will be summarized along with the limitation and implications of this study.

5.1 Main findings & Implications

Based on the pre-experimental data analysis a difference in behavior between gadget and tech-savvy was expected. The results of the experiment show that this was indeed the case. Gadget users made less clicks on the sidebar, visited less pages, visited the main page less, and made less sessions compared to tech-savvy users. This is a confirmation of the pre-experimental data analysis. Another confirmation was found by analyzing the incentive users could win by completing the survey. Users could win a solid state drive and a mobile phone. The SSD was a typical tech-savvy prize, the mobile phone a typical gadget prize. Users categorized as gadget chose the mobile phone significantly more often compared to tech-savvy users. This demonstrated once again that indeed the distinction made between gadget users and tech-savvy users is valid and the difference between the groups is significant.

In terms of behavior, tech-savvy users in congruent conditions made less switches between tech-savvy and gadget content compared to tech-savvy users in the incongruent condition. In fact, gadget users in a congruent condition switched more between tech-savvy and gadget content compared to gadget users in a incongruent condition. No reasonable explanation could be given for this kind of behavior.

Tech-savvy users in the congruent condition used the sidebar more often compared to tech-savvy users in the incongruent condition. This is a very important finding, it means the manipulation was successful and worked at least for tech-savvy users. For gadget users it apparently did not matter what kind of content the sidebar was showing.

In terms of user experience, it was found that tech-savvy users in congruent conditions indeed found the sidebar most accurate. Users who found the sidebar accurate also found the sidebar more varied. This is interesting because this is the opposite of what was expected. The manipulation made the sidebar actually less varied (only items belonging to one segment) but users find the sidebar more varied instead. Thus a higher perceived variety is achieved by only showing items of one particular segment. Users in who found the sidebar more accurate also found the sidebar more appealing. The effect of perceived sidebar accuracy on perceived sidebar appeal was expected given earlier research. The effect of the congruent manipulation on perceived accuracy and appeal of the sidebar led to an increase in perceived sidebar effectiveness. Users who found the sidebar more efficient then also behaved differently. They made fewer switches between segments and used the sidebar more often. This is in line with the behavioral results where it was also found that users in the tech-savvy congruent condition switched less and used the sidebar more often.

Another interesting finding is that even though users categorized as gadget almost exclusively visit pages categorized as gadget, a lot of them still stated they are very interested in tech-savvy content, even though their behavior does not reflect that. The same results can be found for tech-savvy users. A large group of tech-savvy users stated they are also very much interested in gadget content.

Apparently a large group of users think they are interested in both tech-savvy and gadget content, even though they do not behave like it.

5.2 Limitations & Future research

The first limitation to this study is the prediction model. As described in chapter two, only about 70 percent of all users are categorized correctly. A prediction model can never be perfect, but an improvement in that model would perhaps make results clearer.

The clearest limitation however in this study was the manipulation itself. Because the sidebar on Hardware.Info was not a prominent feature and because the manipulation itself was very subtle users might not even have noticed that something had changed. A larger manipulation that would be more salient to the user might have resulted in much clearer results. A replication of this study with a larger manipulation is however difficult: websites with the size and popularity of Hardware.Info will not likely agree to a live experiment that would introduce large design changes that might potentially disrupt the behavior of many of its users. Future research should try to find a way to experimentally test what the effect of a larger manipulation would be on both behavior and user experience.

5.3 Conclusions

There is much evidence presented in this study that showed that users on the Hardware.Info website can indeed be divided up into two segments.

A lot of evidence was found that tech-savvy congruent users found the sidebar significantly more accurate, therefore more appealing, more varied and more effective. Because of this increase in effectiveness users also behaved differently. They made less switches between segments and used the sidebar more often. Even though these effects on user experience were only found for tech-savvy users, this study still provided evidence that the change in behavior, as a result of directly targeting content towards a user, can at least partly be explained by a change in user experience.

There were also direct effects between tech-savvy congruent users and the amount of switches they made and the use of the sidebar. It was shown that tech-savvy users indeed use the sidebar more often when that sidebar is showing content targeted towards the segment of a certain user. This direct change in behavior is a replication of earlier studies done in the field of adaptive hypermedia.

References

Alkharusi, H. (2012). Categorical Variables in Regression Analysis: A Comparison of Dummy and Effect Coding, *International Journal of Education, 4(2)*, 202-210.

Bajraktarevic, N., Hall, W., & Fullick, P. (2003). Incorporating Learning Styles in Hypermedia Environment: Empirical Evaluation. In *Proc. of AH2003: Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, Budapest, Hungary*, 41-52.

Borges, J., & Levene, M. (2007). Evaluating Variable-Length Markov Chain Models for Analysis of User Web Navigation Sessions. *IEEE Transactions on Knowledge and Data Engineering, 19(4)*, 441-452.

Brusilovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. *User Modeling and User Adapted Interaction , 6*, 87-129.

Brusilovsky, P. and Pesin, L. (1998) Adaptive navigation support in educational hypermedia: An evaluation of the ISIS-Tutor. *Journal of Computing and Information Technology , 6 (1)*, 27-38.

Cameron, A. C. and Trivedi, P. K. 1998. *Regression Analysis of Count Data*. New York: Cambridge Press.

Centraal Bureau voor de Statistiek (CBS). (2012). *ICT, Kennis en Economie.*

De Bra, P., Aerts, A., Berden, B., de Lange, B., Rousseau, B., Santic, T., Smits, D. & Stash, N. (2003) AHA! The Adaptive Hypermedia Architecture. *Proceedings of the fourteenth ACM conference on Hypertext and Hypermedia*, 81-8

Dembczynski, K., Kotlowski, W., & Sydow, M. (2007). Effective prediction of web user behaviour with user-level models. *Fundamena Informaticae, 89*, 189-206.

Fan, L. (2004). Adaptation and Personalization in Web-based Learning Support Systems. *WSS04*, 60-66.

Frias-Martinez, E., Chen, S., & Liu, X. (2007). Automatic Cognitive Style Identification of Digital Library Users for Personalization. *Journal of the American Society for Information Science and Technology, 58(2)*, 237-251.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.

Hair, J., Black, W., Babin, B., and Anderson, R. (2010). *Multivariate data analysis (7th ed.).* Upper Saddle River, NJ, USA: Prentice-Hall, inc.

Hassan, M.T., Junejo, K.N., & Karim, A. (2009). Learning and Predicting Key Web Navigation Patterns Using Bayesian Models. *Proceedings of International Conference on Computational Science and Its Applications,* 877-887.

Hassenzahl, M., & Tractinsky, N. (2006). User Experience - a research agenda. *Behavior & Information Technology, 25(2)*, 91-97

Hauser, J., Urban, G., Liberali, G., & Braun, M. (2008). Website Morphing. *Marketing Science, 28*, 202-223.

Herlocker, J.L., & Konstan, J.A., & Riedl, J. (2000). Explaining collaborative filtering recommendations. *Proc.CSCW2000*, 241–250.

Knijnenburg, B.P., & Willemsen, M.C., & Kobsa, A. (2011a). A pragmatic approach to support the user-centric evaluation of recommender systems. *5th ACM Conference on Recommender Systems, Chicago, IL.*

Lu, L., Dunham, M., & Meng, Y. (2005). Discovery of significant usage patterns from clusters of clickstream data. *Proceedings of Webkdd Workshop Notes,* 68-79.

Ma, H., & Yang, H., & Lyu, M.R., & King, I. (2008). SoRec: Social recommendation using probabilistic matrix factorization. *Proc. of CIKM '08*, 931–940.

Mahlke, S. (2005). Understanding users' experience of interaction. *Proceedings of EACR*, 251–254.

Van Schaik, P. & Ling, J. (2008). Modelling user experience with web sites: Usability, hedonic value, beauty and goodness. *Interacting with Computers, 20(3)*, 419-432.

Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., & Chen, Z. (2009). How much can behavioral targeting help online advertising? *WWW*, 261-270.

Appendix I - Categorization product groups

Name Product group	Category
Moederborden	Tech-savvy
Optische drives	Tech-savvy
Processors	Tech-savvy
Harddisks/SSD's	Tech-savvy
Videokaarten	Tech-savvy
3D-chips	Tech-savvy
Behuizingen	Tech-savvy
Monitoren	Tech-savvy
Routers	Tech-savvy
PC speakersets	Tech-savvy
Barebones	Tech-savvy
Printers/All-in-ones	Gadget
Laptops/Tablets	Gadget
CPU-koelers	Tech-savvy
Geheugenmodules	Tech-savvy
Voedingen	Tech-savvy
Storagecontrollers	Tech-savvy
Casefans	Tech-savvy
Servers	Tech-savvy
NAS devices	Tech-savvy
Systemen	Gadget
Muizen	Gadget
Geluidskaarten	Tech-savvy
Muismatten	Gadget
Webcams	Gadget
Toetsenborden	Gadget
USB-flashdrives	Tech-savvy
Geheugenkaartjes	Gadget
Videokaartkoelers	Tech-savvy
TV-kaarten	Tech-savvy
Netwerkadapters	Tech-savvy
Functionpanels	Tech-savvy
Gamecontrollers	Gadget
Optische media	Gadget
Inktcartridges	Gadget
Projectoren	Gadget
Scanners	Gadget
DVD-/Blu-ray-/mediaspe	Gadget
Powerline adapters	Tech-savvy
Headsets	Gadget
Microfoons	Gadget
Portable mediaspelers	Gadget
Externe harddisks	Gadget
Digitale camera's	Gadget
Koelpasta	Tech-savvy
Geheugenkoelers	Tech-savvy
Chipsetkoelers	Tech-savvy
Harddiskkoelers	Tech-savvy
Besturingssystemen	Tech-savvy
Cardreaders	Gadget
Digitale fotolijstjes	Gadget
Televisies	Gadget
Kastverlichting	Tech-savvy
Fanguards	Tech-savvy
UPS'en	Tech-savvy
Laptoptassen	Gadget
USB-hubs	Tech-savvy
Tekentabletten	Gadget

Name Product group	Category
Videocamera's	Gadget
Kabels	Tech-savvy
Home cinema sets	Gadget
Switches	Tech-savvy
TV-ontvangers/Decoders	Gadget
Laptopstands/koelers	Gadget
Versterkers/Receivers	Gadget
Access points	Tech-savvy
Tonercartridges	Tech-savvy
KVM-switches	Tech-savvy
Afstandsbedieningen	Gadget
USB/Firewire-controllers	Tech-savvy
Bluetooth adapters	Tech-savvy
Laptopaccu's	Gadget
Laptopadapters	Gadget
Home automation	Gadget
Anti-virus software	Tech-savvy
Office software	Gadget
Docking stations/Port rep	Tech-savvy
Mobiele telefoons	Gadget
Monitor/TV-steunen	Tech-savvy
Navigatiesystemen	Gadget
Reinigingsproducten	Gadget
Printpapier	Gadget
Spanningbeveiligers	Tech-savvy
Netwerkantennes	Tech-savvy
Printservers	Tech-savvy
Netwerkcamera's	Tech-savvy
Consoles	Gadget
Console-accessoires	Gadget
Telefoons	Gadget
Development software	Tech-savvy
E-readers	Gadget
Cameralenzen	Gadget
Camerafilters	Gadget
Camera-accessoires	Gadget
Cameratassen	Gadget
Cameraflitsers	Gadget
Batterijen/laders	Gadget
Statieven	Gadget
Behuizing accessoires	Tech-savvy
Testers	Tech-savvy
Projectieschermen	Gadget
iPod Docks/Speakers	Gadget
Koeling accessoires	Tech-savvy
Kabel connectoren	Tech-savvy
Games	Gadget
Voice recorders	Gadget
3D-brillen	Gadget
Speakers	Gadget
Image tanks	Tech-savvy
Calibratieproducten	Tech-savvy
Multimedia software	Gadget
Diensten	Gadget
Verrekijkers	Gadget
Tablet docks	Gadget
Tabletcovers/tassen	Gadget

Appendix II - Results prediction models

Model with only the amount of tech-savvy and gadget pages within the first five pages visited within a session.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	73.99%	19.09%	6.93%
	Tech-savvy session	4.32%	88.82%	6.86%
	Not categorized session	19.94%	53.08%	26.98%

Model with the amount of tech-savvy and gadget pages within the first five pages visited and the type of the first page the user visited a session.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	72.18%	20.41%	7.41%
	Tech-savvy session	4.24%	88.91%	6.85%
	Not categorized session	16.03%	38.49%	45.48%

Model with the segment of the first page visited, the segment of the second page visited, the segment of the third page visited, the segment of the fourth page visited and the segment of the fifth page visited.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	64.24%	32.91%	2.85%
	Tech-savvy session	6.37%	91.00%	2.63%
	Not categorized session	18.51%	57.18%	24.31%

Model with the amount of tech-savvy and gadget pages within the first five pages visited and the segment of the first page visited, the segment of the second page visited, the segment of the third page visited, the segment of the fourth page visited and the segment of the fifth page visited.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	72.62%	21.56%	5.82%
	Tech-savvy session	4.08%	91.67%	4.25%
	Not categorized session	16.43%	54.22%	29.35%

Model with the amount of tech-savvy and gadget pages within the first five pages visited, the number of unique segments within the first five pages visited, the number of unique pagetypes within the first five pages visited and the number of switches between segments within the first five pages visited.

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	73.37%	19.30%	7.34%
	Tech-savvy session	4.27%	89.24%	6.49%
	Not categorized session	18.64%	54.16%	27.20%

Model with the amount of tech-savvy and gadget pages within the first five pages visited, the number of unique segments within the first five pages visited and the number of switches between segments within the first five pages visited

		Predicted		
		Gadget session	Tech-savvy session	Not categorized session
Categorized	Gadget session	73.90%	19.29%	6.81%
	Tech-savvy session	4.31%	89.30%	6.39%
	Not categorized session	19.31%	54.39%	26.30%

And last, the model with the amount of tech-savvy and gadget pages within the first five pages visited, the number of product groups seen in the first five pages visited, and the type of the first page the user visited a session. This model was eventually used in the experiment.

		Predicted		
		Gadget user	Tech-savvy user	Not categorized user
Categorized	Gadget user	70.94%	19.51%	9.56%
	Tech-savvy user	3.53%	87.96%	8.51%
	Not categorized user	12.48%	35.06%	52.46%

Appendix III - Questionnaire

General and demographic questions

Dutch wording	Translation	Scale
Wat is jouw leeftijd?	What is your age?	Open
Wat is jouw geslacht?	What is your sex?	Male/Female
Wat is jouw burgelijke staat?	What is your marital status	4 options
Wat is de hoogste opleiding die je genoten hebt?	What is your highest degree?	5 options
Was dit de eerste keer dat je Hardware.Info bezocht?	Was this your first visit to Hardware.info?	Yes/no
Ik ben heel erg geinteresseerd in processoren, moederborden, videokaarten etc	I'm very much interested in processors, motherboards, video cards etc.	5-point Likert scale
Ik ben heel erg geinteresseerd in tablets, mobiele telefoons, laptops etc.	I'm very much interested in tablets, mobile phones, laptops etc.	5-point Likert scale
Ik weet waar ik naar op zoek ben als ik de Hardware.Info website bezoek	I know what I'm looking for when visiting Hardware.Info	5-point Likert scale

Perceived sidebar accuracy

Dutch wording	Translation	Scale
De items in de zijbalk waren niet interessant	The items in the sidebar were not interesting	5-point Likert
De items in de zijbalk bevielen me	I liked the items in the sidebar	5-point Likert
De items in de zijbalk pasten bij mijn voorkeur	The items in the sidebar fitted my interests	5-point Likert
De items in de zijbalk waren relevant	The items in the sidebar were relevant	5-point Likert
De items in de zijbalk waren niet goed gekozen	The items in the sidebar were not chosen well	5-point Likert

Perceived sidebar appeal

Dutch wording	Translation	Scale
De zijbalk is goed ontworpen	The sidebar is designed well	5-point Likert
Het gebruik van de zijbalk is een aangename ervaring	Using the sidebar is a pleasant experience	5-point Likert
De zijbak ziet er niet mooi uit	The sidebar does not look nice	5-point Likert
De zijbalk is erg knap gemaakt	The sidebar is nicely build	5-point Likert
De zijbalk ziet er plezierig uit	The sidebar looks pleasing	5-point Likert
De zijbalk trekt mijn aandacht	The sidebar attracts my attention	5-point Likert

Perceived sidebar variety

Dutch wording	Translation	Scale
De items in de zijbalk waren heel divers	The items in the sidebar were very diverse	5-point Likert
De items in de zijbalk gingen allemaal over hetzelfde	The items in the sidebar were all on the same topic	5-point Likert
De items in de zijbalk gingen over veel verschillende productgroepen	The items in the sidebar covered may differen product groups	5-point Likert
Veel items in de zijbalk verschilden van andere items in de zijbalk	Many items in the sidebar differed from other items in the sidebar	5-point Likert
De items in de zijbalk verschilden van elkaar op vele aspecten	The items in the sidebar differed from eachother on many aspects	5-point Likert
De items in de zijbalk leken allemaal op elkaar	The items in the sidebar all looked alike	5-point Likert
De zijbalk presenteerde artikelen die ik nog niet eerder gelezen heb	The sidebar presented items that I have not read before	5-point Likert

Perceived sidebar effectiveness

Dutch wording	Translation	Scale
Als ik moet gaan besluiten wat het volgende artikel is dat ik ga lezen, maak ik betere keuzes met behulp van de zijbalk	If I have to decide which article I'm going to read next, I make better choices using the sidebar	5-point Likert
Als ik moet gaan besluiten wat het volgende artikel is dat ik ga lezen, beperkt de zijbalk mijn keuzes	If I have to decide which article I'm going to read next, the sidebar limits my options	5-point Likert
Als ik moet gaan besluiten wat het volgende artikel is dat ik ga lezen, maakt de zijbalk mij bewust van de keuzes die ik heb	If I have to decide which article I'm going to read next, the sidebar makes me aware of the choices I have	5-point Likert
Als ik moet gaan besluiten wat het volgende artikel is wat ik ga lezen, kan ik betere opties vinden met behulp van de zijbalk	If I have to decide which article I'm going to read next, I can find better options using the sidebar	5-point Likert
Ik bespaar tijd door gebruik te maken van de zijbalk	I save time by using the sidebar	5-point Likert
De zijbalk helpt mij in het vinden van nieuwe en interessante artikelen	The sidebar helps me in finding new and interesting articles	5-point Likert
Bij het navigeren door de website heeft de zijbalk geen echt voordeel voor mij	The sidebar does not help me when navigating through the website	5-point Likert

Satisfaction with the website

Dutch wording	Translation	Scale
De website is verschrikkelijk	The website is terrible	5-point Likert
De website is moeilijk	The website is difficult	5-point Likert
De website geeft voldoening	The website satisfices me	5-point Likert
De website is saai	The website is boring	5-point Likert
De website is star	The website is rigid	5-point Likert

Satisfaction with the sidebar

Dutch wording	Translation	Scale
De zijbalk is verschrikkelijk	The sidebar is terrible	5-point Likert
De zijbalk is moeilijk	The sidebar is difficult	5-point Likert
De zijbalk geeft voldoening	The sidebar satisfices me	5-point Likert
De zijbalk is saai	The sidebar is boring	5-point Likert
De zijbalk is star	The sidebar is rigid	5-point Likert