The dimension of age and gender as user model demographic factors for automatic personalization in e-commerce sites

Daniel Fernandez-Lanvin; Javier De Andres-Suarez; Martin Gonzalez-Rodriguez; Beatriz Pariente-Martinez

Abstract

Personalization in e-commerce increases sales by improving customer perception of site quality. However, some demographic data about customers (crucial for the success of the personalization process) not always can be obtained explicitly, as is the case of anonymous web site visitors.

The paper describes a user study focused on determining whether it would be possible to categorize the age and gender of individual visitors of a web site through the automatic analysis of their behavior. Three tasks commonly found in e-commerce sites (*Point & Click*, *Drag & Drop* and *Item Selection*) were tested by 592 volunteers and their performance was analyzed using different statistical methods. The study found consistencies in the execution times of individuals across the different tasks and revealed that age and gender are sufficiently determining factors to support an automatic profiling. Results also showed that relevant information about gender and age can be extracted separately through the individual analysis of each one of the mentioned interaction tasks

Keywords — Personalization, User Model, GOMS, Fitts' law, Hicks-Hyman's law, Salthouse' regularities.

1. Introduction

The success of online marketing is determined, –among other factors— by the level of *personalization* of the e-commerce sites. That is, the process of making a unique user experience for each customer. Personalization is a dominant business model in online marketing strategies (Chen & Hsieh 2012) and is used to establish relationships between customers and sellers (Shen 2014). Its capability to provide recommendations to the customers is acknowledged to be an important feature of online shopping (Choi et al. 2011) as it enhances customer retention and increases sales (Woo & Shirmohammadi 2008).

Addressing similarities and differences among consumers is critical as differences in demographic factors may be associated to different tastes and therefore to different purchasing patterns (Xu 2006).

There is strong empirical evidence showing that differences caused by gender and age, influence online shopping preferences (Haque et al. 2006; Lee 2009). Several authors pointed these two variables as key elements in the personalization process (Kim et al. 2007; Weiser 2000; Freudenthal 2001; Cheong et al. 2013; Agudo et al. 2010).

Personalization requires to collect relevant information about users and this information has a great relevance in the success of e-commerce (Sebora et al. 2008; Turban et al. 2008; Ardissono & Goy 2000; Alpert et al. 2003). However, data gathering is not a trivial issue. It can be explicitly made (e.g. getting this information through registration forms) or implicitly (e.g. monitoring customers' purchase patterns) (Woo & Shirmohammadi 2008).

Although most of the demographic factors are explicitly collected through the registration process (Zhang & Ghorbani 2007; Yang & Claramunt 2007; Woo & Shirmohammadi 2008), this approach may result into biases and/or outdated data (Woo & Shirmohammadi 2008). Users may withhold information due to privacy, social or cultural issues (Zhang & Ghorbani 2007). So for example, underage users may avoid controls in adult-oriented sites or social networks just providing fake data (Strom et al. 2012). Even more important, the outcome of the data gathered explicitly may be limited given that most of the potential customers in online sites are anonymous and/or first visit users. The use of an implicit data gathering approach may help to surpass these limitations.

Building implicit data gathering systems able to estimate the age and/or gender of their users requires the prior identification of the interaction factors that make a user to be unique. This work explores the execution time required to perform basic interaction tasks as a candidate factor to build such kind of systems.

Although previous research efforts suggest an independent influence of both age and gender on the execution time, ((Hill et al. 2011), (De Andrés-Suárez et al. 2015), (Rohr 2006; Beckwith & Burnett 2007)), elegant literature lacks an evaluation of their combined effect in e-commerce applications. Therefore, this work attempts to assess the degree of association between gender and age on execution time conducting a joint evaluation of the effect of age and gender on interaction. It also introduces a combined analysis on how other variables such as the user's laterality (left handed or right handed) or the user's prior experience in the use of computers might influence the time required by the users to complete basic interaction tasks.

The study analyzes the performance of 592 volunteers executing three usability design patterns commonly found in the design of e-commerce interactive systems like Amazon, DeviantArt, Alibaba, etc. These patterns are: (i) *Point & Click*, (ii) *Drag & Drop* and (iii) *Item Selection*. Although this study is focused on the analysis of the users' performance in the use of e-commerce usability patterns, its findings may impact other domains based on these patterns too (e.g. education and entertainment).

The goals of the research are to determine whether the influence of age and gender on the execution time is significant enough to infer its value through behavior analysis and to analyze which of the interaction tasks mentioned before would be the most appropriate to build such kind of personalization systems. If it is possible to infer the user's age and/or gender through the quick inspection of the performance measured in the execution of basic interaction tasks, it would not only be possible to adapt marketing messages to a specific age and gender range, but also the development of tools targeted to prevent certain crimes, such as pedophilia, or illegal access to web sites. Adults pretending to be children in social networks could be detected through the analysis of their interactions with the user interface. A similar approach may be used to detect children or teenagers accessing adult web sites.

The remainder of the paper is structured as follows. Section two discusses prior literature about studies based on the influence of age and gender on user's performance. Section three designs the hypotheses, describes the empirical study used to test them and discusses the statistical methods employed. The fourth section is devoted to the presentation and discussion of the results. Finally, the main conclusions, practical implications and future research are presented.

2. Related Work

Determining users' age and gender by observing their interactions with an e-commerce web site to update the site's user model dynamically, and therefore improving its marketing capabilities, is the overall goal of this work.

User modeling is the process of constructing user models (Zhang & Ghorbani 2007). A user model is an explicit representation of the properties of individual users or user classes. It allows the adaptation of the system to the user needs and preferences (Liu et al. 2008). This process involves both static and dynamic user information. Static user information refers to basic characteristics (e.g. demographics) explicitly presented by the user during a registration procedure. On the other hand, dynamic user information is collected by observing user's behavior and it is recorded in log files or in a list of objects visited by the web user (Yang & Claramunt 2007).

Automated user profiling has been studied in previous works with different approaches and results. Most of these works deal with semantic information gathered from users' interaction. For instance, Woo & Shirmohammadi (Woo & Shirmohammadi 2008) proposed an automatic user personality categorization model based on their digital personality. The authors collect information through the observation of user interacting with products. Yang & Claramunt (Yang & Claramunt 2007) proposed a semantic user model that uses both static and dynamic user information to predict user features relevant for a specific application domain. Such et al. (Such et al. 2013) analyzed the automated user profiling techniques and proposed an approach to control buyer profiling. However, their goal is to prevent users to be automatically identified, just the opposite to the motivation of this work. Fijalkowsky (Fijałkowski 2011) proposes an e-commerce web system that collects data obtained from social network profiles and uses it to provide purchase recommendations to its users.

All these works are focused on user behavior information at the semantic level, under the assumption that demographic information will be explicitly collected and relying in user's collaboration. Ghazarian and Noorhosseini (Ghazarian & Noorhosseini 2010) face the same problem from a lower abstraction level, using machine learning algorithms to adapt user interfaces to the needs of user groups with different levels of skills detected through the analysis of mouse motions. Garg et al (Garg et al. 2006) also tries to identify users depending on interaction behavior like mouse movements and clicks, typing speed and system background processes. The authors extract relevant mouse related features such as average distance, speed, angles of movement and number of clicks during a session, and then they utilize Support Vector Machines (SVM) to learn the user specific feature sets. They conclude that this information is relevant enough to identify and/or differentiate users behind different aliases, but they use it only to prevent masquerading attacks to web sites.

None of these works are focused on the identification of the demographic factors, so their results do not analyze the feasibility of this approach. On the contrary, most of them assume that relevant data should be collected through the registration process, with the limitations that this strategy involves.

2.2. Ageing

Ageing negatively impacts the ability to use computers (Fisk et al. 2009; Hill et al. 2011) and is typically defined in Human Computer-Interaction through an emphasis on declines in abilities and associated reductions in performance when using technology (Vines et al. 2015). It produces a poorer motor control and sensory deficits (Dickinson et al. 2007). Related studies show that older people have slower reaction times (Fozard et al. 1994), delayed movements, a decline in motor skills (Walker et al. 1997). Reduced mobility, caused by a loss in muscle strength (Stubbs et al. 1993), produces difficulties in the execution of movements (Walker et al. 1997). This process of losing muscle strength begins in people aged over 40 (Metter et al. 1997).

Other studies revealed that ageing negatively influences the learning strategies used to operate online systems, as perception and cognition declines (Morrell 2001; Xie 2003). Senior users have been found to be slower than young adults when retrieving information (Nap et al. 2005; Freudenthal 2001), performing 3D navigation on desktop systems (Sayers 2004) or browsing the web (Neerincx et al. 2000).

Studies analyzing information search behavior (Tullis 2007; Hill et al. 2011) enforce the importance of ageing. The study of the behavior of expert older adults daily using the web, compared to their younger colleagues, concluded that age is a determining factor (Hill et al. 2011). This work is particularly interesting because it is specifically focused in web interaction. However, the analyzed activities (search behavior and related) require different operators than those involved in mouse motion.

On the other hand, very young users reveal a poor execution time in the development of certain tasks. Basic interaction tasks like *Drag & Drop* are especially difficult for them as keeping the finger pressed down while controlling its trajectory requires a high demand of motor skills (MacKenzie et al. 1991), perception and cognitive skills (Chadwick-Dias et al. 2002; Czaja & Lee 2006). The execution time slows down if it is possible to replace *Drag & Drop* by *Point & Click*, as the operation can be easily resumed from the last pointing task in case of failure (Joiner et al. 1998). Attaching and lifting objects in the real world causes some difficulties to children under 8 years old as these tasks requires subtle hand-eye coordination (Kuhtz-Buschbeck et al. 1998). At that age, the coordinate movements are further determined by cognitive factors rather than motor skills (Agudo et al. 2010). The speed of such coordinate movements evolves up to the age of 12 years (Kuhtz-Buschbeck et al. 1998).

Some authors reported how different interfaces influence the interaction of specific groups of users regarding their age (Carvalho et al. 2015), but no study was found about whether there are significant differences between the time required to execute different –alternative— interaction tasks conducting to the same result (e.g. *Point & Click* as an alternative to *Drag & Drop* to obtain the same result).

If these differences do not exist and the execution time keeps coherence in each basic interaction task, that is, if the time required by each group of users is similar in each task (*Point & Click*, *Drag & Drop* and *Item Selection*) it would not be necessary to analyze the three interaction tasks in the same user interface to detect the users' age. It would be enough to analyze the users' performance in only one of them. However, if those differences exist, it would be necessary to measure and to analyze the users' performance in all the three different proposed interaction tasks to categorize users according with their age.

2.3. Gender

Women process information in different ways than men (Beckwith et al. 2006). Gender-associated differences in decision making, learning, and problem solving can be a determining factor in user's effectiveness (Beckwith 2003; Beckwith & Burnett 2004). Even more, it has been observed that the self-perceptions concerning computer competence as well as the level of ICT-related social interactions is different for boys and girls (Christoph et al. 2014).

It was observed that men's performance in navigating through virtual environments is better than women's when small displays are used. The use of larger displays reduces the gender performance gap since the women's performance improves while the men's performance is not negatively affected (Tan et al. 2003; Czerwinski et al. 2002).

Inkpen (Inkpen 2001) compared *Drag & Drop* tasks as opposed to *Point & Click* in children. Although there was no any significant gender difference in the overall movement time and/or general error rates, there were relevant differences in pickup and drop errors. The girls performed poorly when executing *Drag & Drop* tasks, as opposed to *Point & Click*. There were also performance correlation differences between gender and target size.

Rohr (Rohr 2006) evidenced that gender-specific movement biases emphasize speed for men and accuracy for women. Wahlstrom et al. (Wahlström et al. 2000) observed that when operating the mouse, women worked with greater extension and had a greater range of motion in the wrist when compared to men. This observation could explain Rohr's results regarding speed versus accuracy. They also found gender differences for musculoskeletal load. For most of the measured variables, women worked with higher loads than men. These differences are not limited to the low-level interaction. Collazos et al (Collazos, César; Guerrero, Luis A. ; Llana, Mónica; Oetzel 2016) found significant differences in the way woman and men face collaborative work in computer-mediated communication.

3. Design of the Empirical Study

3.1 Hypothesis

The related previous studies evidence that there are significant differences between the times required by children and adults to execute different basic interactions tasks. However, to date, we found no studies evidencing these differences in adults, something that lead us to conjecture that the performance of one specific adult in these tasks could be correlated. If so, it would mean that the analysis of performance in one of them would be enough to identify adults, simplifying users age classification. On the other hand, even though there are not evidences of differences between genders in adults for these basic interaction tasks, some studies identified some differences between men and women in other activities that could determine the correlation between the performances in these basic interaction tasks. That lead us to wonder whether these correlations could be determined by user's gender. Therefore, we formulate the following hypotheses to be verified/refuted by the empirical study:

(h_1) The execution time of the different tasks increases with the age of the subject under study

(h_2) Women's execution time for the different tasks is longer than men's

(h_3) The execution times of basic interaction tasks (Point & Click, Drag & Drop and Item Selection) are significantly correlated

To assess whether the hypothesis formulated in the prior section hold, the performance of 592 individuals was analyzed in the execution of three basic interaction tasks.

3.2 Object of Study

The tasks analyzed in this study were *Point & Click*, *Drag & Drop* and *Item Selection*. They were selected because of the crucial role they play in the usability patterns behind the design of successful e-commerce sites.

Point & Click is used to move the mouse pointer over an image or over CTA items (*Call To Action*) to click on it. It is commonly used by customers of online shopping sites like Amazon, eBay, ModCloth, Zappos, etc. to retrieve information about appealing products or to include them in the shopping basket.

Drag & Drop is mostly used to collect vast number of items to place them into the shopping basket. This task is commonly found as part of the usability patterns used in art/photo e-commerce sites like UXPin, DeviantArt, etc.

Finally, *Item Selection* is used to browse through small navigation bars or menus to select item categories. It is commonly used is popular sites like Alibaba, Walmart, Asos, Etsy, etc. Users were encouraged to complete these tasks achieving interaction goals in the minimum amount of time.

Their behavior was recorded by data gathering agents that measured the execution time required by everyone to complete every single task proposed he measurement of the user's execution time in the different tests proposed was based on GOMS (Goals, Operators, Methods, and Selection rules). This analysis method was designed to estimate the users' performance when they interact with different interfaces (Card et al. 1983). The method has been successfully used to estimate user performance in many different scenarios including interaction with automobiles (Xiang & Xiaoli 2010), touch screens (Abdulin 2011) and online web sites (Schrepp 2010; Oyewole & Haight 2011), among several others.

GOMS splits complex interaction tasks into low level components called *operators*. These operators include actions like mouse pointing (denoted as P), dragging (D), key typing (K), decision taking (M), etc. The execution of each operator requires a specific amount of time (denoted respectively as T_P, T_D, T_K, T_M, etc.), so GOMS estimates the execution time of complex interaction tasks as the sum of the execution times of the different operators required to complete the tasks. So, for example, the estimated execution time for a *Drag & Drop* interaction task would be $T_P + T_K + T_D + T_K$; that is, the time needed to move the mouse pointer over the movable object (T_P) plus the time required to press the mouse' button once the pointer is over the target (T_K) plus the time used to drag the object to a new position (T_D) plus the time required to release the mouse's button (T_K).

The execution time for each operator (T_P, T_D, T_K, T_M, etc.) is estimated using well-known psychological laws and regularities such as the Fitts' law (T_P and T_D), the Salthouse's regularities (T_K), the Hicks-Hyman's law (T_M), etc.

Fitts' law estimates the time needed to move a pointing object (the users' finger, the mouse pointer, a joystick, etc.) over a target as $a + b \, Log_2 \, (D/S)$. Where 'D' is the distance to the target, 'S' represents the target's size and 'a' and 'b' are user dependent correction factors ((Zhai 2004), (Guiard et al. 2011)). Salthouse's regularities predict the time required by different kind of users (ranging from novices to experts) to type texts of a known length (Salthouse 1984). The Hicks-Hyman's law estimates the time required to take a decision (such as the selection of a menu item) as $a + b \, Log_2(n+1)$ where 'n' represents the number of available options and 'a' and 'b' again are user dependent correction factors ((Rosatti L 2013), (Schneider & Anderson 2011)).

Although these laws help to estimate the execution time required by an average user, they have to be adapted to the specific needs of individual users. That is the case of the correction factors used by the Fitts' law and the Hicks-Hyman's law which have to be obtained through the analysis of performance records previously obtained for specific users. The values for these correction factors rely on the external variables under analysis in this research, as it is the case of the age and gender.

The use of GOMS in this context has two main advantages. First, it helps to structure the study of the different interaction tasks using a common research framework to other similar studies. Second, the experimental measurement of the users' runtime for each specific task, to a high degree of accuracy, facilitates a quick and accurate estimation of the global execution time for e-commerce sites whose user interfaces combine several of these interaction tasks.

Each test was designed to replicate the behavior of a real e-commerce application but hiding the features that might allow the user to identify it, thus avoiding the effect that the familiarity with the real product might have on the measurements to be obtained. Hence, the lexical and semantical levels (related to mouse movement and object recognition and perception) of the user interface of the corresponded e-commerce application were recreated in the most realistic way, while the semantic (iconic representation) and conceptual (final goal of the application) levels of the interface were ignored or hidden to avoid the mention familiarity.

The first test (task 1) was designed to analyze the behavior of users executing the *Point & Click* tasks required to select objects in web documents by moving the mouse pointer across the display to click on links, buttons, scrolling boxes, etc.

The test showed a sequential series of rectangles in different locations across the screen. Participants in the test had to click inside each rectangle to make it disappear before a new one appeared in a different location. The test used fourteen different rectangles distributed in positions that followed a Z pattern layout to keep a fair balance between left-handed and right-handed users. Along the test, the location of the different targets was changed using the horizontal (left to right, right to left) and vertical dimensions (top to bottom, bottom to top).

At the same time, Fitts' law was used to increase the difficulty of each interaction, increasing the distance to the target (D) and reducing its size (S), thus increasing the time required to click on the target by a factor of $\log_2(D/S)$.

To click on the target, users had to use two GOMS operators: P and K (see Figure 1). First, the users moved the mouse over the display to place the pointer over the square using T_P units of time (step 1 in Figure 1). Next, users needed to click pressing the mouse button using a K operator (step 2 in Figure 1). The time estimated by GOMS to complete each *Point & Click* action is therefore $T_P + T_K$. The time required to complete each point and click action ($T_P + T_K$) was recorded (in milliseconds) for each click interaction. The sum of the execution times required to complete the full test was recorded for later statistical analysis.

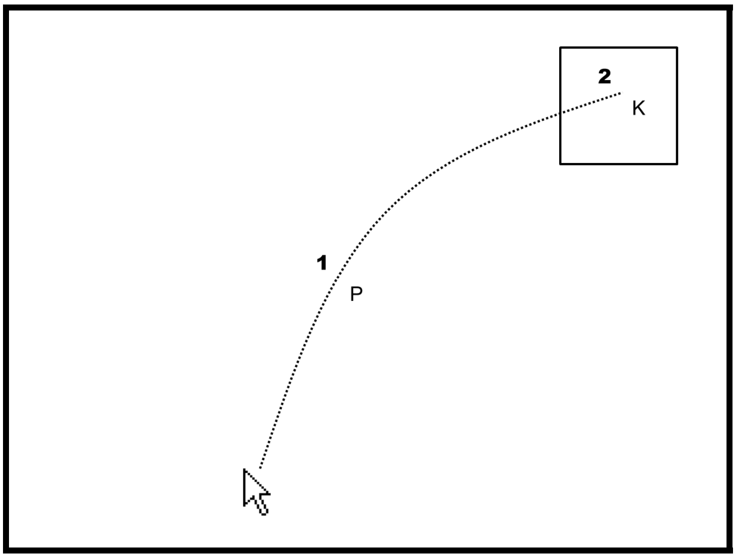

Figure 1. GOMS operators required to complete a *Point & Click* task. Step 1: users move the mouse *Pointing* (P) to the target. Step 2: the user clicks on the target *Key pressing* (K) the mouse button.

The second test (task 2) was designed to measure the time required to complete *Drag & Drop* tasks, commonly used to drag items into the shopping basket in electronic commerce applications.

In this second test users were asked to drag a red rectangle over a second one, which had a size two thirds bigger than the red one. Every time the user completed the task, both rectangles disappeared, and two new rectangles appeared in separate locations of the display. The process was repeated along fourteen interactions. Each time, the rectangles were distributed using a Z shaped layout to keep a fair balance between left-handed and right-handed users. The distance between objects was incremented and its size was reduced in each interaction, using the Fitts' law to increase the time required by the users to complete each interaction.

To drag the first rectangle over the second, the users had to select it first. Therefore, they needed to use the GOMS operators required in a *Point & Click* task. The P operator is required to point to the rectangle (step 1 in Figure 2) and the K operator is needed to select it (step 2 in Figure 2) clicking the mouse button. Next, users had to drag the rectangle using the dragging operator (D) until the first rectangle was over the second one (step 3 in Figure 2) releasing it with a mouse button action (T_K). The time estimated by GOMS to complete each *Drag & Drop* action is therefore $T_P + T_D + 2T_K$. The time required to drag the object (T_D) was recorded (in milliseconds) to be used in the statistical analysis.

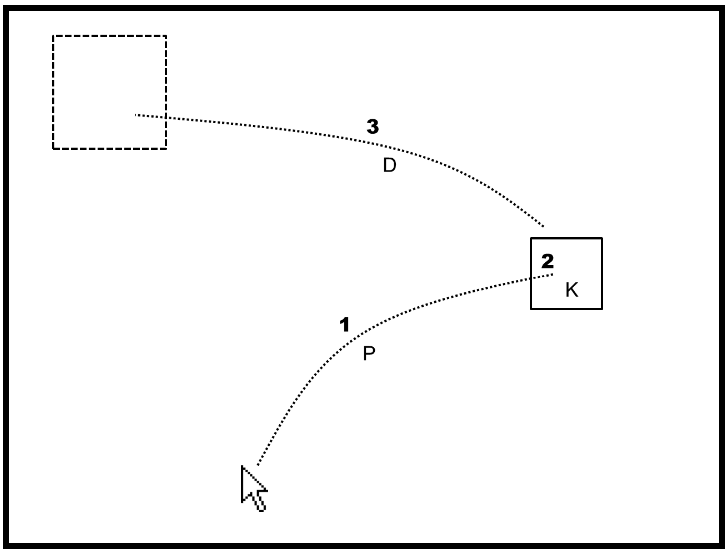

Figure 2. GOMS operators required to complete a *Drag & Drop* task in the test application. Steps 1 and 2 are the same as in the *Point & Click* task described in **Figure 1**. In step 3 users had to drag (D) the small rectangle over the big one.

The third and last test (task 3) was designed to evaluate the user performance in the execution of *Item Selection* tasks, which are used to select items in a user interface (e.g. menus, combo boxes, radio button groups, etc.). In this test, users were asked to select a given color in a popup menu. To achieve this operation participants needed to execute a *Point & Click* task to display the menu items available clicking on the menu's title. Then, users were asked to select a specific menu item whose name was displayed in the screen. Then, participants executed a second *Point & Click* task to click on the menu item corresponding to the asked color. The process was repeated ten times. In each interaction, the menu was placed in a different position using the Z shaped layout described before. Each menu contained five items. Each volunteer had to select each menu item two times across the interactions.

The time required to achieve the first *Point & Click* task (see Figure 3) was denoted by $T_{P1} + T_{K1}$. It represents the Point (P_1) and Key pressing (K_2) operators required to activate the menu. The second runtime was denoted by $T_{P2} + T_{K2}$. Finally, the time needed by the mental operator M to take the decision (selecting which menu item satisfies the search constraints) was denoted by T_M. The resulting execution time predicted by GOMS for the entire test process, denoted by $T_{P1} + T_{K1} + T_M + T_{P2} + T_{K2}$ was recorded to later statistical analysis**.**

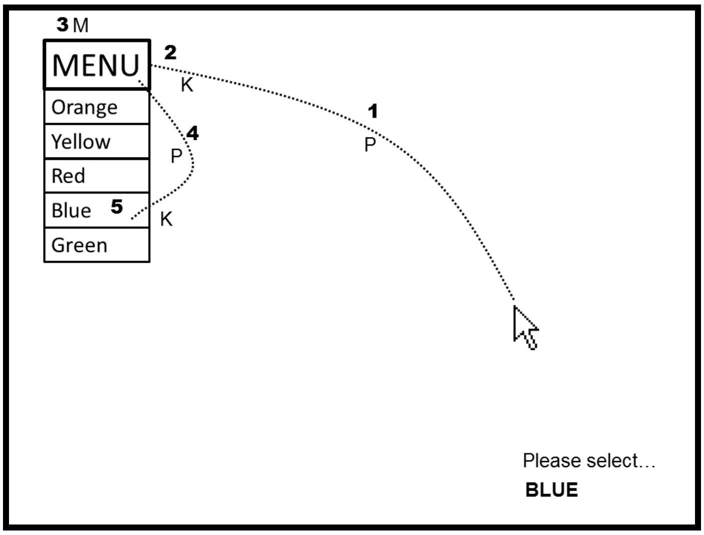

Figure 3. GOMS operators required to complete an *Item Selection* task. Steps 1 and 2 define the *Point & Click* task required to activate the menu clicking on its title. Once the menu items are displayed, a Mental (M) operator is executed (step 3) to select the required color (displayed in the bottom-right corner of the screen). Steps 4 and 5 represent the Pointing (P) and Key pressing (K) GOMS's operators required to complete the *Point & Click* task needed to select the menu item.

3.3 Subjects

GOMS assumes that the volunteers know how to use the web system under evaluation (either because they got some previous training or because they have used the system previously). GOMS also assumes that users will not commit any error during the process. Due to this high degree of expertise, users are supposed to interact as fast as possible. Based on these precepts, GOMS is a reliable tool to estimate the user's effectiveness (execution time) instead of estimating the user's efficiency (success/failure rate).

To meet these strong requirements, the 592 individuals participating in the case study were recruited through Twitter and *Foro Coches* (http://www.forocoches.com/), the most popular general purpose online community in Spain, thereby ensuring that participants were familiar with the basic interaction tasks frequently found in online systems. Therefore, participants could execute *Point & Click*, *Drag & Drop* and *Item Selection* tasks in a so natural way that they did not need to think about the steps needed to complete them.

This approach not only complied with the GOMS requirements but also allowed the participation of a high number of users. The sample used in this study include 592 individuals. It is large when compared with the samples used in the studies described in the *Related Work* section, which were mostly based on samples whose size ranges between 10 and 20 individuals.

This high number allowed the use of multivariate regression analysis to obtain more accurate results when compared with those of prior studies. In addition, it allowed the inclusion of some variables in the model that may bias the results if they are not adequately controlled for (handedness and prior experience with computers).

3.4 Variables of the study

Apart from the variables used to test our first two hypotheses (age and gender) and the execution times of the analyzed tasks we considered some additional variables for the testing of h_1 and h_2.

Specifically, we included handedness and previous user experience with computers. Several studies reported differences regarding movement between left and right handed individuals (Lenhard & Hoffmann 2007; Mieschke et al. 2001; Velay & Benoit-Dubrocard 1999), movement preparation (Bestelmeyer & Carey 2004; Boulinguez et al. 2000; Neely et al. 2005; Helsen et al. 1998), stimulus velocity effect (Rodrigues et al. 2012) and interactions between hand preference and hand performance (Peters & Ivanoff 1999). Besides this, other studies suggest that skill performance and the amount of practice are correlated (Howard 2014) following an exponential law (Heathcote et al. 2000).

These precedents suggest that users' handedness and the users' experience may have a sensible influence on the user behavior and therefore in their execution time. As a result, these two variables (handedness and amount of practice) were incorporated as control variables in the regression models that are explained in section 3.5.

As was noted above, the experiment was completed by 592 participants. As a summary, we indicate in Table 1 the variables used in the study.

Before participating in the tests, users were asked to fulfill a questionnaire to provide information about their age, gender, handedness (tendency to use either the right or the left hand) and experience in the use of computers. This last parameter was provided in terms of the number of weekly hours spent by the users interacting with computers. Some of the users were reluctant to provide their actual age (especially older users). As a result, we were forced to discretize the age value in ranges of 5 years. This way we sacrifice some of the statistical analysis to obtain this parameter from all the users participating in the tests.

Name	**Definition**		
Dependent Variables			
Point & Click	Time (T_P) required to pointing (P) each object during the test (measured in milliseconds)		
Drag & Drop	Time (T_D) required to drag (D) each object during the test (measured in milliseconds)		
Item Selection	Time required selecting each menu item during the test. It was calculated as $T_{P1} + T_{K1} + T_M + T_{P2} + T_{K2}$ (measured in milliseconds)		
Independent Variables			
Age	Age Group	Minimum Age	Maximum Age
	0	0	15
	1	16	20
	2	21	25
	3	26	30
	4	31	35
	5	36	40
	6	41	45
	7	46	50
	8	51	55
	9	56	60
	10	61	65
	11	>= 66	

HoursUse	Weekly number of hours interacting with computers.
Gender	1 Female, 0 Male
Handedness	1 Left handed, 0 Right handed

Table 1. Variables in the study.

3.5. Statistical Methods

First, we computed some descriptive statistics about both the dependent and independent variables. The exam of such data gives us a first idea of the features of the individuals in the sample and their behavior in the experiment.

Second, to test hypotheses h_1 and h_2, we estimated a Linear Regression model for each of the tasks. The regression equations have the following form:

$$Task_i = a_0 + a_1 \times Age + a_2 \times Gender + a_3 \times HoursUse + a_4 \times LeftHanded + \varepsilon_i$$

Where $Task_i$ is the dependent variable in each one of the models, a_0 is the intercept term, a_1 to a_4 are the coefficients of the independent variables in the models and ε_i is the error term.

Regarding these equations, and as prior robustness checks, we tested for multicollinearity and heteroskedasticity. Multicollinearity tests were conducted through the calculation of the Condition Indices (CI) and the Variance Inflation Factors (VIF). To assess whether heteroskedasticity represents a problem we used the Cook and Weisberg test (Cook & Weisberg 1983).

Furthermore, we also conducted some post-estimation additional tests which allow shedding light on specific concerns about whether a) there are extreme values which have an abnormal influence on the results, b) the model is not correctly specified and c) results are sensitive about the browser/operating system used. First, to detect the presence of influential cases we computed Cook's D statistic for each data point in the regressions. Second, and regarding model specification, we tested for the existence of non-linear effects for the age variable (that is, whether middle-age users perform better than both younger and older users). This was done by adding a quadratic term (Age^2) to the equations and reestimating the models. Finally, we also re-estimated the models for different subsamples defined considering the browser used for the test (three subsamples: Chrome, Firefox, IExplorer, as the number of persons using other navigators was not enough to allow regression equation estimation) and the operating system (Windows, Linux and Mac).

Finally, and to know whether the execution time of individuals about one task is related to the performance in the other tasks (h_3) we conducted a correlation analysis. We computed Nonparametric correlation coefficients (Spearman's Rho) to avoid the problems caused by nonnormality of data. To test normality of data we used the Lilliefors test, and in all cases data distributions departed significantly from normality (results not reported due to space limitations). For the calculations of these statistics, as well as for all the other tests and equations indicated above, we used the statistical package STATA 11.

4 Results

4.1. Descriptive statistics

	Mean	Std. Dev.	Minimum	Maximum
Point & Click	16864.77	4294.92	9319	45792
Drag & Drop	32832.77	10615.61	19595	159867
Item Selection	61139.34	14069.62	38351	147630

Table 2. Descriptive statistics for the dependent variables in the study (execution times are measured in milliseconds).

Age	Number of observations
0	2
1	85
2	182
3	145
4	77
5	38
6	25
7	13
8	12
9	9
10	2
11	2
TOTAL	592

Table 3. Frequency distribution for Age.

Gender	Number of observations
0 (Male)	462
1 (Female)	130
TOTAL	592

Table 4. Frequency distribution for Gender.

HoursUse	Number of observations
0	1
1	16
2	58
3	67
4	63
5	33
6	354
TOTAL	592

Table 5. Frequency distribution for HoursUse.

Handedness	**Number of observations**
0 (Right Handed)	524
1 (Left Handed)	68
TOTAL	592

Table 3. Frequency distribution for handedness.

As we might expect the average execution time depends on the complexity of the test. As mentioned previously, some authors (MacKenzie et al. 1991; Chadwick-Dias et al. 2002; Czaja & Lee 2006) reported a higher level of complexity in the execution of *Drag & Drop* tasks when compared with *Point & Click*. Thereby, Table 2 shows that *Drag & Drop* tasks required a higher amount of time that Point & Click tasks.

Furthermore, the runtime of menu selection tasks is higher than that of the other two. This result is consistent with predictions provided by GOMS analysis studied before. Notice that while the *Point & Click* and the *Drag & Drop* tasks required the execution of single P or D operator, the *Item Selection* tasks requires the execution of two P operators (one for menu activation and another one for item selection). Besides that, item selection requires the execution of a complex M operator to take the decision of what item to select.

With regard to the sample descriptive indicators (tables 3 to 6) it is noticeable that the sample is mainly composed by individuals which are young, male and have intensive experience in the use of computers. However, the number of observations that correspond to the other types of web applications users (women, elder and low experienced users) is sufficient to conduct a valid statistical study. Furthermore, and regarding handedness, around 11% of the individuals in the sample are left-handed. This value is consistent with the global rate of left-handed people, that is estimated between 10% and 13% (Raymond et al. 1996).

4.2. Regression analysis results (hypotheses h_1 and h_2)

Table 4 indicates the main results of the three regression models and the related tests. Prior to the comment of the results we must highlight that all CIs of the different variables in the three regression models are below 15. In accordance to this, all VIFs are below 10. These values are common thresholds to discard the presence of significant multicollinearity among the variables of a linear regression model (Nachtsheim et al. 2004). For the sake of clarity in the presentation of the results we do not include CI and VIF values in table 7, but data are available from the authors upon request.

Results of the Cook-Weisberg test for heteroscedasticity are detailed in the last row of the table. We indicate the values of the chi-squared test statistic and the corresponding p value. As the null hypothesis for this test is that variance is constant we can conclude that such hypothesis is rejected in the three cases and heteroscedasticity is significant. So, we repeated the estimation of the regression equations using a robust estimation procedure, which consisted in the calculation of robust standard errors for the coefficients in the different regression equations and robust p-values, including White's correction (White 1980). Such results are those displayed in table 7.

The layout of the rest of the table is as follows: in each column, we show the statistics for each one of the regression equations (where, the dependent variables are, respectively, time for completion of point & click, drag & drop, and item selection tasks). The first five cells of each column contain the estimates for each one of the independent variables and the intercept of each model. In each cell, the upper figure is the coefficient estimate, that in the middle is the robust t statistic (computed using the standard error that includes White's correction) and the figure shown in the lower part of the cell is the robust p-value. In addition, the table displays for each model, the F statistic for the test of the joint significance of the coefficients and its p-value, as well as the adjusted R^2 and the results of the aforementioned Cook-Weisberg test for heteroscedasticity.

With regard to the results, we must first underline that although R^2s are not very high, conducted F tests evidence the jointly significance of the coefficients of the variables, that is, the set of variables, considered as a unit, influence the performance in all the tests.

Regarding the parameter estimates, coefficients for age are always positive and significant. These results give support to the first hypothesis (h_1) as it is evidenced that older users perform worse for all the tasks (needing more time to complete. So, the performance decline regarding the age is confirmed.

Furthermore, the gender coefficient is significantly positive in all cases. So, women perform worse on *Point & Click*, *Drag & Drop* and *Item Selection* tasks. These results suggest that hypothesis h_2 also holds corroborating the observations made by Inkpen (Inkpen 2001) regarding girls having difficulty with *Drag & Drop* tasks. However, our data does not support Inkpen's other observations related to the absence of any significant gender difference in the overall movement time. That leads us to conclude that Inkpen's results are more related to children's different learning styles than directly to the gender.

	Point & Click	*Drag & Drop*	*Item Selection*
Age Group	Parameter estimate: 500.74 t statistic: 4.40 p-value: <0.001	Parameter estimate: 1721.47 t statistic: 6.86 p-value: <0.001	Parameter estimate: 3077.63 t statistic: 8.82 p-value: <0.001
Gender	Parameter estimate: 1742.03 t statistic: 4.29 p-value: <0.001	Parameter estimate: 5077.81 t statistic: 3.91 p-value: <0.001	Parameter estimate: 6293.93 t statistic: 4.32 p-value: <0.001
HoursUse	Parameter estimate: -394.89 t statistic: -3.02 p-value: 0.003	Parameter estimate: -895.07 t statistic: -2.46 p-value: 0.014	Parameter estimate: -1481.98 t statistic: -4.73 p-value: <0.001

Handedness	Parameter estimate: 707.98 t statistic: 1.18 p-value: 0.238	Parameter estimate: 2986.57 t statistic: 1.53 p-value: 0.127	Parameter estimate: 107.61 t statistic: 0.07 p-value: 0.946
Intercept	Parameter estimate: 16759.52 t statistic: 19.21 p-value: <0.001	Parameter estimate: 30363.66 t statistic: 16.99 p-value: <0.001	Parameter estimate: 57363.51 t statistic: 27.88 p-value: <0.001
F test	F-statistic: 17.48 p-value: <0.001	F-statistic: 13.46 p-value: <0.001	F-statistic: 33.27 p-value: <0.001
Adjusted R^2	11.41%	17.59%	25.49%
Cook-Weisberg test for heteroscedasticity	Chi-squared: 7.07 p-value: 0.007	Chi-squared: 388.80 p-value: <0.001	Chi-squared: 44.60 p-value: <0.001

Table 4. Regressions results and related tests.

Regarding the control variables in the model, it is first noticeable that prior experience with computers is significant in all cases. Coefficients for *HoursUse* are significantly negative in all cases, meaning that more hours of computer use always imply a better performance. These results are similar to those obtained by a prior study (De Andrés-Suárez et al. 2015), which was focused on observations based on the execution of top level interaction tasks, mostly related with cognition and perception. Our findings confirm that the same effect is observable at the low level of interaction required by the GOMS analysis, which is mostly based on the human motoric system. With regard to the other control variable, handedness does not seem to have an influence, as left-handed users perform neither significantly best nor significantly worse than right-handed.

With respect to the additional post-estimation tests, we must first underline that Cook's D values are always lower than 1 for all the individuals in all the regressions so there are no influential cases in the models. Second, none of the quadratic terms (Age^2) that we included in alternative versions of the equations was found to be significant. So, we can reject the existence of non-linear effects for the age variable.

Finally, the re-estimation of the models for different subsamples defined considering the browser used for the test (Chrome, Firefox, IExplorer,) and the operating system (Windows, Linux, Mac) produced results which are qualitatively the same as those displayed in

Table 4. For the sake of brevity, we did not include in the paper the results. However, they are available from the authors upon request.

4.3. Correlation analysis results (hypothesis h_3)

The results of the correlation analysis we conducted to assess whether individuals that perform well in a certain task also perform well in the others (h_3) are shown in table 8. In each of the cells we display the non-parametric Spearman correlation coefficient (upper figure) and the corresponding p-value (lower figure). Cells below the main diagonal contain the results of tests.

	Point & Click	*Drag & Drop*	*Item Selection*
Point & Click			
Drag & Drop	0.731 0.000		
Item Selection	0.660 0.000	0.674 0.000	

Table 5. Results of the correlation analysis.

Data in table 8 evidence that correlations are significant among all the tasks. This finding supports hypothesis h_3, suggesting that the execution time (performance) of an individual in a specific task, keeps its coherence in the other tasks as well. So, for example, if a person has superior performance in the execution *Point & Click* tasks, she/he is expected to also have superior performance in the execution of *Drag & Drop* and *Item Selection* tasks.

This finding may have a relevant impact in the future design of automatic user modeling algorithms. As the three proposed interaction tasks have the same usefulness in terms of user categorization, any of them can be used separately to achieve this goal. Moreover, the amount of data required to automatically infer the type of user may be notably reduced (as only one task is analyzed), which is crucial for the execution of real time algorithms.

5. Limitations, Future Directions, and Conclusions

This work had two interrelated goals. First, we wanted to assess whether the gender and age are sufficiently significant determining factors to support an automatic profiling system based on the analysis of mouse motion behavior when executing *Point & Click*, *Drag & Drop* and *Item Selection* tasks. Second, to figure out whether the individuals perform consistently across these basic interaction tasks, that is, if their performance in one of them are extrapolable (or not) to the others.

Regarding the first, the results of the empirical study reveal that both age and gender factors are significantly determinant. While older users performed worse than younger in each the interaction tasks, men obtained better results than women. On the other hand, in relation with the analysis of correlations between the execution times of the target basic interaction tasks, data gathered in the tests revealed consistencies in the execution times of individuals across them. User's performance measured in any of these tasks is coherent to their execution time in the other tasks.

These results open the door to implement a system that automatically classifies users in age and gender groups by observing the way they interact and perform in these basic interaction tasks with any web interface. However, these evidences must be taken carefully, given that the data was gathered through artificial and isolated adhoc tests, and not in a real web interface where the behavior of the user can differ from the one evidenced during the tests. On the other hand, the existing correlation between the way individuals perform across the different interaction tasks makes it more flexible not only to integrate the data gathering processes into the final system (since developers do not need to force the use of all of them), but also it expands the data gathering possibilities to a number of observations whose results could be combined in a hybrid voting algorithm or a machine learning based system.

The possible benefits of such a classification system are straightly applicable in e-commerce sites, the main target of this work, since the information architecture of the site (and the list of products or sales offered) could be adapted accordingly to the preferences of this target user. However, there are other possible applications like preventing some users to claim the identities of other users or from pretending being a different age and/or have a different gender. In addition, detecting old users would support the automatic adaptation of the interface to the specific features of this group, using for example bigger fonts and simpler interfaces.

Besides the design, implementation and evaluation of this system in a real environment, we consider there are other factors that could somehow determine user's performance in basic interaction tasks and that should be considered to extend this study in the future. One of them is the cultural factor. The sample used is limited to western cultures individuals. Some studies suggest that the culture of an individual could determine his/her performance. Ford et al.(Ford & Gelderblom 2003) designed an experiment to evaluate if any of the Hofstede's cultural dimensions can affect human performance while interaction with computers. Even though their study did not provide sufficient evidence to reach any determining conclusion, we consider it would be interesting to extend this work to a multicultural sample of individuals to study such influence in these specific types of interaction.

6. Acknowledgments

This work has been funded by the Department of Science and Technology (Spain) under the National Program for Research, Development and Innovation: project TIN2009-12132, entitled

SHUBAI: Augmented Accessibility for Handicapped Users in Ambient Intelligence and in Urban Computing Environments.

References

Abdulin, E., 2011. Using the keystroke-level model for designing user interface on middle-sized touch screens. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*. ACM, pp. 673–686.

Agudo, J.E., Sánchez, H. & Rico, M., 2010. Playing Games on the Screen: Adapting Mouse Interaction at Early Ages. *2010 10th IEEE International Conference on Advanced Learning Technologies*, pp.493–497.

Alpert, S.R. et al., 2003. User attitudes regarding a user-adaptive ecommerce web site. *User Modelling and User-Adapted Interaction*.

De Andrés-Suárez, J. et al., 2015. Towards an automatic user profiling system for online information sites. *Online Information Review*, 39(1), pp.61–80. Available at: http://www.emeraldinsight.com/doi/pdfplus/10.1108/OIR-06-2014-0134.

Ardissono, L. & Goy, A., 2000. Tailoring the interaction with users in Web stores. *User Modelling and User-Adapted Interaction*.

Beckwith, L. et al., 2006. Gender HCI: What About the Software? *Computer*, 39(11), pp.97–101.

Beckwith, L., 2003. Gender HCI issues in end-user software engineering. In *IEEE Symposium on Human Centric Computing Languages and Environments, 2003. Proceedings. 2003*. IEEE, pp. 273–274.

Beckwith, L. & Burnett, M., 2004. Gender: An Important Factor in End-User Programming Environments? *2004 IEEE Symposium on Visual Languages - Human Centric Computing*, pp.107–114.

Beckwith, L. & Burnett, M., 2007. Gender HCI Issues in End-User

Software Engineering Environments. ... *Software Engineering*.

Bestelmeyer, P.E.G. & Carey, D.P., 2004. Processing biases towards the preferred hand: valid and invalid cueing of left- versus right-hand movements. *Neuropsychologia*, 42(9), pp.1162–7.

Boulinguez, P., Barthélémy, S. & Debû, B., 2000. Influence of the movement parameter to be controlled on manual RT asymmetries in right-handers. *Brain and cognition*, 44(3), pp.653–61.

Card, S., Thomas P. Moran & Allen Newell, 1983. *The Psychology of Human Computer Interaction*, Lawrence Erlbaum Associates.

Carvalho, D.C.M. et al., 2015. Interaction Paradigms Versus Age-Related User Profiles : an Evaluation on Content Selection. *IEEE Latin America Transactions*, 13(2), pp.532–539.

Chadwick-Dias, A., McNulty, M. & Tullis, T., 2002. Web usability and age: how design changes can improve performance. *ACM SIGCAPH Computers and the Physically Handicapped*, (Issue 73-74, June-September 2002), pp.30–37. Available at: http://dl.acm.org/citation.cfm?doid=960201.957212 [Accessed March 19, 2014].

Chen, P.T. & Hsieh, H.P., 2012. Personalized mobile advertising: Its key attributes, trends, and social impact. *Technological Forecasting and Social Change*, 79(3), pp.543–557.

Cheong, Y., Shehab, R.L. & Ling, C., 2013. Effects of age and psychomotor ability on kinematics of mouse-mediated aiming movement. *Ergonomics*, 56(6), pp.1006–20.

Choi, J., Lee, H.J. & Kim, Y.C., 2011. The Influence of Social Presence on Customer Intention to Reuse Online Recommender Systems: The Roles of Personalization and Product Type. *International Journal of Electronic Commerce*, 16(1), pp.129–154.

Christoph, G. et al., 2014. Adolescents' computer performance: The role of self-concept and motivational aspects. *Computers & Education*, 81(Februrary 2015), pp.1–12.

Collazos, César; Guerrero, Luis A. ; Llana, Mónica; Oetzel, J., 2016. Gender: An influence in the collaborative work process in computer-mediated communication. *International Conference on Nanotechnology and Environment Engineering*, (June 2014).

Cook, R.D. & Weisberg, S., 1983. Diagnostics for Heteroscedasticity in Regression. *Biometrika*, 70(1), pp.1–10.

Czaja, S.J. & Lee, C.C., 2006. The impact of aging on access to technology. *Universal Access in the Information Society*, 5(4), pp.341–349.

Czerwinski, M., Tan, D.S. & Robertson, G.G., 2002. Women take a wider view. In *Proceedings of the SIGCHI conference on Human factors in computing systems Changing our world, changing ourselves - CHI '02*. New York, New York, USA: ACM Press, p. 195.

Dickinson, a., Arnott, J. & Prior, S., 2007. Methods for human – computer interaction research with older people. *Behaviour & Information Technology*, 26(4), pp.343–352.

Fijałkowski, D., 2011. An architecture of a Web recommender system using social network user profiles for e-commerce. , pp.287–290.

Fisk, A.D. et al., 2009. *Designing for Older Adults: Principles and Creative Human Factors Approaches* Second., CRC Press.

Ford, G. & Gelderblom, H., 2003. The Effects of Culture on Performance Achieved through the use of Human Computer Interaction. *Human-Computer Interaction*.

Fozard, J.L. et al., 1994. Age differences and changes in reaction time: the Baltimore Longitudinal Study of Aging. *Journal of gerontology*, 49(4), pp.P179-89.

Freudenthal, D., 2001. Age differences in the performance of information retrieval tasks. *Behaviour & Information Technology*, 20(1), pp.9–22.

Garg, A. et al., 2006. Workshop on Information Assurance Profiling Users in GUI Based Systems for Masquerade Detection

Workshop on Information Assurance. , pp.48–54.

Ghazarian, A. & Noorhosseini, S.M., 2010. Automatic detection of users' skill levels using high-frequency user interface events. *User Modeling and User-Adapted Interaction*.

Guiard, Y., Olafsdottir, H.B. & Perrault, S.T., 2011. Fitt's law as an explicit time/error trade-off. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. New York, New York, USA: ACM Press, p. 1619.

Haque, A., Sadeghzadeh, J. & Khatibi, A., 2006. Identifying potentiality online sales in Malaysia : a study on customer relationships online shopping. *The journal of applied business research.*, 22(4), pp.119–130.

Heathcote, A., Brown, S. & Mewhort, D.J., 2000. The power law repealed: the case for an exponential law of practice. *Psychonomic bulletin & review*, 7(2), pp.185–207.

Helsen, W.F. et al., 1998. Manual Asymmetries and Saccadic Eye Movements in Right-Handers During Single and Reciprocal Aiming Movements. *Cortex*, 34(4), pp.513–530.

Hill, R.L. et al., 2011. Older web users' eye movements. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. New York, New York, USA: ACM Press, p. 1151.

Howard, R.W., 2014. Learning curves in highly skilled chess players: A test of the generality of the power law of practice. *Acta Psychologica*, 151, pp.16–23.

Inkpen, K.M., 2001. Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM Transactions on Computer-Human Interaction*, 8(1), pp.1–33.

Joiner, R. et al., 1998. It is best to point for young children: a comparison of children's pointing and dragging. *Computers in Human Behavior*, 14(3), pp.513–529.

Kim, H.-J., Bae, K. & Yoon, H.-S., 2007. Age and Gender Classification for a Home-Robot Service. *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human*

Interactive Communication, pp.122–126.

Kuhtz-Buschbeck, J.P. et al., 1998. Development of prehension movements in children: a kinematic study. *Experimental Brain Research*, 122(4), pp.424–432.

Lee, M.-C., 2009. Understanding the behavioural intention to play online games: An extension of the theory of planned behaviour. *Online Information Review*, 33(5), pp.849–872.

Lenhard, A. & Hoffmann, J., 2007. Constant error in aiming movements without visual feedback is higher in the preferred hand. *Laterality*, 12(3), pp.227–38.

Liu, W., Jin, F. & Zhang, X., 2008. Ontology-Based User Modeling for E-Commerce System. In *Pervasive Computing and Applications*. IEEE, pp. 260–263.

MacKenzie, I.S., Sellen, A. & Buxton, W.A.S., 1991. A comparison of input devices in element pointing and dragging tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems Reaching through technology - CHI '91*. New York, New York, USA: ACM Press, pp. 161–166.

Metter, E.J. et al., 1997. Age-Associated Loss of Power and Strength in the Upper Extremities in Women and Men. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 52A(5), pp.B267–B276.

Mieschke, P.E. et al., 2001. Manual asymmetries in the preparation and control of goal-directed movements. *Brain and cognition*, 45(1), pp.129–40.

Morrell, R.W., 2001. *Older Adults, Health Information, and the World Wide Web*, L. Erlbaum Associates Inc.

Nachtsheim, C.J. et al., 2004. Applied linear statistical models. *Journal Of The Royal Statistical Society Series A General*, Fifth, p.1396.

Nap, H.H., Greef, H.P. De & Bouwhuis, D.G., 2005. Access for all by cognitive engineering. *Gerontechnology*, 3(4), p.258.

Neely, K., Binsted, G. & Heath, M., 2005. Manual asymmetries in bimanual reaching: the influence of spatial compatibility and visuospatial attention. *Brain and cognition*, 57(1), pp.102–5.

Neerincx, M. et al., 2000. A practical cognitive theory of Web-navigation: Explaining age-related performance differences. Position paper. In *Workshop Basic Research Symposium CHI2000*. ACM.

Oyewole, S.A. & Haight, J.M., 2011. Determination of optimal paths to task goals using expert system based on GOMS model. *Computers in Human Behavior*, 27(2), pp.823–833.

Peters, M. & Ivanoff, J., 1999. Performance Asymmetries in Computer Mouse Control of Right-Handers, and Left-Handers with Left- and Right-Handed Mouse Experience. *Journal of motor behavior*, 31(1), pp.86–94.

Raymond, M. et al., 1996. Frequency-dependent maintenance of left handedness in humans. *Proceedings. Biological sciences / The Royal Society*, 263(1377), pp.1627–1633.

Rodrigues, P.C. et al., 2012. Stimulus velocity effect in a complex interceptive task in right- and left-handers. *European Journal of Sport Science*, 12(2), pp.130–138.

Rohr, L.E., 2006. Gender-specific movement strategies using a computer-pointing task. *Journal of motor behavior*, 38(6), pp.431–7.

Rosatti L, 2013. How to design interfaces for choice: Hick-Hyman law and classification for Information Architecture. In Slavic A, Akdag Salah A., & Davles S., eds. *Proceedings of the International UDC Seminar,13*. The Hague: Ergon, pp. 125–138. Available at: https://www.academia.edu/6549784/How_to_design_interfaces_for_choice_Hick-Hyman_law_and_classification_for_information_architecture.

Salthouse, T.A., 1984. Effects of age and skill in typing. *Journal of experimental psychology. General*, 113(3), pp.345–71.

Sayers, H., 2004. Desktop virtual environments: a study of navigation and age. *Interacting with Computers*, 16(5), pp.939–956.

Schneider, D.W. & Anderson, J.R., 2011. A memory-based model of Hick's law. *Cognitive psychology*, 62(3), pp.193–222.

Schrepp, M., 2010. GOMS analysis as a tool to investigate the usability of web units for disabled users. *Universal Access in the Information Society*, 9(1), pp.77–86.

Sebora, T.C., Lee, S.M. & Sukasame, N., 2008. Critical success factors for e-commerce entrepreneurship: an empirical study of Thailand. *Small Business Economics*.

Shen, a Y., 2014. Recommendations as personalized marketing: insights from customer experiences. *Journal of Services Marketing*, 28, pp.414–427.

Strom, P.S. et al., 2012. Cyberbullying: Assessment of Student Experience for Continuous Improvement Planning. *NASSP Bulletin*, 96(2), pp.137–153.

Stubbs, N.B., Fernandez, J.E. & Glenn, W.M., 1993. Normative data on joint ranges of motion of 25- to 54-year-old males. *International Journal of Industrial Ergonomics*, 12(4), pp.265–272.

Such, J.M., Garcia-fornes, A. & Botti, V., 2013. Electronic Commerce Research and Applications Automated buyer profiling control based on human privacy attitudes. *Electronic Commerce Research and Applications*, 12(6), pp.386–396. Available at: http://dx.doi.org/10.1016/j.elerap.2013.05.001.

Tan, D.S., Czerwinski, M. & Robertson, G., 2003. Women go with the (optical) flow. In *Proceedings of the conference on Human factors in computing systems - CHI '03*. New York, New York, USA: ACM Press, p. 209.

Tullis, T.S., 2007. Older adults and the web: lessons learned from eye-tracking. In *UAHCI'07 Proceedings of the 4th international conference on Universal access in human computer interaction: coping with diversity*. Springer-Verlag, pp. 1030–1039.

Turban, E. et al., 2008. *Electronic Commerce: a managerial perspective*, Velay, J.L. & Benoit-Dubrocard, S., 1999. Hemispheric asymmetry and interhemispheric transfer in reaching programming. *Neuropsychologia*, 37(8), pp.895–903.

Vines, J. et al., 2015. An Age-Old Problem : Examining the Discourses of Ageing in HCI and Strategies for Future Research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(1), pp.1–27.

Wahlström, J. et al., 2000. Differences between work methods and gender in computer mouse use. *Scandinavian Journal of Work, Environment & Health*, 26(5), pp.390–397.

Walker, N., Philbin, D.A. & Fisk, A.D., 1997. Age-related differences in movement control: adjusting submovement structure to optimize performance. *The journals of gerontology. Series B, Psychological sciences and social sciences*, 52(1), pp.P40-52.

Weiser, E.B., 2000. Gender Differences in Internet Use Patterns and Internet Application Preferences: A Two-Sample Comparison. *CyberPsychology & Behavior*, 3(2), pp.167–178.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), pp.817–838. Available at: http://www.jstor.org/stable/1912934.

Woo, N.H. & Shirmohammadi, S., 2008. Modeling and Measurement of Personality for E-commerce Systems.

Xiang, L.X.L. & Xiaoli, C.X.C., 2010. The research on performance of automobile human-machine interface based on BHR-GOMS behavior model. *Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on*, 2.

Xie, B., 2003. Older adults, computers, and the internet: Future directions. *Gerontechnology*, 2(4), pp.289–305.

Xu, H., 2006. Consumers ' attitudes of e-commerce in China.

Yang, Y. & Claramunt, C., 2007. Towards a DL-based semantic user model for web personalization. , pp.3–8.

Zhai, S., 2004. Characterizing computer input with Fitts' law parameters—the information and non-information aspects of pointing. *International Journal of Human-Computer Studies*, 61(6), pp.791–809.

Zhang, J. & Ghorbani, A.A., 2007. GUMSAWS : A Generic User Modeling Server for Adaptive Web Systems.