"UEQ-MD: a Protocol to Compare User Experience in Mobile Applications Made with Google's Material Design."

Luz María Estrada Irigoyen

A research paper submitted to the University of Dublin, in partial fulfillment of the degree of Master of Science Interactive Digital Media.

May 2016

Declaration

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: http://www.tcd.ie/calendar

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at http://tcd- ie.libguides.com/plagiarism/ready-steady-write

I declare that the work described in this research Paper is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university.

Signed: _____ Luz María Estrada Irigoyen

15/05/2016

Permission to lend and/or copy

I agree that Trinity College Library may lend or copy this research Paper upon request.

Signed: _____ Luz María Estrada Irigoyen

15/05/2016

Acknowledgements

I would like to thank my supervisor Donal O Mahony for his guidance in this project. And I would like to thank my family for their unwavering support.

Thank you all

Summary

In 2014 Google published their Material Design proposal. It is described as a *visual language* that makes use of classic principles of good design (Google design guidelines, 2014). The Google Design specification is available for the general public. It is arguably a comprehensive tool, effective in the creation of successful designs. However this is not the only instrument available, and it is important to assess which tools are more adequate for the needs of each particular project.

This research project wishes to address the question: Is Material Design the best alternative compared to other available design frameworks?. With this in mind a user experience evaluation protocol that can compare mobile applications made with Material Design against other mobile applications with different design guidelines was developed and is presented in this paper. It is a practical instrument for designers and developers to asses the tools available, particularly Material Design. The User Experience Questionnaire - Material Design (UEQ-MD) protocol proposed in this paper has a particular focus on integrating qualitative data that can provide insights into users emotional responses and overall interaction experience. One of the major successes of the UEQ-MD is that it builds upon a validated instrument (UEQ) and that this validation was further developed with the feedback of an expert panel. It is important to continue academic work on instruments that evaluate user experience. It is also increasingly important as technology permeates more and more aspects of everyday life, and as more people around the world have access to interactive devices. The development of critical tools promotes transparency, expands our understanding of human-computer interaction, and pushes the boundaries of UI design.

Table of Contents

Tables and Illustrative Material

Abbreviations

UEQ-MD - User Experience Questionnaire - Material Design

UEQ - User Experience Questionnaire

UX - User Experience

UI - User Interface

API - Application programming Interface

SDK - Software Development Kit

RWD - Responsive Web Design

KDD - Knowledge Discovery in Databases

MD - Material Design

HTML - Hypertext Markup Language

URL - Uniform Resource Locator

IT - Information teCHNOLOGY

"Good language alone will not save mankind. But seeing the things behind the names will help us understand the structure of the world we live in."
Stuart Chase, The Tyranny of Words

Introduction

In modern days we rely on technology to manage our day-to-day life. The worldwide liberalisation, privatisation and deregulation of telecommunication markets (Dunnewijk & Hultén, 2007), the development of mobile technologies, broad band and the Internet revolution created a situation in which a large portion of the world population is now "in constant touch" (Agar, 2003). We make use of devices that allow us to be online at all times. These interactions take place through different graphic user interfaces (UI´s). However we only notice them when something goes wrong, when we do not have a clear understanding of what actions can be taken, when, where and to what end. When this happens users might experience emotions such as frustration, anger, or anxiety. This is one of the reasons why recently User Experience and particularly aspects of experience that relate to emotion have become a focus in the study of interactive products. Although UX testing is common in the industry these techniques and the result of their research are scarcely available (Lai-Chong et al; 2009). Most of the existing information comes from the academic environment and is more readily discussed and shared.

Google is one of the most prominent technology companies in the world with a client base in the millions. In 2014 they published their Material Design proposal. It is described as a *visual language* that makes use of classic principles of good design (Google design guidelines, 2014). The Google Design specification is available for the general public. It is arguably a comprehensive tool, effective in the creation of successful designs. However this is not the only instrument available, and it is important to assess which tools are more adequate for the needs of each particular project. Several questions arise from the idea of having an underlying language throughout the web and interactive devices. For instance, What are the effects on user experience?, Is Material Design a tool that allows and/or promotes originality, branding and identity? This research project wishes to address the question: Is

Material Design the best alternative compared to other available design frameworks?. With this in mind a user experience evaluation protocol that can compare mobile applications made with Material Design against other mobile applications with different design guidelines was developed and is presented in this paper. It is a practical instrument for designers and developers to asses the tools available, particularly Material Design. The User Experience Questionnaire - Material Design (UEQ-MD) protocol proposed in this paper has a particular focus on integrating qualitative data that can provide insights into users emotional responses and overall interaction experience. The current version of the questionnaire (UEQ-MD) currently compares two mobile calendar applications: Google Calendar, which is made with Material Design and Cal Calendar developed by Any.do however it can esily be adapted to test other applications.

The following chapters look at the principles of good design based on cognitive science and print design, identify user interface (UI) design trends that influenced Material Design, examines the context in which Material Design emerged, looks at the internal process within the company that allowed for Material Design to come to existance, study the principles and specifications of Material Design, analyzes user experience and the available methods to evaluate it, and presents the UEQ-MD protocol.

Chapter 1 - What is Good Design Today?

In "The Design of Everyday Life" Don Norman claims that the challenge of good design is to "enhance lives" and "add to our pleasure and enjoyment" (Norman, 2013). In order to identify the principles of good design it is useful to look at cognitive science and usability engineering, and the concepts that allow us to further understand the interactions between people and objects or systems. Some of these concepts are affordances, feedback, conceptual models, and flow.

Google's designers often refer to the idea of *affordances;* this concept describes the relationship between a physical object and a person. It is commonly misunderstood as the intrinsic properties of an object but in fact refers to the properties of both the object and the person (or agent). Affordances refer to "(…) the relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used (Norman, 2013, p 11)." Where affordances describe a relationship, signifiers communicate where the action should take place, in other words, a signifier is any sound, mark or perceivable indicator that communicates to a person the appropriate behaviours that can take place. Creative design incorporates the signifying aspect of design into a consistent experience.

Another basic concept in Interaction Design (and in the Human-Computer-Interaction tradition) is *feedback*, which refers to communicating a result of an action to an agent. In order to have the desired effect, the response must be immediate, even a delay of a tenth of a second can be disconcerting (Norman, 2014). Today technology has reached a point where in many cases feedback is not only immediate but also parallel to user input. For example in 2009 the typical search returns on the Google browser was less than 0.2 seconds (Google Research Blog, 2009), today results appear simultaneously as users type in the search box.

Another useful concept in understanding how people interact with designs is *Conceptual models.* These are a relevant factor in the way a person reacts to a system. It is generated based on previous interactions, the affordances and documentation a person has access to. A good conceptual model allows people to predict the effect of their actions. As Don Norman explains most of the brain's operations are subconscious. It is only during the learning period where a conscious

effort has to be made. Continuous practice produces what is known as "overlearning", this allows performance to become automatic. To summarize, some relevant characteristics of subconscious thought are the tendency to match patterns based on previous experience, detecting general trends, generalizing and making predictions. Arguably Material Design builds upon a combination of physical rules[1] and skills that most users have "overlearned" (such as clicking, swiping, pinching, etc.) which allows the language to become quite intuitive.

Emotions are increasingly important in assessing users experience and measures to evaluate them are being integrated into contemporary research techniques (Vermeeren et al; 2010). "Cognition and emotion cannot be separated. Cognitive thoughts lead to emotions: emotions drive cognitive thoughts. The brain is structured to act upon the world, and every action comes with expectations, and these expectations drive emotions (Norman, 2013, p 46)." As mentioned before Google is a company with a robust infrastructure, technology has reached a point where emotions such as frustration have been drastically minimized, with fast response times, the large capability to render a huge amount of pixels and colours, and the ability to display multimedia content in an efficient manner. However, avoiding frustration is not enough to promote outstanding interactions. Mihaly Csikszentmihalyi´s concept of *flow* is useful to describe a state of consciousness in which people lose track of time and their outside environment[2] (Csikszentmihalyi, 1990); good designs should aspire to promote this kind of immersive experience.

Finally, it is important to notice how, when a person encounters a design (physical objects and artefacts) for the first time; it is often possible to obtain precise behaviour from imprecise knowledge (Norman, 2013), which is to say people are usually able to figure out what actions are possible without the need of training or instruction manuals. As Norman explains there are four reasons for this; the first is that knowledge is both in the world and in people's minds; by combining these two most interactions develop intuitively. The second is that many tasks do not require a lot of precision, so even if the action is not performed exactly the desired effect will

---

[1] Many of the principals of Material Design are based on physical rules.
[2] It also refers to a process of learning where the difficulty level provides a challenge but is not so difficult as to generate frustration.

still be obtained, in UI an example of this is predictive typing which takes spelling and orthographic mistakes into account and is able in many cases to produce the desired result. The third is the fact that in the physical world there are natural constraints, which is to say the perceived properties of an object allow for a person to discern which actions are possible, in the case of UIs these cues must be incorporated through the use of signifiers (arrows, buttons, etc.) that indicate where actions are possible. The fourth reason refers to people´s knowledge of cultural constraints and conventions which influence many reactions and limit certain behaviours.

Another relevant phenomenon is described by Norman as *learned helplessness* which refers to a situation in which after repeated failure of a task a person assumes the task cannot be achieved by them. Many designs often promote this. "When people have trouble using technology, especially when they perceive (usually incorrectly) that nobody else is having the same problems they tend to blame themselves" (Norman, 2014, p63). However it is poor design that usually generates these situations.

The concepts and ideas described above raise interesting questions for UI and graphic design and particularly for Google's Material Design proposal. For instance, how can these natural constraints that appear in the physical world be represented in UI designs? And since Google´s client base includes a very large variety of people from around the world, how can cultural conventions be managed?, how does Material Design avoid or promote *learned helplessness*? The following chapters will attempt to shed some light on these questions.

1.1 Utility and Usability

The combination of *utility* and *usability* ensure the usefulness of a design, system or product (Nilesen, 2012). Both concepts have been relevant in the computer-human interaction tradition for decades. The term *utility* refers to the design's functionality, in other words, it should answer the questions: does it do what the user needs? And does it provide the adequate features to achieve it? (Nilesen, 2012). On the other hand, the term *usability* refers to how easy user interfaces are to use. According to usability expert Jakob Nielsen, *usability* is defined by five quality components,

learnability, *efficiency*, *memorability*, *errors* and *satisfaction*. *Learnability* has to do with how easy it is for users to do basic tasks at the moment of their first interaction with the system or design. *Efficiency* refers to the speed with which the users are able to perform those tasks. *Memorability* refers to what happens if a user stops using the design for a period of time and then comes back to it, how quickly can the user re-establish proficiency?. *Errors* means keeping track of the amount of mistakes a user makes, the severity of them and the amount of time it takes to recover from them (Nielsen, 2012) usually to reevaluate the system, modify it , and improve it. The fifth component is *satisfaction* and refers to how pleasant the overall experience was for the user; this last component has been further developed by user experience methods that will be further described in chapter five.

Chapter 2 - Trends in UI design

Arguably material design has been influenced by several design sources, from print to other user interface (UI) design styles. It is useful to have a brief look of the history of graphical user interfaces in order to identify elements and patterns that endure today and why.

2.1 History of Graphical User Interface

In 1973, the first graphical user interface was built at PARC, using the desktop as a metaphor. The UI introduced windows, icons, menus, file management, and tool palettes (Parc, 2016). The 1981 Xerox Star PC interface, was an important influence for the UI´s developed in companies like Microsoft and Apple (Rawlinson, 2016). In the 80´s and 90´s the metaphor was extended, not very successfully, to include other tangible objects and environments such as offices, hallways, people and animals in Magic Cap and Microsoft BOB (Kruzeniski, 2011). In more modern PC interfaces these initial metaphors cemented themselves and endured. As the technological advances started to allow for the rendering of more pixels and colours, the interfaces continued to reflect as best they could aspects of the real world in the form of bevels, glows, shadows and reflections. User interface controls were also made to resemble those of the physical world such as of knobs, switches, dials and buttons. The Web 2.0 aesthetic cleaned up the look and feel of the desktop, but many of the metaphors still remained (Graham, 2005). This mirroring of the physical world approach made sense when interfaces were first built because of the need to communicate what they were to the users, for this end direct representation was appropriate. However, as time passed and people became accustomed to these new technologies the need to continue to use these metaphors to communicate the interface mechanics decreased substantially. Today, most content on the screen is assumed to be interactive. "In an age where our interactions are information-based rather than tool oriented, a visual communication language that is hinged on arcane artefacts is no longer relevant (Kruzeniski, 2011)". Today the focus is on the content and the message that is trying to be conveyed.

2.1.1 Skeuomorphic Design

One of the design trends that exemplify the mentioned emulating of the physical world is Skeuomorphic Design. It is the technical term for incorporating old, familiar ideas into new technologies, even though they don't necessarily play a functional role (Norman, 2013). Often making the digital element resemble its physical counterpart as much as possible. "One way of overcoming the fear of the new is to make it look like the old (Norman, 2013, p.170)". One of the benefits of this kind of design is that it helps ease the transition from old to new. When the learning stage is over it is important to question the use of these metaphors. Several design trends emerged as a counter reaction to the limitations of Skeuomorphic design one of the more successful styles was Flat Design which was heavily influenced by Swiss design.

2.2 Swiss or international Design

The Swiss or International style is one of the most influential design systems today. It is the basis of the development of graphic design during the 20th century (Budrick, 2015). Swiss design originated in the 1940´s and 50´s in Switzerland but had its roots in the 1920's in countries including Russia, Germany and the Netherlands (Terror, 2009). "These pioneering graphic designers saw design as part of industrial production and searched for anonymous, objective visual communication (Hollis, 2006, p. 158)". Some of their more noticeable contributions where the use of mathematical grids for organising information, a tight structure, strong use of sans-serif typography, and the use of iconography that is recognizable and memorable, simplicity and clarity through the use of only the most essential elements, objectivity through photography and attention to detail; most of these principles were adopted internationally and are still in use today.

In sum, Swiss Design and the Print Design tradition in general have provided several important contributions. Four of them have been particularly influential to UI design. The first is the use of grids as a system that organises the information in a logical, consistent and meaningful framework. Grids improve readability and allow for a visual hierarchy. The second is the use of white space, giving content space, allows

areas with information to have a stronger impact, in addition a balanced page makes the content more easily digestible. The third is the use of only essential elements in the design, avoiding clutter helps the users identify the primary tasks of the interface. The fourth is typography, which has the ability to reinforce hierarchy in the structure of the content. It also provides direction and rhythm through size and weight.

2.3 Towards Material Design

Material Design was influenced by the developments in UI design. Their design guidelines make particular emphasis in the print principles mentioned above as being crucial to their proposed design language. The next paragraphs describe the context in which Material Design was developed and the trends that were predominant at the time.

2.3.1 Metro Design

Microsoft's Metro Design language was a reaction to Apple's *iOS* and Google's *Android*. It has a strong influence from print, particularly Swiss or International Design. It was released in late 2006 with the *Zune Media Player* . This design style was characterised by the use of lower-case menu typography and background imagery based on the content. This style was maintained in the desktop software with the likely objective of promoting integration and consistency across its displays (Greene, 2012). This new design drastically contrasted other Microsoft software such as Windows. In 2010 with the release of *Windows Phone 7* a similar design style was applied that included the use of grid shapes with bright colours, simple sans-serif typography and flat icons.

2.3.2 Flat Design

Flat design gained momentum in 2013. As the name implies this style seeks to remove stylistic characters such as drop shadows, gradients, textures and any other type of design element that creates a three-dimensional effect. It owes a lot of its success to the fact that some of the big players, such as Microsoft and Apple, adopted it early on (Turner, 2014). Flat design was influenced by Swiss design and also by Minimalism. The influence of Swiss design can bee seen in the use of grids

to lay out the information and also in the heavy focus on typography to create hierarchies and allow for better scanning of the content. On the other hand Minimalism, which was popular in architecture, visual art and design much longer before the Web became popular, provided the use of few, essential elements, geometric shapes, bright colours and clean lines (Sandu, 2015). One of the benefits of Flat design is the fact that it works well with modern frameworks and interfaces. As responsive design developed, it was apparent that designs that relied heavily on textures, drop shadows, and fixed imagery did not translate well to smaller screen sizes. However the new tendencies of flat design integrate elements that disrupt an entirely flat design, shadows became a useful tool to add depth and to emphasize certain elements.

The use of Card design derived from Flat design and became very prominent as designers made use of colour blocking to divide the content into a grid. Often using different colours to create a card-like mosaic. These "cards" were usually combined with hover effects to provide visual feedback and guide the interactions.

2.3.3 Card Design

Card as an information disseminator medium have been around for a very long time, as they are a fast and practical way of communicating information. Cards by nature are easily shareable and easy to manipulate. And the virtual version of cards inherits these characteristics. They can be folded for a summary, expanded for more detailed information; they can be flipped around or stacked together (Adams, 2013). They are compatible with animation and have recently integrated multimedia content. The development of mobile technologies made the usefulness of Card design apparent. The information contained in the cards is sourced from different tools like APIs and SDKs[3] (Adams, 2013). The aggregation of content is influenced by several factors such as the interest, preferences and behaviour of the user as well as the location and environmental context. In addition contacts and friends interest, preferences and behaviours also influence the aggregate. There are several examples of websites and companies making use of card for their interface designs

---

[3] API is an interface that allows software programs to interact with each other. SDK is a set of tool that can be used to develop software applications for a specific platform.

Twitter cards is one such example (Twitter cards, 2014). Other examples of the use of cards in UI design is Airdrop iOS 7 is and Google Now. Material Design continues to make use of card design but has adapted it to fit with their material metaphor by focusing on shadows, light, and motion.

2.4 The Influence of Mobile and Responsive Design

Since the iPhone´s release in 2007 the importance of designing for mobile became clear. Most website owners understood that a tailored site delivered a better user experience. Mobile requires to have a clean, simple interface were the important things are available at a glance. The widespread use of Responsive Web Design (RWD) had an overall positive influence on mobile usability, as it increased awareness of basic mobile-usability principles. RWD brought awareness to the need for consistency of content and features across device types. It also brought attention to the fact that content should not be arbitrarily limited, which is to say the answer to a user's question should be the same regardless of where the question is asked: mobile, desktop, or tablet (Budiu, 2015). The prominence of RWD also helped designers seek a balance between chrome, which refers to the UI elements (buttons, menus, links, etc.), and content.

As can be observed from the styles and trends of UI design, Google's Material Design emerged in an environment where there were lessons to be learned. Design trends had shifted from rich design and Skeumorphic tendencies to a version of Flat Design that had reintegrated some three-dimensional elements, in particular shadows. The development of mobile technologies and the surge of responsive design made the use of cards to organise content, focus attention and prioritize actions and messages prominent among most technology companies. In addition, user experience became the centre of attention when developing successful UI´s. Mobile user experience will be further developed in chapter five.

Chapter 3 - Google's Business Model

3.1 Google's Business Model

It is likely that Material Design´s influence in UI and UX design will continue to grow, for this reason, it is important to understand the context and implications of Material Design being developed and owned by Google. By the end of 2000, Google was the most visited search engine available. At that point advertising programs AdWords and AdSense were launched to provide businesses with ways to promote their products and services (Sutherland, 2011). AdWords is based on a bidding system, and AdSense is based on cost-per-click. Another dominant business strategy started early on: acquiring other businesses. Although acquisitions is a common strategy among technology firms in Silicon Valley, Google stands out from other companies like Facebook, Microsoft, Yahoo, Amazon and Apple because of the massive amount of acquisitions it has done since the early stages of its appearance. Some of there high profile aquisitions include 2001 Dejas.com´s Usenet Discussion service, Pyra Labs, the creators of Blogger in 2003, Picasa advanced technology for photo display in 2004, YouTube in 2006, and DoubleClick a digital marketing company in 2007 among others. As of 2015, Google has acquired more than 200 companies (Geis, 2015). This is done not only to expand the workforce and launch new products but arguably to sustain the pace of innovation. Most of Google's products are free, however an exchange still takes place. By using these products and agreeing to the terms and conditions, users allow the service provider to gather their information and use it for advertisement purposes. More than 97% of Google's income comes from the advertising services they provide (Geis, 2015). Arguably as the range of Google product and services have expanded into more and more areas (web, mobile, business, media, geo, home and office, social) the amount of information they have access to has grown exponentially. However Google is able to provide highly targeted advertising and relevant products not only because it has access to a vast amounts of data regarding user's behaviours online but also because it is able to convert it into useful information.

3.2 Knowledge Discovery in Databases (KDD)

Knowledge discovery in databases refers to the process of transforming data into useful information or knowledge. At the core of the KDD process is the application of specific data-mining methods for pattern discovery and extraction. Database marketing systems analyse customer groups and forecast their behaviour. Most businesses including Google use this knowledge to gain competitive advantage. The company leaders in Google had the foresight to appreciate how valuable that information can be in producing and selling products that respond to the specific markets and their needs. Google competes with many technology firms in several sub industries such as computing, software, mobile devices, etc. although each company takes a different approach in their business strategies, Google has a unique Internet status that allows them to enjoy the kind of repeat-usage that other companies crave for. Until 2015 Apple based their strategies on product innovation and customer loyalty and on the interconnectivity of their devices or so called Apple - ecosystem lock. In comparison Microsoft's business model was built, for many years, on servicing businesses and licensing software (particularly Windows OS and the Microsoft Office suit). However in response to Google's disruptive no-fee model they have shifted towards an emphasis on product integration and on "freemium" software packaging. *Freemium* designates "a business model using two products or services, or a combination of products and services. In such a combination, one item is provided at no charge while a complementary item is sold at a positive price to the same general group of customers (Pujol, 2010, p.1)".

By looking at Google's business model it is possible to set Material Design in context, to acknowledge among other things that it is a product owned by a company and might not remain available for a long period of time, or that conditions for its use may change. On the other hand the acquisitions aspect of their business model had direct implications in regards to design within the company, bringing different companies in meant having increased diversity and noise, these are some of the challenges the company had to overcome before a concept such as Material Design was possible. The next section describes this internal process.

3.3 The story of Design in Google

For a long time Google did not show a particular interest in design as compared to other technology companies such as Apple or Olivetti[4]; their focus was around efficiency and speed. Design was done based on small incremental changes rather than having a larger design vision to guide decisions across the product teams (Bohn & Hamburger, 2013). This internal structure where the product teams have a high level of autonomy made efforts to unify Google's visual style across their large range of products quite challenging. Before Material Design there had been attempts to improve consistency and branding. An example of this is in 2008 project *Kanna* led by visual designer Evelyn Kim. Kim and a small team of designers redesigned Mail, Maps, and Search to have a consistent visual style; however this effort did not take of (Kuang, 2015).

3.3.1 Kennedy Project

In 2011 Larry Page took over the CEO role at the company and one of his main priorities was to redesign Google's major applications. In response to this directive, project *Kennedy* was launched, with a focus on the iOS market. One of the main goals of this venture was to create a consistent look and feel across platforms while improving functionality for users. The collaborating project teams identified "simplicity" as one of the strong qualities of the Google brand, and decided to carry this across to other designs (Certin, 2013). To keep to this and avoid clutter only relevant options were displayed on the screen, hiding any fields or buttons that were not needed. Another objective of the Kennedy project was to improve functionality across platforms, in order to do this it was necessary to recognise the differences between platforms and find ways to bring them together through the design (Certin, 2013).

3.3.2 Google Now Cards

Google *Now Cards* predate the appearance of Material Design. They are an example of efforts to improve functionality and apply simplicity in Google's design choices.

---

[4] Italian manufacturer of typewriters, computers, tablets, smartphones, and other electronic devices.

Now cards are customizable cards that organise information automatically and presents it to the users without the need of a search (Google now, 2014). *Now cards* track web and app activity (including searches, Chrome history, and content browsed) and location history (creates a map of where the user goes with logged-in devices) to provide relevant information.

To further understand these changes it is useful to look at the context in which they were happening. On the one hand Apple´s success in the market created a great incentive for Google to rethink its design strategy, their success marked the general public inclination towards well thought designs. On the other hand, the development of telecommunication technologies in particular mobile technologies, created an environment in which Material Design was not only possible but needed in order to continue to compete in the market and improve their products and services.

Chapter 4 - Material Design

"Material Design is a new perspective on what the human and device relationship can be"
Jessica Huang, Senior Visual Designer at Google

In 2014 Google published the Material Design specification. A team of designers was heading the Material Design project; including VP designer Matias Duarte, principal designers Nicholas Jitkoff and Jon Wiley, design manager Jonathan Lee, senior interaction designer Bethany Fong, senior visual designer Christian Robertson among others. On Google's dedicated material design website two main goals are stated. The first is to "Create a visual language that synthesizes classic principles of good design with the innovation and possibilities of technology and science" (Google design guidelines, 2014) The second goal is arguably more ambitious in nature, and is to "Develop a single underlying system that allows for a unified experience across platforms and device sizes. Mobile precepts are fundamental, but touch, voice, mouse, and keyboard are first class input methods" (Google design guidelines, 2014). In this chapter these objectives, the prototyping techniques, and the three design principles will be analysed. The specifications included in the document are extensive and highly detailed, which would make it impossible to cover in this project, however I will highlight particular aspects that are relevant to the research question of this paper.

4.1 Prototyping Techniques

In their research process the Google design team made use of traditional design methods such as brainstorming, sketching, creating wireframes and digital prototypes on software like After Effects. However, they also included prototyping techniques based around photography and studio lighting. Images of lighting and photography techniques can be observed in Figure 1.
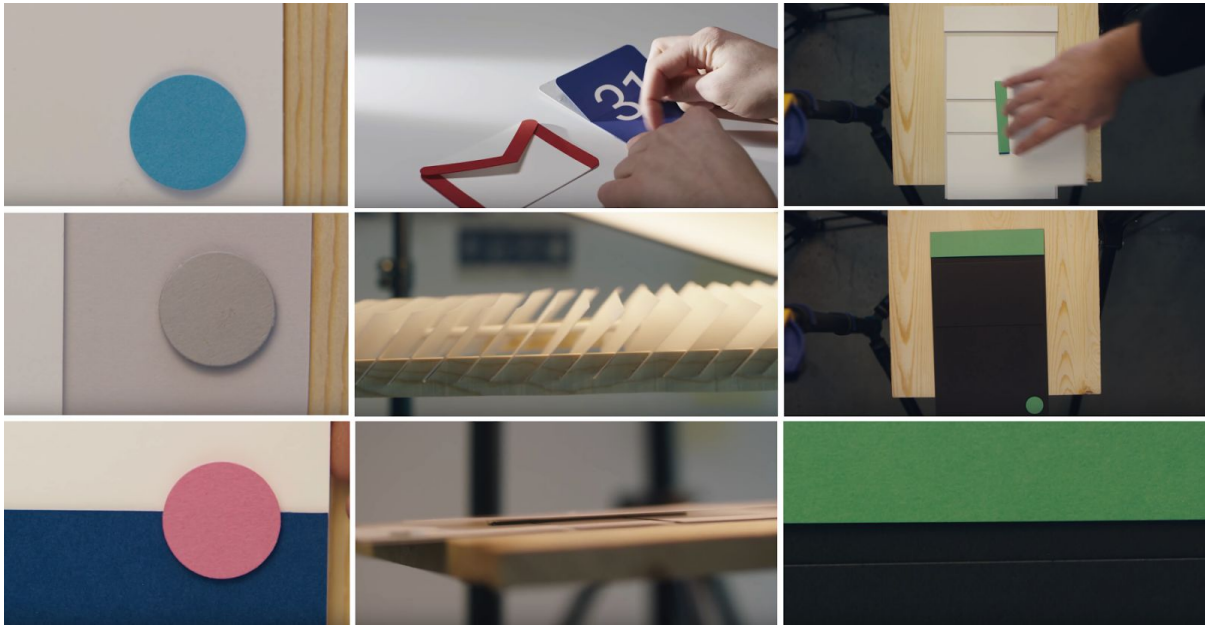
Figure 1. Material Design Prototyping Techniques (Source: Stills from "Making Material Design" Video)

These techniques involved setting up light rigs with paper models of different applications and testing different lighting arrangements. As senior visual designer Christian Robertson explains, the light was kept at the same angle to maintain the shadows consistent from the top to the bottom of the screen. This in turn communicates the relationships between the surfaces and their position in space. "Surfaces and edges of the materials provide visual cues that are grounded in reality. The use of familiar tactile attributes helps users quickly understand affordances" (Google design guidelines, 2014). According to the Material Design team these experiments with physical objects and light allowed them to gain insight into the different ways in which shadows are cast. It also allowed them to identify depth cues, and to learn how to represent and communicate surface in UIs.

There are three principles described in the Material Design document that explain the motivations behind the spec decisions.

4.2 Principal One - Material is the metaphor

The word "material" in Material Design is used as a metaphor to refer to the fact that the design principles were developed based on tactile reality and informed by their prototyping techniques that studied paper and ink. "A material metaphor is the

unifying theory of a rationalized space and a system of motion." (Google design guidelines, 2014) According to the Material Design specification, the objective behind focusing on surfaces and edges is to provide users with *affordances* they can quickly understand and interpret into actions. For example material can vary its dimensions in the x and y axis but has a constant thickness of 7dp, multiple materials can not occupy the same space, and can not go through each other. These properties are understandable as they are rooted in physical properties of paper.

However Material Design´s "smart paper" has behaviours that go beyond the capabilities of real paper as it can heal, shrink, join together, spontaneously generate or be destroyed. This flexibility according to their published document creates new affordances that surpass those of the physical world. To further ground this metaphor the visual treatment of material Design has a focus on light, surface, and movement to communicate how the elements of the interface move, interact, and relate to each other in space.

4.3 Principal Two - Bold, graphic, intentional

This principle refers to taking on the contributions of print design and applying them to the creation of engaging graphic user interfaces. These fundamentals as mentioned before include the use of typography, grids, space, and imagery to create hierarchy, meaning, and to direct and focus attention. Another goal is to create an immersive experience for the user and to emphasize on user input to allow for ease of navigation. The Style section of the specification covers most of the traditional print elements as it includes guidelines on the use of colour, icons, imagery, typography and writing.

The Colour and Typography section describes their colour system as "powerful, immersive and adaptable to any application." It is based around the use of primary and accent colours. The second has the objective of encouraging user interaction. They encourage designers and developers to start from the primary 500´s, which scales from light to dark. Then scale up to the 700´s for status bars or down to the 300´s for secondary information. All the HEX codes and pallets are available to download and refer to.

On the other hand Typography is provided based on a language categorization divided into three: English and English-like languages (Latin, Greek, and Cyrillic scripts), Tall (South Asian and Middle-Eastern languages) and Dense (Chinese, Japanese, and Korean). In this section they mention Roboto a typeface developed to work across the wider set of supported platforms. According to them it gives greater clarity and because it is wider and rounder is perceived as more optimistic.

The goal of creating and underlying language in the web and other interactive devices is truly showcased by integrating all types of languages into the specification and doing so in a consistent manner. Addressing a global audience is also shown by the use of general psychological rules for the use of colour.

4.4 Principal Three - Motion provides meaning

As a general rule in UI design motion should be meaningful and appropriate. In the case of Material Design (MD) UIs the design is transformed based only on user interactions, it is also used to focus and guide the attention of the user. According to the MD guidelines Motion is the ideal way to provide subtle and clear feedback. In this sense, the use of the z-motion is typically a result of user interaction with material. In their words it is meant to "(…) encourages deeper exploration of an app by creating timely, logical screen reactions to the user input". Examples of this visual feedback are ink-like surface reactions, material lifting when touched (to indicate an active state) and direct manipulations such as dragging or flinging sheets of material. Through focus on user interactions a shift from just transmitting information to promoting immersiveness and meaningful communication takes place.

The following chapters define user experience and describe some of the methods available to evaluate it.

Chapter 5 - User Experience

5.1 Defining User Experience

As reviewed in chapter 1, usability is a well-known concept in human-computer interaction. Since the mid 80´s several methods have been developed in order to measure usability aspects, such as learnability and efficiency of a particular application or product. In the past decade, it has been increasingly recognized that users needs go beyond usability and utility, leading to a shift in focus towards a more general experiential perspective (Hassenzahl and Tractinsky, 2006). This comprehensive view includes softer aspects of user interaction such as sensation, meaning, and value. Furthermore as can be recalled from chapter 4 the Material Design principles and specifications have an emphasis on user actions, and have the overall goal of immersing the user in the experience (Google design guidelines, 2014). An immersive experience goes beyond functionality and can be referred to Csikszentmihalyi´s *flow* concept discussed in chapter 1.

ISO provides a useful definition of UX "A person's perceptions and responses that result from the use or anticipated use of a product, system or service" (ISO, 2008). The scope of this concept should include products, systems, services, and objects that a person interacts with through a user interface (Lai-Chong et al; 2009). UX is generally understood as inherently dynamic, because of the constant flux of internal and emotional state a person goes through, this generates differences in the circumstances during and after an interaction with a product.

One of the reasons for this shift in focus is that both in the industry and in academia, limitations of the traditional usability framework became apparent (Lai-Chong et al; 2009). "Traditional methods often focus on usability criteria (…) which correspond roughly to the concepts of usability goals or pragmatic quality. More recent approaches increasingly give attention to the subjective reactions, also including emotional aspects of the user's experience" (Laugwitz, 2016). Criteria in regards to usability are often referred to as *pragmatic quality aspects* while the second group is known as *hedonic quality aspects* (Rauschenberger et al; 2013). Hedonic aspects cover emotions and subjective impressions. Recent research results indicate that

visual aesthetics may be capable of improving performance (Moshagen, Mosch, and Goritz; 2009). For example, aesthetically pleasing web sites seem to influence a variety of constructs related to its capability to promote positive affect, such as satisfaction, preference, customer loyalty, etc. (Norman, 2004). Another aspect relevant to UX is the fact that an increasingly amount of products work in combination with other systems and are integrated across devices. This has implications in regards to consistency of content, chrome and navigation. *Brand Experience* is another factor that influences UX directly, it encompases interactions with the branded products, with the company, and its products and services. It also includes any information that might come from a variety of sources such as the media, peer recommendation or criticism, etc. For example, users tend to forgive the flaws of a brand they love, and may become fixated on flaws of a brand they dislike (Lai-Chong et al; 2009). User experience is a growing field in the industry and in academia, for this reason, several methods have been developed in order to evaluate the quality of users interactions, some of which will be described in the following chapter.

Chapter 6 - Methods to Evaluate UX

There is a wide variety of methods available for UX design and evaluation. As can be inferred from the definitions above user experience should not only be seen as something evaluable after interacting with an object, but also before, and during the interaction. User´s values affect their experiences with products and services, and this relationship should be taken into account in the design process from the beginning (Vujala and Vaananen-Vainio-Mattila, 2009).

Within the European Union (EU) the ENGAGE project collected UX design and evaluation methods between 2004 and 2006. The collected methods were classified into Generative and Evaluative. The Evaluative methods were subdivided into three groups depending on what kind of measures the methods focused on: Sensory characteristics, Expression and Meaning, and Emotional reactions (ENGAGE, 2006).

This project has served as a basis for the development of a diverse range of methods. The following paragraphs describe some of the particular methods that were considered when building the UEQ-MD protocol proposed for this research project.

6.1 A/B Testing

In the industry context, companies frequently make use of A/B Testing online experiments to determine whether a software or design change should be implemented. Online controlled experiments became popular in the late 1990s with the development of networked technologies (Kohavi & Longbotham, 2015). The basic method for A/B testing entails splitting the users randomly between the variants. For example, to compare two different versions of a layout of a web page the users would be split 50% for each version. The interactions with the site are instrumented and the metrics computed (Kohavi & Longbotham, 2015). This method was considered for the protocol because it compares two different elements, however it is not adequate for two completely different products so it was discarded.

6.2 Global self-assessment techniques

Global self-assessment is a broad category of measurement techniques, such as Likert scales and Semantic Differentials, these techniques include many variations of a seemingly simple process. In self-assessment tests participants are asked to provide information regarding behaviour, attitudes, feelings, emotions, etc. One of the main reasons why self- assessment tests are widely used in Psychology and Consumer Psychology are their simplicity and ease of administration (Vermeeren et al; 2006). However, behind what appears to be a simple process there are several steps that take place before the end result. These phases include, understanding and interpreting the questions, searching memory for relevant information, constructing an answer, translating the answer into a meaningful response and editing the response for that particular audience (Lucas and Baird, 2006). Understanding these internal processes can aid in the design and application of the tests and in identifying unwanted variances. A self-report can be considered flawed when it is not logical or when it is influenced by some feature or stimulus that is unrelated to the attitude being evaluated (Lucas and Baird, 2006). There are particular issues underlying each of the three phase of the self-report judgement process which will be expanded on in the paragraphs below.

6.2.1 Understanding the question

Participants of a self-assessment study must first understand the literal meaning of the question. In order to do so it is important to avoid vague or unfamiliar words, complicated sentence structure or any other factor that might lead participants to misunderstand the question (Schwarz, 2004). The next step is to discern the pragmatic meaning of a question, at this point one of the risks is that the participants may try to infer what the person applying the exam had in mind (Lucas and Baird, 2006). Although a degree of error is unavoidable there are statistical and procedural methods to compensate for them.

6.2.2 Formulating a response

In order to provide an answer participants must search their memory for relevant information in regards to the object, product, or service being evaluated. The participant may also compare the object to some relevant standard of comparison, and finally make a judgment toward that object (Schwarz, 2004). Most of the formulating a response processes are subconscious and happen in a few second.

6.2.3 Reporting a response

Anything that impedes accuracy of communication will affect the validity of the report. Although errors are most likely to occur, according to Lucas and Baird these errors often do not severely limit the validity of the measures.

In summary self-report methods evaluate complicated processes, which not always occur in a logical and consistent manner, factors like idiosyncrasy, inaccurate memory recollection, careless responses may introduce unwanted variances. In spite of this these methods have proven to be flexible, efficient, and able to access information that would otherwise be difficult or impossible to obtain (Lucas and Baird, 2006).

6.3 Likert scale

Likert scales were developed in the 1930´s as a five-point bipolar response . These scales range from a group of categories, that go from least to most. Participants must indicate how much they agree or disagree, approve or disapprove, or believe to be true or false. The scale most include at least five response categories (Allen & Seaman, 2007). However it is often extended by adding "very" to the respective top and bottom of the scale. Likert scales were considered for the protocol because of their ease of administration, however semantic differential proved to be more useful to pinpoint aesthetic aspects and emotional responses.

6.4 Stroop Test

In order to compensate for the limitations of self-report scales, many researchers are turning towards implicit methods to assess an individual's predisposition. The classical Stroop test involves having participants look at color words, but asking them to name the ink color, not the word itself. Word and ink color are either

congruent or incongruent. The typical response is that the subjects response is delayed in the incongruent condition this is known as the Stroop effect (MacLeod, 1991). Although the Stroop test dates back to the beginning of experimental psychology in the late 1800s (Jensen and Rohwer, 1966), more modern research has put it to use in the area of computer technologies. In their study, professors Schmettow, Noordzij, and Mundt, apply the Stroop priming test to measure associations that are activated when seeing an interactive computing device, for example a smart phone. Based on their results they describe three categories of users: hedonic, utilitarian, and "geek".

6.4.1 Utilitarian

A utilitarian individual, thinks of technology as a tool to complete a task and achieve particular goals. For people in this category functionality and usability are more important compared to aspects such as aesthetic, and emotion.

6.4.2 Hedonic

Hedonic users are the opposite of utilitarians, they place more importance to the surface features of a product, such as a brand and visual appeal (Schmettow, Noordzij, and Mundt, 2015).

6.4.3 "Geek"

Geek is a term coined by the authors that refers to a person who is a technology enthusiast, it describes individuals that have a particular interest in understanding the inner working of a computer system (Schmettow, Noordzij, and Mundt, 2015). Future endeavours of the present research would attempt to integrate the Stroop priming test to the UEQ-MD protocol.

Chapter Eight will extensively describe the UEQ-MD.

Chapter 7 - User Experience Questionnaire (UEQ)

7.1 Description of the User Experience Questionnaire

The User Experience Questionnaire (UEQ) falls within the self- assessment methods category and makes use of semantic differentials. It was developed by researchers at the University of Applied Science Emden/Leer in Germany. Its purpose is to evaluate continuous user experience of an end-user with the use of 26 bipolar items (Rauschenberger et al; 2013). It is currently available in several different languages including German, English, French, Italian, and Spanish. The questionnaire is made to measures both usability aspects like efficiency, perspicuity, and dependability, but also user experience aspects like stimulation or originality. The UEQ has three main objectives: to deliver a quick assessment, a comprehensive impression of user experience, and a simple and immediate impression of the feel of the interactions between the user and a product (Laugwits and Held, 2016). The UEQ format supports the user´s response to immediately express feelings, impressions, and attitudes that arise when they use a product. The items and scales of the UEQ were created by data analytical approach (Rauschenberger & Schrepp, 2013). The UEQ covers hedonic and pragmatic qualities, according to the authors these are two aspects of the product accept or reject dimension also known as *attractiveness*. Both aspects are important to measure the overall satisfaction of a user with a product or service (Vermeeren et al; 2006). Figure 2 shows the English version of the UEQ.

	1	...	7		
annoying	O	...	O	enjoyable	1
not understandable	O	...	O	understandable	2
creative	O	...	O	dull	3
easy to learn	O	...	O	difficult to learn	4
valuable	O	...	O	inferior	5
boring	O	...	O	exciting	6
not interesting	O	...	O	interesting	7
unpredictable	O	...	O	predictable	8
fast	O	...	O	slow	9
inventive	O	...	O	conventional	10
obstructive	O	...	O	supportive	11
good	O	...	O	bad	12
complicated	O	...	O	easy	13
unlikable	O	...	O	pleasing	14
usual	O	...	O	leading edge	15
unpleasant	O	...	O	pleasant	16
secure	O	...	O	not secure	17
motivating	O	...	O	demotivating	18
meets expectations	O	...	O	does not meet expectations	19
inefficient	O	...	O	efficient	20
clear	O	...	O	confusing	21
impractical	O	...	O	practical	22
organized	O	...	O	cluttered	23
attractive	O	...	O	unattractive	24
friendly	O	...	O	unfriendly	25
conservative	O	...	O	innovative	26

Figure 2. English Version of the UEQ

There are six scales included in the UEQ, Figure 3 describes the relationship between the qualities and the scales. As can be observed, perspicuity, efficiency, and dependability belong to pragmatic qualities and stimulation and novelty correspond to hedonic qualities.

Qualities	Scales
Pragmatic	Perspicuity Efficiency Dependability
Hedonic	Stimulation Novelty

Figure 3. UEQ Quality and Scale relationship

Figure 2 provides examples of differentials corresponding to each scale.

Scales	Examples of differentials
Attractiveness	Annoying/enjoyable
Efficiency	Fast/slow Efficient/impractical Organised/cluttered
Perspicuity	Understandable/not understandable Easy to learn/difficult to learn Clear/confusing
Dependability	Unpredictable/predictable Secure/not secure
Stimulation	Boring/exciting Interesting/not interesting Conservative/innovative
Novelty	Creative/dull Inventive/conventional

Figure 4. Examples of scales in UEQ

7.2 Validity of the Questionnaire

To achieve high scientific quality, when dealing with the results of UEQ research the first step is to confirm the Cronbach´s Alpha data, which describes the consistency of the items of the scale (i.e if all items in the scale measure the same quality) the goal is to obtain a coefficient higher or equal to 0.7 that can confirm high consistency in the scales. (Rauschenberger & Schrepp, 2013). There are two well-known effects that can cause a small value of the Alpha-coefficient for a scale, the first is misinterpretation of some items, for example secure/not secure can be misinterpreted as related to spyware rather than the dependability aspect of a product, system or service (Rauschenberger & Schrepp, 2013). This aspect was particularly relevant for certain items in the UEQ-MD version and will be further explained in Chapter Eight. On the other hand, an item in the scale may also be influenced by a context specific effect due to a special target group, for example age can be a definitive factor in how a product is perceived and there for in how the questionnaire is answered. An example of this is that younger people are less likely to be impressed by certain technologies while people from an older age range might encounter the same technology for the first time and find it impressive. (Rauschenberger & Schrepp, 2013). In regards to the procedure there are additional considerations that should be taken into account such as avoiding deeper rational analysis and retrospective evaluation (Laugwits and Held, 2016).

7.3 Interpreting the Overall Results

The UEQ has a nine steps scale to avoid the well known central tendency bias for items (Heise, 1970). The items are scaled from -3 to +3, -3 represents the most negative answer, 0 a neutral answer, and +3 the most positive answer. The scales perspicuity, efficiency, and dependability should show a strong negative correlation in regards to the completion of the task. (Laugwits and Held, 2016).

Overall the UEQ is a quick, inexpensive, flexible, and effective method of gathering information related to a user's overall experience when interacting with interactive products and services. It provides a solid and reliable base on which to built a self-

report method that is adequate to compare different mobile applications including ones made with Google's visual language Material Design.

Chapter 8 - User Experience Questionnaire - Material Design (UEQ-MD)

Like any method the UEQ has limitations, as Laugwits et al acknowledges, the items of valence do not provide information concerning reasons for the acceptance or rejection of a product (Laugwits and Held, 2016). In addition several researchers agree that a multi-method approach that allows for collection of different types of data is more likely to provide a bigger picture of UX (Vermeeren et al; 2006, Lucas and Baird, 2006). For these reasons and in order to further adapt the questionnaire to the particular goals of this research additional sections where integrated into the UEQ to create the UEQ-MD version. The UEQ-MD evaluates continuous user experience of an end-user with the use of 19 bipolar items and five open-ended questions. In order to test the instrument a ten expert panel with a Computer Science and Graphic Design background analysed the questionnaire and provided feedback.

8.1 Adapting UEQ

The main objectives of the UEQ-MD is to provide a practical tool for independent designers and developers to compare and evaluate the user experience in different mobile applications. It is particularly tailored to asses if the stated goals of Material Design are being met, and if it is the best tool for their particular audience.

As can be recalled from Chapter 4 some of Material Design´s stated goals are to utilize classic principles of good design, to provide *affordances* for users to quickly understand what actions are possible, and to provide the user with the information he needs to understand the relationships between UI objects by focusing on light, surface, and movement (Google design guidelines, 2014).

8.2 UEQ-MD for Mobile Applications

The UEQ-MD protocol proposed in this paper has a focus on mobile applications. Although any design framework, including Material Design, can be used to create interfaces in any device type including desktop and tablet, Material Design has been more widely adopted in the development of mobile applications (Android Apps - Top

charts, 2016). Another reason for focusing on mobile applications is the design challenges it poses. Mobile user experience is affected by several elements such as, *functionality*, *information architecture*, *content*, *design*, *user input*, *mobile context*, *usability*, *trustworthiness*, *feedback*, *help*, *social*, and *marketing*.

Functionality refers to the tools and features that enable users to complete the desired tasks. *Information Architecture* is the logical arrangement of functionality and content to allow users to access information and achieve their goals (Carejo, 2012), this includes navigation and search. As can be inferred *Design* refers to the layout, composition, and visual arrangement of the UI elements. *User input* is data entered by the user by any means, including typing, voice commands, etc. User Input should be minimized on mobile device and should not require the use of both hands (Carejo, 2012). *Usability* was addressed in previous chapters and refers to how easy it is for a user to achieve his goals. *Trustworthiness* describes the level of confidence and trust a user feels when interacting with the application (Carejo, 2012). *Feedback* can be used to attract the user's attention, and also provides the user with cues to let him know the status of the system. *Help* refers to the options, products, and services that are available to assist the user in using the application. *Social* relates to the content and features that promote establishing social networks and that facilitate sharing. The *Marketing* elements are strategic factors that encourage repeated usage. *Mobile context* in particular is complex and has substantial influence in the range of needs a user might have. As illustrated by Figure 4 the user is frequently multitasking in a dynamic environment.

Figure 5. The context of Mobile Interaction

(Source: smashingmagzine.com Author: Nadav Savio)

There are three main types of mobile applications: Native, Mobile, and Hybrid apps.

8.2.1 Types of Mobile Applications

Native apps are installed through an application store like Google Play or Apple's App store. Once installed the app icon remains on the device home screen. Native apps are developed specifically for one platform, and are able to access all the device features (i.e camera, GPS, accelerometer, compass, list of contacts, etc.) In addition Native apps are able to use the device´s notification system and is able to work offline. Mobile web applications are websites that are implemented as a mobile application. They are accessed and rendered by a browser and are typically written with HTML5. Users have the option of keeping the app icon on their home screen by

creating a bookmark to that page, through the use of a special URL. As the name suggests Hybrid applications are part native and part web. They are accessed via an app store. In order to work they rely on the HTML being rendered on a browser; but the browser is embedded in the app. Hybrid apps are able to access all the device features available.

As can be expected each type of application has advantages and disadvantages. Web and hybrid applications have lower maintenance costs and have to deal with less content restrictions, approval processes, etc. Discoverability is highest for hybrid apps as they can be found both by an app store or regular web searches. Native apps are particularly fast, and provide an experience that is consistent with the operating system (Budiu, 2013).

Although arguably the UEQ-MD is effective in comparing any of these types of mobile application it was particularly developed for Native and Hybrid Apps that have access to all the device features as this provides designers with more resources to create rich interactions.

8.3 UEQ-MD Format

The following sections describe the structure, added sections, and relevant details of the UEQ-MD format.

8.3.1 Test Structure

The proposed structure of the test is divided into three sections. First the participant is asked to perform a task, Figure 7 provides examples of these tasks. The second step is a "think-aloud" exercise where the participant is given a minute of free interaction to explore the application's features and functionality. The third part of the test is to answer the UEQ-MD form, which is printed out for each participant. The task and free-interaction sections are video recorded to be analysed, this can provide insights into how well *affordances* and visual cues are functioning. All three sections are timed, the time to answer each app quantitative section is recorded and considered as an additional variable. The demographic information required from the

participants is name, age, and occupation. A notice at the top of the test would also be included to guarantee the participants´ anonymity. In addition, the notice would state that the information would be used for research purposes only.

Type of Application	Task
Calendar	You are meeting Gaby and Harry for coffee next Monday at 15:00 hrs. at the Science Gallery Coffee shop and you want to add the event to you calendar and set a reminder for one hour before.
Weather	You are planning a hiking trip for next Saturday and you want to know what the weather will be like so you can prepare.

Figure 6. Task Examples

8.3.2 Adding Qualitative measures to the UEQ

As mentioned before the UEQ-MD is a modified version of the UEQ. One of the main differences is the open-ended question section. Figure 8 shows five questions that compare the two applications and ask what application the participant currently uses. From these questions qualitative data, such as keywords, could be extracted and analysed to gain insight into the experience of interacting with each application.

Number	Question
1	Which of the two applications did you find more memorable, and why?
2	Which of the two applications did you find more visually appealing, and why?
3	Which of the two applications would you prefer to use, and why?
4	Which of the two applications would you recommend, and why?
5	Which calendar application do you use now?

Figure 7 - UEQ-MD Open-ended Questions

The UEQ items were modified in order to eliminate those that were likely to be misunderstood, that were irrelevant for the particular assessment of a calendar mobile application, or that were rendered irrelevant in the context of the user group being tested. Figure 6 shows which items were maintained, discarded and added and to which scale they correspond.

Scales	Maintained Pairs	Discarded Pairs	Added Pairs
Pragmatic	Annoying/enjoyable Easy to learn/difficult to learn Fast/slow Complicated/easy Inefficient/efficient Clear/confusing Impractical/practical Organized/cluttered	Unpredictable/predictable Obstructive/supportive Secure/not secure Meets expectations/does not meet expectations	
Hedonic	Annoying/enjoyable Creative/dull Boring/exciting Not interesting/interesting Inventive/conventional Unlikeable/pleasing Attractive/unattractive Friendly/unfriendly Unpleasant/pleasurable	Good/bad Motivating/demotivating Conservative/innovative	Tensed/relaxed Dull/engaging

Figure 8. Discarded, Maintained, and Added Pairs

All together the UEQ-MD has nineteen items, eight correspond to the pragmatic scale, and eleven to the hedonic scale. The questionnaire was separated into two sections, the first inquires about particular aspects of the interactions with the application and the second has four items that assess the overall experience, this is done twice once for each application that is being compared. Figure 7 shows the quantitative section of UEQ-MD.

Figure 9. UEQ-MD Quantitative Section

8.3.3 Participants

Participants for testing represent an average user: western men and women within the age range of 15 to 45. In this protocol, an average user is considered to be comfortable with technology and can navigate an interface with relative ease[5].

---

[5] The number of smartphone users in 2016 is forecast to reach 2.08 billion (Statista, 2016).

Exclusion criteria includes people with low Information Technology (IT) experience, people that have already used one or both applications, and non-native English speakers[6].

8.4 Testing the Instrument

As mentioned before, in order to test the instrument ten experts with Computer Science and Graphic Design backgrounds examined the UEQ-MD with two Calendar applications and provided feedback. In this section the test and feedback from the experts will be described.

8.4.1 Selecting Applications

Several factors should be taken into account when selecting applications to compare. For instance it is necessary for them to have similar features, for instance, in the case of two calendars apps both applications should be available for iOS and Android devices, and have common basic features such as real time synchronization to all calendars, voice entry and predictive text, integrated contacts, navigation with Google maps, and customizable reminders. Figure 10 shows two mobile calendar applications with similar features and concepts that are examples of adequate candidates for testing with the UEQ - MD. On the left Google Calendar, which was made using Material Design, and on the right Cal Calendar developed by design agency Any.do.

---

[6] since the UEQ is a semantic differential it is especially important that the user see the items in his native language (Rauschenberger and Schrepp, 2013, p.40).
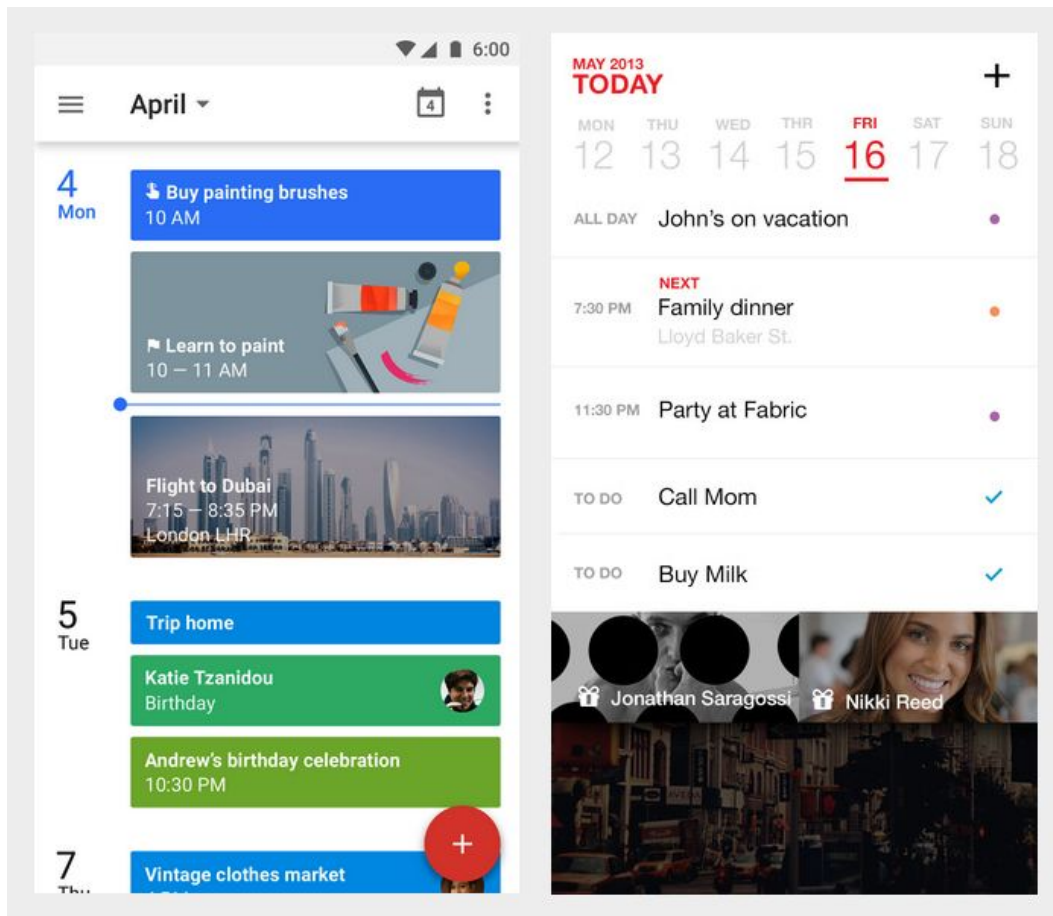
Figure 10. Images Calendar Applications - left Google Calendar, right Cal Calendar
(Source: Google App Store Website)

Another relevant factor when selecting applications for the test is that the overall concept is not too different (i.e humorous VS minimalistic). Finally, it is ideal to compare applications that have a different aesthetic style in order to evaluate identity, originality, and Branding.

8.4.2 Expert Panel Feedback

The panel provided written and verbal feedback in regards to the questionnaire. It was also reviewed against the literature. The feedback was then classified into different categories to identify the protocol's strengths and weaknesses. Figure 8 shows the categories and their corresponding feedback.

Category	Feedback
Integration	● Ecosystem- lock ● Privacy Issues
Special Requirements	● Hardware ● Storing Method
Period of Experience	● Momentary ● Single-Episode
Potential Bias	● Social Desirability ● Participant Predisposition
Test Format	● Duration
Location	● Field ● Lab

Figure 11. Categories and Feedback

The following paragraphs expand on each feedback category.

Integration

Interestingly, Integration was considered to be positive for some experts and negative for others. The main argument for integration as a positive incentive for usage was the automatic dissemination across devices that saves the user from having to input the information more than once. The counterpart of integration had to do with privacy issues that arise from giving further access to technology corporations. This contradicting views were helpful to remind how some features can be interpreted in very different manners and to try to foresee what the implications are for the criteria.

Special Requirements

The UEQ-MD testing requires a smartphone and a video recording device, Cloud storage and external harddrive to store video material and software for the applications being tested (many applications are free or fall within the Freemium model). Special Requirements such as the need for a specific hardware and software affect the practicability of the test and must be taken into account.

Period of Experience

An important characteristic of a method is, what period of user experience it studies. The UEQ.MD is designed to evaluate short-term usage. According to authors Vermeeren et al the test can be considered *Momentary* as it makes use of questionnaire, self-report methods, and think-aloud exercises. It also falls within the scope of *Episode-test* as it makes use of retrospective video analysis. Although this is useful to gain insight on what interactions are like for users that encounter these applications for the first time, and to gain input on how visually attractive they find the interfaces it does not provide information on how the user experience would transform over weeks, months, or years.

Potential Bias

Social Desirability affects the communication of self-reported judgements, it refers to the participants tendency to endorse items that are considered socially positive. The impact of social desirability can be reduced by ensuring anonymity to the participants and also with various statistical techniques (Lucas and Baird, 2006). On the other hand participants predisposition can be assessed and taken into account by integrating the Stroop priming test into the UEQ-MD protocol. This would allow to classifying participants into utilitarian or hedonic and to account for some unwanted variances.

Test Format

Most of the panel agreed that the shorter the test time the less taxing for the participants. If participants become exhausted this could potentially compromise the reliability of the data (vermeeren et al; 2016). Based on the panel's feedback the

average total time for administering a test, including all three sections, was 13 minutes. However by removing the five questions the time was reduced by approximately 5 minutes. In order to get a valid impression the test requires at least thirty participants (Raushenberger et al; 2013). For future application two versions of the UEQ-MD would be designed the first including the section with open-ended question and the second without them (see Figure 9). Only a small percentage (depending on resources available) of the overall participants would be asked to fill out the full version. This would improve the practicality aspect of the protocol, while still providing qualitative data.

Location

Field and laboratory locations were considered in order to assess the potential differences. The field testing would be useful to provide realistic situations in which factors such as noise, different light conditions, etc. affect the interactions. On the other hand observation and video recording would be considerably more efficient in a controlled laboratory environment. It could also reduce the amount of potential stress and distractions the participant might experience. Further instrument testing in different locations including field and laboratory can be conducted to determine which context would provide the best insight into user experience.

Chapter 9 - Conclusions

As mentioned before, Google is first and foremost a business, it is indeed one of the most powerful and influential companies in the intersection between technology and consumer access (Desjardins, 2016). As such they have a large range and variety of products and services, among them Material Design. Google is particularly unpredictable in its business decisions compared to other companies, and product cancellations are common (Mace, 2013). For this reason, designers and developers should not rely on Material Design being permanently available. In addition, styles and trends in design often come and go, it is impossible to know if Material Design will remain and for how long but the fact is that it will transform, as it is "a living document" that will continue to be adapted and modified (Google design guidelines, 2014), and as independent designers and developers contribute to its specification.

It is likely that Material Design´s influence in UI and UX design will continue to grow, and making critical use of its resources and learning from its different applications is useful and important. The analysis from the chapters above indicate that the Material Design language is effective in constructing successful UIs. It has an original and logical approach for composing user interfaces, with the use of material as a metaphor to guide design decisions. It is particularly effective in making proper use of affordances, and cues for the user to access functionalities of an interactive system with ease in a relatively intuitive manner. Nevertheless, the human device relationship is complex, as Don Norman explains, successful design must take place in all three levels: visceral, behavioral, and reflective (Norman, 2004). Without a critical use of the Material Design framework one of the potential risks is having a homogeneous web and interactive device experience, and to a degree loosing identity and originality in UI designs. It is important to continue asking the question what is good design? And continue to challenge the tools we have available.

9.1 Scoping: Further Development of the UEQ-MD

The UEQ-MD protocol is a tool for designers and developers to asses if the Material Design framework is indeed appropriate for their particular project, needs, and audience, and if it provides the desired user experience. One of the major successes

of the UEQ-MD is that it builds upon a validated instrument (UEQ) and that this validation was further developed with the feedback of an expert panel, and by reviewing against the literature and other available research techniques. It also has the benefit of being enriched with qualitative measures to gain better insight into user experience emotional aspects. The UEQ-MD takes phenomenons such as *learned helplessness* into its considerations and tries to mitigate it by selecting simple tasks, and taking other contextual factors into account.

However, as mentioned in chapter 7 all methods have their limitations and the UEQ-MD is not an exception. It requires further testing and a more rigorous validation of the instrument including, assessing the internal consistency with the use of the Alpha - coefficient and other relevant statistical techniques. Criterion validity to compare results against the behaviors of users in day to day life, and predictive validity to allow for follow up measures that provide a more broad sense of how user experience changes over longer periods of time, can also be further developed and improved. Integrating the Stroop priming test is another possible improvement to the protocol as it can provide insights to better understand participants predisposition and what it means for the criteria.

Since the industry does not often share its UX testing methods or their results with the general public, it is important to continue academic work on instruments that evaluate user experience. It is also increasingly important as technology permeates more and more aspects of everyday life, and as more people around the world have access to interactive devices. The development of critical tools promotes transparency, expands our understanding of human-computer interaction, and pushes the boundaries of UI design.

References

2016, S. (2016) *Number of smartphone users worldwide from 2014 to 2019 (in millions)*. Available at: http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ (Accessed: 2 May 2016).

About PARC (no date) Available at: http://www.parc.com/about/ (Accessed: 6 May 2016).

Adams, P. (2013) *Why cards are the future of the web - inside intercom*. Available at: https://blog.intercom.io/why-cards-are-the-future-of-the-web/ (Accessed: 8 May 2016).

Agar, J. (2014) 'Constant touch: A global history of the mobile phone',*Choice Reviews Online*, 51(08), pp. 51–4484–51–4484. doi: 10.5860/choice.51-4484.

Android Apps - Top Chart (2016) Available at: https://play.google.com/store/apps/top?hl=en (Accessed: 7 May 2016).

Any.do Cal - your life. Your calendar (no date) Available at: http://www.any.do/cal/ (Accessed: 7 May 2016).

Bargas-Avila, J.A. and Hornbæk, K. (2011) 'Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience', *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, , pp. 2689–2698. doi: 10.1145/1978942.1979336.

Bohn, D. (2013) *Redesigning Google: How Larry Page engineered a beautiful revolution*. Available at: http://www.theverge.com/2013/1/24/3904134/google-redesign-how-larry-page-engineered-beautiful-revolution (Accessed: 12 May 2016).

Bright, P. (2012) *Microsoft: 'Metro' out, 'windows 8-style UI' in, amid rumors of a trademark dispute*. Available at: http://arstechnica.com/information-technology/2012/08/microsoft-metro-out-windows-8-style-ui-in-amid-rumors-of-a-trademark-dispute/ (Accessed: 22 April 2016).

Budiu, R. (1998) *The state of mobile user experience*. Available at:
https://www.nngroup.com/articles/mobile-usability-update/ (Accessed: 23 April 2016).

Budiu, R. (2013) *Mobile: Native Apps, web Apps, and hybrid Apps*. Available at:
https://www.nngroup.com/articles/mobile-native-apps/ (Accessed: 23 April 2016).

Budrick, C. (2015) *Swiss style: The principles, the typefaces & the designers - print magazine*. Available at:
http://www.printmag.com/typography/swiss-style-principles-typefaces-designers/ .

Cerejo, L. (2012) *The elements of the mobile user experience – smashing magazine*.
Available at:
https://www.smashingmagazine.com/2012/07/elements-mobile-user-experience/ (Accessed: 2 May 2016).

Connolly, E. (2015) *Has visual design fallen flat? - inside intercom*. Available at:
https://blog.intercom.io/has-visual-design-fallen-flat/ (Accessed: 15 May 2016).

Csikszentmihalyi, M. (1990) 'Flow: The psychology of optimal experience', *Choice Reviews Online*, 28(01), pp. 28–0597–28–0597. doi: 10.5860/choice.28-0597.

Curtin, M. (2013) *Unified design: Inside Google's Kennedy project*. Available at:
https://www.unleashed-technologies.com/blog/2013/02/13/unified-design-inside-google%E2%80%99s-kennedy-project (Accessed: 12 May 2016).

Desjardins, J. (2016) *Every 10 Seconds, the World's Tech Giants Make This Much Money…*. Available at:
http://www.visualcapitalist.com/tech-giants-visualizing-profits-for-every-10-seconds/ (Accessed: 30 April 2016).

Eadicicco, L. (2014) *LEAKED: A closer look at Microsoft's big plan to fix the disaster that was windows 8*. Available at:
http://uk.businessinsider.com/windows-10-screenshots-leak-2014-12?r=US&IR=T (Accessed: 12 May 2016).

ENGAGE, Report on the evaluation of generative tools and methods for 'emotional design' (2006) *Deliverable D15.3. EU project Engage 520998*, .

Google Design (2015) *Making material design*. Available at: https://www.youtube.com/watch?v=rrT6v5sOwJg (Accessed: 7 May 2016).

Google now (no date) Available at: https://www.google.com/landing/now/#whatisit (Accessed: 6 May 2016).

Google (PPC) pay-per-click online advertising (no date) Available at: https://www.google.ie/adwords/ (Accessed: 6 May 2016).

Graham, P. (2005) *Web 2.0*. Available at: http://www.paulgraham.com/web20.html (Accessed: 8 May 2016).

Greene, J. (2012) *Why metro now rules at Microsoft*. Available at: http://www.cnet.com/uk/news/why-metro-now-rules-at-microsoft/ (Accessed: 8 May 2016).

Hassenzahl, M. and Tractinsky, N. (2006) 'User experience - a research agenda', *Behaviour & Information Technology*, 25(2), pp. 91–97. doi: 10.1080/01449290500330331.
Heise, D.R. (1970) *Semantic differential*. Available at: http://www.indiana.edu/~socpsy/papers/AttMeasure/attitude..htm (Accessed: 2 May 2016).

History.com (1958) *Swiss design: Design is history*. Available at: http://www.designishistory.com/home/swiss/ (Accessed: 12 May 2016).

Hollis, R. (2006) *Swiss graphic design: The origins and growth of an international style*. Available at: http://www.jannuzzismith.com/downloads/swissgraphicdesign_pr.pdf (Accessed: 12 February 2016).

Introduction - material design - Google design guidelines (2014) Available at: https://www.google.com/design/spec/material-design/introduction.html (Accessed: 14 March 2016).

Jensen, A.R. and Rohwer, W.D. (1966) 'The stroop color-word test: A review', *Acta Psychologica*, 25, pp. 36–93. doi: 10.1016/0001-6918(66)90004-7.

Kruzeniski, M. (2011) *Mattspost*. Available at: http://www.slideshare.net/mikekruzeniski/how-print-design-is-the-future-of-interaction (Accessed: 26 February 2016).

Kuang, C. (2016) *How Google finally got design*. Available at: http://www.fastcodesign.com/3046512/how-google-finally-got-design (Accessed: 7 February 2016).

Kujala, S. and Väänänen-Vainio-Mattila, K. (2009) 'Value of Information Systems and Products: Understanding the Users' Perspective and Values', *Journal of Information Technology Theory and Application (JITTA)*, , pp. 521–530.

Lai-Chong, E., Roto, V., Hassenzahl, M. and Kort, J. (2009) 'Understanding, scoping and defining user experience', , pp. 719–728. doi: 10.1145/1518701.1518813.

Laugwitz, B. and Held, T. (2016) 'Construction and Evaluation of a User Experience Questionnaire.', *Lecture Notes in Computer Science HCI and Usability for Education and Work: . Print.*, , pp. 63–76.

Laurans, G., Hekkert, P. and Desmet, P. (2009) 'The Emotion Slider: A Self-report Device for the Continuous Measurement of Emotion.', *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (2009)*, .

Learn how AdSense works (2016) Available at: https://www.google.com/adsense/start/how-it-works/ (Accessed: 6 May 2016).

Lucas, R.E. and Baird, B.M. (2006) *Handbook of Multimethod measurement in psychology*. Edited by Michael Eid and Edward Diener. United States: American Psychological Association.

Mace, M. (2013) *Google logic: Why Google does the things it does the way it does*. Available at: https://www.theguardian.com/technology/2013/jul/09/google-android-reader-why (Accessed: 12 April 2016).

MacLeod, C.M. (1991) 'Half a century of research on the Stroop effect: An integrative review', *Psychological Bulletin*, 109(2), pp. 163–203. doi: 10.1037/0033-2909.109.2.163. Moshagen, M., Musch, J. and Göritz, A.S. (2009) 'A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance', *Ergonomics*, 52(10), pp. 1311–1320. doi: 10.1080/00140130903061717.

Müller, F. (2016) *Flat design: A history, past, present and future*. Available at: http://www.onextrapixel.com/2014/11/07/flat-design-a-history-past-present-and-future/ (Accessed: 3 March 2016).

Nakamura, J. and Csikszentmihalyi, M. (2014) *The concept of flow*. Available at: http://nuovoeutile.it/wp-content/uploads/2015/12/2002-Flow.pdf (Accessed: 8 May 2016).

Nilesen, J. (1998) *Usability 101: Introduction to usability*. Available at: https://www.nngroup.com/articles/usability-101-introduction-to-usability/ (Accessed: 12 May 2016).

Norman, D.A. (2003) *Emotional design: Why we love (or hate) everyday things*. 3rd edn. New York: Basic Books.

Norman, D.A. (2013) *The design of everyday things: Revised and expanded edition*. New York: Basic Books.

Rauschenberger, M., Schrepp, M., Perez-Cota, M., Olschner, S. and Thomaschewski, J. (2013) 'Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (UEQ).Example: Spanish language version', *International Journal of Interactive Multimedia and Artificial Intelligence*, 2(1), p. 39. doi: 10.9781/ijimai.2013.215.

Rawlinson, N. (2016) *Apple turns 40 today - here's its history from 1 april 1976 to 2016*.
Available at:

http://www.macworld.co.uk/feature/apple/history-of-apple-steve-jobs-what-happened-mac-co
mputer-3606104/ .

Sandu, B. (2015) *Flat design is not A trend, it's been around for some time*. Available at:
http://www.designyourway.net/blog/inspiration/flat-design-is-not-a-trend-its-been-around-for-
some-time/ (Accessed: 8 May 2016).

Schmettow, M., Noordzij, M.L. and Mundt, M. (2015) *An implicit test of UX: Individuals differ
in what they associate with computers*. Available at: http://doc.utwente.nl/87275/ (Accessed:
6 May 2016).

Schwarz, N. (2004) 'Metacognitive experiences in consumer judgment and decision making',
Journal of Consumer Psychology, 14(4), pp. 332–348. doi: 10.1207/s15327663jcp1404_2.
Statistics Roundtable: Likert Scales and Data Analyses (2016) Statistics Roundtable: Likert
Scales and Data Analyses.

Terror, D. (2009) *Lessons from Swiss style graphic design – smashing magazine*. Available
at: https://www.smashingmagazine.com/2009/07/lessons-from-swiss-style-graphic-design/
(Accessed: 12 May 2016).

The Microsoft Timeline (2016) Available at:
https://www.microsoft.com/about/companyinformation/visitorcenter/timeline.htm (Accessed:
18 April 2016).

Turner, A.L. (2014a) *The history of flat design: Efficiency, minimalism, Trendiness*. Available
at:
http://thenextweb.com/dd/2014/03/19/history-flat-design-efficiency-minimalism-made-digital-
world-flat/#gref (Accessed: 8 May 2016).

Twitter cards (2014) Available at: https://dev.twitter.com/cards/overview (Accessed: 8 May
2016).

Vermeeren, A.P., Law, E.L.C., Roto, V., Obrist, M., Hoonhout, J. and Väänänen-Vainio-Mattila, K. (2010) 'User experience evaluation methods: current state and development needs, APA', *In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, , pp. 521–530.