Gender Human-Computer Interaction: *Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective*

By

Gulnaz Kaur Randhawa

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (M.Sc.) in Computational Sciences

The Office of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian Université/Université Laurentienne
Office of Graduate Studies/Bureau des études supérieures

Title of Thesis Titre de la thèse	Gender Human-Computer Interaction: Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective
Name of Candidate Nom du candidat	Randhawa, Gulnaz Kaur
Degree Diplôme	Master of Science
Department/Program Département/Programme	Computational Sciences
Date of Defence Date de la soutenance	November 04, 2021

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:

Dr. Ratvinder Grewal
(Supervisor/Directeur(trice) de thèse)

Dr. Ramesh Subramanian
(Committee member/Membre du comité)

Dr. Kalpdrum Passi
(Committee member/Membre du comité)

Dr. Pradeep Atrey
(External Examiner/Examinateur externe)

Approved for the Office of Graduate Studies
Approuvé pour le Bureau des études supérieures
Tammy Eger, PhD
Vice-President Research (Office of Graduate Studies)
Vice-rectrice à la recherche (Bureau des études supérieures)
Laurentian University / Université Laurentienne

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

Abstract

Consumer market has witnessed remarkable shift from designer's preferences to what actual users want. Not only has this led to hike in market profit but also resulted in products with enhanced User-Experience. In the field of HCI, designers have been working to make their end products usable to large group of users belonging to different ages, gender as well as tech-savvy and non-computer literates. Considering 'gender', much research has been done and is still being conducted to explore its' relevance to Human-Computer Interaction. This research addresses the researcher's vision to identify gender differences in User-Experience received from using mobile applications. The idea is to experimentally prove whether there exists same-gender or opposite gender credibility in rating the usability of mobile applications. If no statistically significant differences are noted in male and female respondents' respective ratings of applications from male and female designer, then null hypothesis will be accepted. This would imply that UI designers did not target users of any particular gender during the design phase thereby producing gender-inclusive designs. On the contrary, if differences are found then alternate hypothesis shall be accepted which will be an indication of gender bias being propagated in designs.

Data collected from 30 participants (15 males and 15 females) showed no statistically significant differences in the ratings of male and female designer's interfaces. Future research has been proposed with greater sample size along with additional test variants such as virtual gendered avatars.

Keywords: Inclusive Design, Gender Bias, User-Experience, Gender Stereotypes

Acknowledgement

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Ratvinder Grewal for giving me this golden opportunity to conduct research in the field of my utmost interest. His perfect blend of professionalism, outstanding subject knowledge, expertise, and motivation has inspired me always. I feel extremely grateful for being mentored by Dr. Grewal under whose invaluable guidance at every step in my research, I learnt the core concepts related to research methodology in a true sense and worked on my weaknesses. I would like to extend my heartfelt thanks and dedicate my entire journey of two years at Laurentian University to Dr. Grewal for accepting me as his research student and believing in my vision.

I am extremely thankful to my parents for their love, support and giving me a chance to fulfill my dream of pursuing education abroad. I will forever be grateful to them for their tremendous support throughout my time as research student. Special note of appreciation to my father for always letting that spark shine inside me which has been my guiding light since the beginning of master's and to my mother for her words of encouragement and keeping me in high spirits.

Special thanks to Laurentian University's incredible library and database services for providing access to books, research papers and journals. I am also thankful to all the professors from whom I have had a chance of being taught the Graduate courses. Last but not the least, thanks to all the members of Computer-Human Interaction Lab at Laurentian University.

Table of Contents

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

1. HCI: Human-Computer Interaction
2. UX: User Experience
3. UI: User Interface
4. TSA: Tactile Situation Awareness System
5. SUS: System Usability Scale
6. SA: Situation Awareness
7. App: Application
8. FAPP.F: Female designed app evaluated by females
9. FAPP.M: Female designed app evaluated by males
10. MAPP.F: Male designed app evaluated by females
11. MAPP.M: Male designed app evaluated by males

Chapter 1: Introduction

1.1 Background of Gender-HCI

Gender Human Computer Interaction (henceforth referred to as Gender HCI) is a sub-field of HCI which explores the relationship between gender and the design of interfaces, computer software and technological artifacts. HCI communities involved in gender research have been actively analyzing the role of gender in the design and development of technology. Gender awareness has paved the way for design teams to make interfaces with gender diversity in consideration.

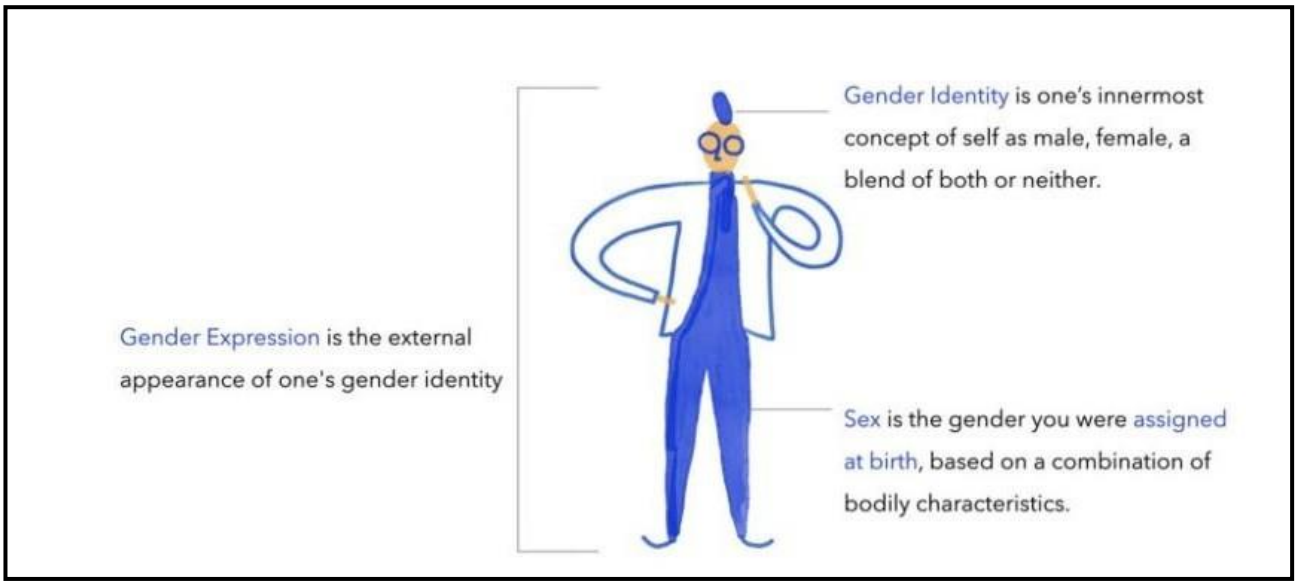
Broadly categorizing various spheres where the roots of Gender HCI have been planted, or in other words, the rising impact of gender on technology has been examined:

1. Gender differences in the use of interfaces and the User Experience perceived.
2. Evaluation of technological devices and interfaces from the perspective of males and females. Examples include car seat belt design, chest strap to monitor heart rate etc.
3. Improvisation in the design of video games to target audience of both the genders.
4. Implementation and introduction of gender-inclusive/gender-neutral designs.
5. Feminist-HCI bringing awareness about reshaping existing products and designs to make them more women-friendly.

1.1.1 What is Gender?

Whether we are conscious or not, our identity is greatly shaped by how we identify ourselves on gender spectrum. We use the word, 'gender-spectrum' rather than gender because as per the latest trends, people's perception on gender are changing rapidly. Gen Z population believes strongly that one's birth gender does not define one's gender identity. Due to the same reason, in today's rapidly evolving society, gender is best defined as a spectrum with multiple degrees of gender identity and gender expression.

Figure 1.1: Sex, Gender Identity and Gender Expression. Source: [13]



Chapter 1. Introduction

Some important terms and concepts:

1. Gender vs Sex: It is important to make distinction between the terms gender and sex. By 'sex' it refers to biological characteristics which identifies whether a person is male or female, that is, biological sex assigned at birth. 'Gender' on the other hand is associated with social construct perspective which is more personal and individual, beyond the labels of man and woman. Gender identification and expression might not necessarily align with biological sex as it is the representation of masculine or feminine characteristics irrespective of one's sex. Consider it through this illustration, a person born as male might identify as female with orientation towards feminine aesthetics.

2. Gender Identity: refers to how an individual view their own gender internally irrespective of their biological sex assigned at birth

3. Gender Expression: refers to how an individual expresses their gender externally in the form of looks, acts and attire which influences other's perceptions of their gender.

4. Trans: are those individuals whose biological sex does not align with their gender identity.

4. Cis gender: are those group of people whose biological sex (assigned at birth) matches their gender identity.

1.1.2. What is Inclusive Design?

Inclusive design is a methodology which results in products being more useful to largest diversity of people by eliminating biases and obstacles that users might face. The term 'inclusive design' was first coined to support social inclusivity towards ageing and disability population [14]. In present times, it is blending well into business and companies where mainstream products are designed to be more inclusive in order to reach a wider market [15].

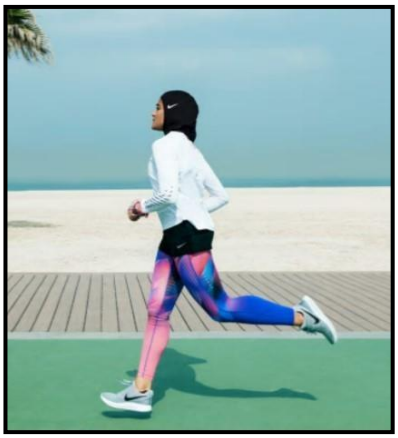
The first step towards an inclusive design is an inclusive team. Diverse team leads to collaboration of ideas and feedback from multiple angles. Accepting people from all backgrounds, gender, age, cultures lead to a more colorful palette of ideas being intermixed which is a crucial step towards developing universal designs. Diversity in team ensures that same problem is tackled from contrasting perspectives. The terms inclusive design, universal design and design for all, are a call or reminder for the designers to consider numerous forms of human differences which could be related to age, abilities, gender, ethnic backgrounds, lifestyle, culture etc. [16]

Quoting certain examples of inclusive design [17]:

1. Inclusive Design at Nike:

As an act of encouraging Muslim girls into sports, Nike created Pro Hijab which would make it easier for hijab wearing women to participate in athletics. Believing that around 800 Muslim women around the world are neither foreigner nor strangers but valuable customers, Nike relentlessly supported the production of this product.

Figure 1.2: Pro Hijab. Source: [17]



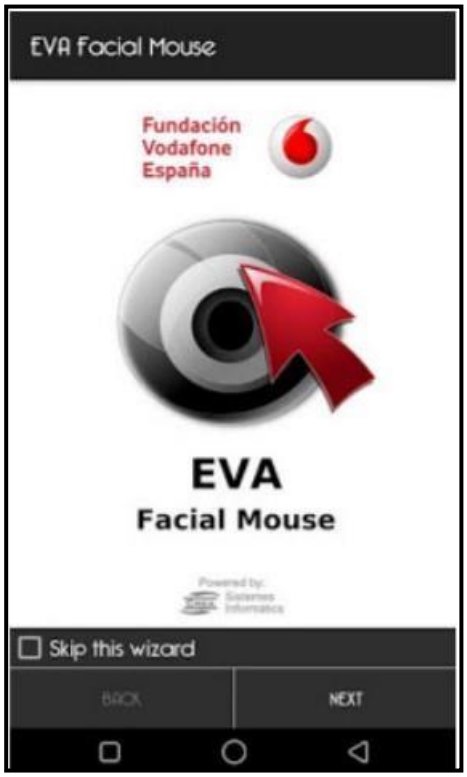
2. Bank of America:

With more than 13,000 members, Bank of America's Hispanic Latino Organization for Leadership and Advancement (HOLA) has proved to be a driving force in the company's success. This has been attributed to positive community engagement. Recognizing the fact that greater percentage of new chequing account holders are Hispanic/Latino, HOLA put forward the appeal to have more resources available in Spanish. Therefore, the mobile app of Bank of America is now available in Spanish with more than 1.3 million active users.

3.EVA Facial Mouse:

Developed by CREA Software Systems, EVA Facial Mouse is a free app for users unable to access touchscreen, for example, individuals with amputations, spinal injuries, cerebral palsy or other disabilities. Through the medium of an interaction channel built within the app, users can control an Android device by their face tracking. App's front camera captures the face movements, hence allowing users to control pointer on the screen which functions like a mouse pointer. This way users can operate the elements on the interface giving them a sense of independence.

Figure 1.3: EVA. Source: [17]



1.2 What is Gender Inclusion in Technology?

Gender inclusion is a concept which explicitly addresses the issue of bias propagation in current technological products. With universal adoption of technology in every sector, decision-making processes, and day to day life, it has been viewed as a catalyst in the direction of achieving gender equality. Gender inclusivity in tech industry implies that no gender group feels biased by the design or ergonomics of any product. Stereotypical gender roles prevalent in the society have earlier been transcended into the development of ubiquitous computing thereby influencing the perception of what a 'man' or 'woman' should be. In early 90's, majority game developers being males, created game content particularly seeking male audience.

The characters in those games represented typical macho image, for example 'Duke Nukem' [18]. This was not received well among female game players, and they disliked the stereotypical notion related to how avatars should look like [19]. When it comes to playing computer games, females emphasize on 'sense of belonging' and 'ownership' [20]. In 1997, traditional perceptions of female characters were re-inforced when Lara Croft by Tomb Raider was introduced. The character Lara depicted a stereotypical erotic appearance to make the game more appealing to young male players. Cal Jones in 'PC Gaming World' said that the underlying issue with Lara was that it was designed by men for men [21]. Commenting about Tomb Raider's success through the use of sexist female avatar, Kate Roberts from Corrosive Software put forward a question, 'had Lara Croft been shown in nice warm sweater and sweatpants, would the company Tomb Raider have been successful in selling as many copies [22].

Figure 1.4: Lara Croft. Source: [23]

The typical slogan, ‘as long as it is pink, it is for women’ failed to view gender as dynamic and evolving [23]. When designs are invented to especially target either males or females, it results in that gender being ghettoized leaving the impression that it needs ‘help’. The emerging alternative is to discard the definition of ‘who a female user is’ and ’who a male user is’ based on societal roles and expectations.

On this stance, the topic of gender-inclusion has been extended in the design and development of ubiquitous technologies that should otherwise be gender-neutral but have proved not to be equally acceptable among men and women.

1.3 Statement of the Study Problem

Evaluation of a final product based on User Experience derived from functionality,
has always been a criterion to determine its’ success or failure in the market. Over the past recent years, successful marketing has relied on what users want. Narrowing down on the term user, it is further classified into male and female users. Same product when subject to evaluation by opposite genders might reveal contrasting user opinions. As an illustration, an average smartphone is 5.5 inches which means that it is an ideal fit for a man to operate it single-handedly –but not practically relevant for an average woman’s hand. This can said to be the result of ‘one size fits men’ approach instead of ‘one size fits all’ [11]. Connecting the dots, such drawback in design has paved the way for another debate on the definition of ‘user’ in the design market. Previously conducted research have successfully pointed out that traditional concept of ‘user’ meant only a ‘man’ which resulted in gender stereotypes being enforced in designs as well.

Rhonda Boyle [3] illustrated the same idea on keyboards. A standard keyboard is forty- eight inches with octaves being 7.4 inches wide. After relating the hand span with the capacity to stretch a given interval on the conventional keyboard; findings revealed that an octave cannot be played comfortably by a significant minority of females. Moreover, a ninth as well as a tenth cannot be played comfortably by a significant majority of females not even ‘on the edge’. Contrary to this, significant majority of males could comfortably play an octave, a ninth and a tenth ‘on the edge’.

The above-mentioned case of smartphones and keyboard clearly indicate dissatisfying User Experience of female users. Such scenarios drift the attention towards the design team of these otherwise highly successful and impressive devices. The question which arises here is that did the design team fail to consider the female consumer group. If viewed from a different angle, should female designers invent products for women owing to their greater ability of understanding female psychology and physiological differences while male designers design for men products. Gloria Moss [4] in her work demonstrated that men and women preferred to choose Christmas cards designed by members of their own gender. In another experiment, as a part of same research, she explored the significant differences in men's' and women's' business cards. M.D.C. Stilma [5] extended the work from Gloria Moss's two-dimensional products to three-dimensional products. The experimental works carried out by both the researchers significantly proved that male and female designer's follow different criteria while designing.

In present times, especially when business, educational and numerous other sectors are functioning online due to Covid-19 restrictions, the dependency on mobile applications has drastically increased for online shopping, virtual meetings etc. Taking into consideration that previous research [4] [5] have been mainly limited to 2-D, 3D products and websites only, this current research aims to examine gendered perspectives of male and female designers during the design phase of mobile applications and the gendered opinions of users of those applications. This shall bring to notice whether the designers succeed in creating gender-inclusive design. The primary objective is to understand if there exists statistically significant difference in the ratings given for the apps designed by male and female designers.

1.4 Hypothesis

Null Hypotheses (H0):

1. For Male Testers: No statistically significant difference exists in the mean SUS (System Usability Scale) ratings given for male designer's app vis-à-vis female designers' app.

2. For Female Testers: No statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

Alternate Hypotheses (H1):

1. For Male Testers: Statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

2. For Female Testers: Statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

SUS is further discussed in detail in Sections 3.5 and 3.11.

1.5 Purpose of Study

This research attempts to shed light on the question whether male and female users rate the UX derived while using an interface originating from male designer vs from female designer significantly different. This shall identify any synthesis between gender of the source and gender of the receiver while rating the perceived credibility of mobile interface. This can be viewed as a step in the direction to determine whether gender bias exists in the world of design. Another angle to view the same is noticing the differences in male's and female's design strategies and if they include more male oriented or female oriented elements in their respective User Interface (UI). Observations that will be drawn from testers' s rating scores and post-experiment questions will depict if the theory of same-gender credibility is valid for mobile interface. In case the 'null' hypothesis is true, this would imply that findings from [4] are not applicable in applications designed for smartphones. However, if alternate hypothesis is true, this would be an indication that interfaces are not gender inclusive. Therefore, the results can guide the future research to further explore:

1. Gender bias being propagated in gender-neutral mobile apps.
2. Why products or interfaces which are supposed to be gender inclusive, appeal more to only one gender.
3. How future technological designs be made more gender inclusive.

1.6 Limitations

Present research has some limitations which must be addressed. In a true sense, the researcher feels that these limitations are the aspects which have been intentionally excluded from the entire experimental setup due to possible complexities to be encountered during testing. To begin with, the mobile application to be designed will be designed to run only on Android smartphones or other Android devices only. Reason being that these will be coded in Android studio and not designed as hybrid applications which run on both iOS and Android operating systems. Developing two mobile applications compatible with both operating systems can be extremely time consuming. The same is true when at the time of testing, errors of different nature will be introduced in Android versus iOS devices. Because of this, only those participants were recruited for testing who owned an Android device.

Secondly, the current research excludes the observations or involvement of 'other' gender groups. The experimenter immensely supports the LGBTQ community and believes that their suggestions would give a new dimension to the future research in Gender HCI. However, due to unavoidable constrains such as remote testing and insufficient literature involving perspective of LGBTQ or individuals identifying as 'non-binary', the current research is focused on only gender binary of male and female.

1.7 Thesis Organization

The first part of thesis explains the key concepts related to gender and inclusive design along with certain examples of inclusive design. The chapter 1 also includes statement of study problem and hypothesis. Purpose behind conducting this research has also been included in this chapter followed by its limitations. The second chapter provides an insight into literature review. Chapter 3 describes the methodology and procedures followed in this experimental study. In chapter 4, quantitative and qualitative results obtained after conducting the experiment have been outlined. Chapter 5 discusses the conclusions drawn from the research work and future research directions have been laid out in chapter 6.

Chapter 2: Literature Review

2.1 Motivating Factors Behind Gender Inclusive Research Studies

It would be unfair to say that the term 'Gender HCI' is a fairly new as the concept or just the mere realization of gendered point of view towards technology, has been always prevalent in the society. Chuck Huff [1] in 1987 conducted the research in the field of Gender HCI through their experimental investigation of the question why computers are more alluring to boys than to girls? The experimental design involved educators (both male and female; skilled in 2.2 programming languages on average) designing an educational software for grade 7 pupils with the aim to teach the use of commas. Results revealed that software designed for males appeared like a 'game' while the one designed for females took the resemblance of a 'learning tool'.

The research work 'Gender Inclusive HCI Research and Design: A Conceptual Review' [7] provides an insight into the importance of adopting 'gender inclusive' approach in HCI designs. Such an endeavor implies avoiding variant designs that target or serve the needs of a specific user-group. In the recent years, emerging awareness of 'gender research' has directed the evaluation of software, websites and several technical devices from a gendered perspective. This paper presented three motivating factors which have acted as a base for work done on gender in inclusive design: Economic, Ethical/Inclusivity and Political/Feminist motivations.

1. Economic Motivation: Software and other digital developments which were previously dominated by male users, are now witnessing an equal or nearly equal percentage of female users. For example, LinkedIn's prominent male user base shifted to 44 percent women and 56 percent men in the year 2019 [8]. Interestingly, social media has a greater number of female audiences who even dominated men in the early adoption of healthy living social technology.

Figure 2.1: Gender Breakdowns for Social Networking Platforms. Source: [7]

Social networking platform	Women (%)	Men (%)
Overall	76%	72%
Facebook	77%	66%
Twitter	21%	24%
	(17% in 2013)	(18% in 2013)
Pinterest	42%	13%
Instagram	29%	22%
Reddit*	4% (2013)	8% (2013)
	36.3% (2014)*	63.7% (2014)*
LinkedIn	27%	28%

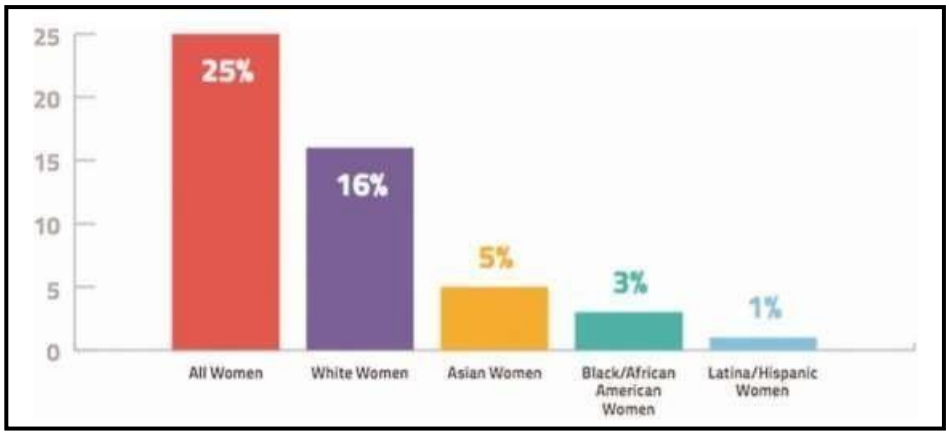
The connection between gender and technological artifacts is not a recent discovery. Before the launch of present time's sophisticated computational interfaces, product designers back in 90's acknowledged the fact that women are an important consumer section of electronic products. Consider the example of telephone which was originally marketed as a business tool [30] but with time soon paved its way into domestic sphere. The advertisements would show the imagery of women engaged in telephonic conversation with their friends while painting their nails. Gradually, the designs of telephones began reflecting feminine aesthetics. One of the extreme cases to mention here would be of iconic telephone 'Hot Lips'.

Figure 2.2: The Hot Lips Telephone. Source: [30]

Van Oost [2] described in their work how gender attribution to materialistic objects such as electric shaver led to their designs being re-designed by a renowned company 'Phillips' for female users. Initially, Phillips launched their electric shavers with the aim of only targeting the male population. Slowly as women began using the shavers due to concerns regarding body hair, Phillips designed their first electric shaver for women with practical changes from the original men's shavers to account for differences in women's hair and skin. Men's shaver was marketed as an electric product whereas the woman's version as a cosmetic tool.

2. Ethical Motivation: When viewed from ethical perspective, if any gender is marginalized by technology, then the situation can be problematic. HCI designers are considered to be following fair policies while designing so that their designs can be equally usable as well as useful for both the genders. Lack of inclusiveness can also influence the career choices for people.

Figure 2.3: Computing Occupations held by Women. Source: [11]

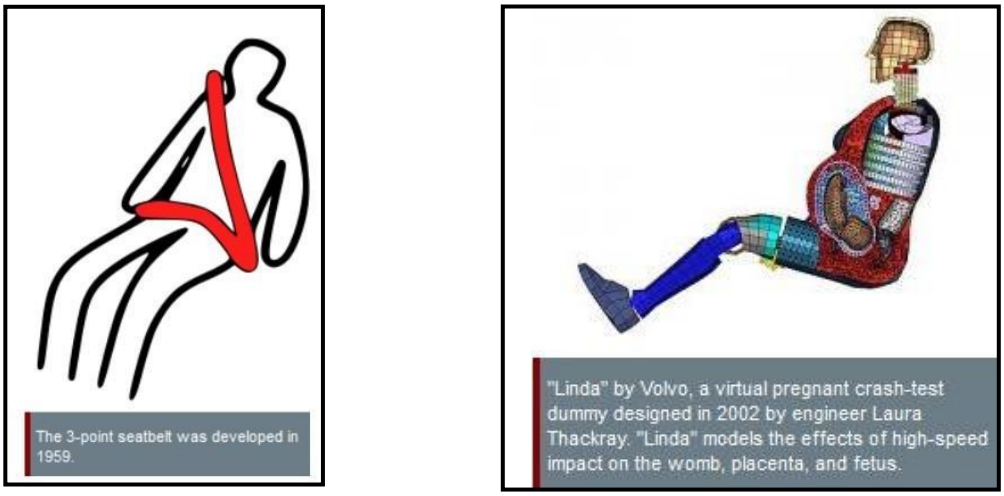


3. Political Motivation: Feminism and Feminist HCI states that technological designs are not gender-balanced and there exists assumptions in designs which result in poor User-Experience for females. This sub-field of HCI aims to prevent masculinization of technological products.

   A prominent example which comes into the picture with this statement is of car crash test dummies [9]. Commonly, a 50th percentile male dummy is used for testing. In 1966, a 5th percentile female dummy was introduced which was merely a scaled down version of male dummy. Even till date, a 50th percentile female dummy is lacking. Hence, when women are involved in a car crash, they sustain more serious injuries than men even while using seat belts. Besides, the conventional seatbelts do not fit pregnant women properly due to which motor crashes are the leading cause of fetal death because of maternal trauma. It is suggested by current research that pregnant women should use 3-pont seatbelt; yet for women who carry low, 3-point seatbelts ride up on the belly. In case of a crash, this results in an increased force transmission to abdomen by three or four-fold relative to the force transmitted when the belt is worn below the uterus resulting in an increased risk of fetal injury.

   In 2002, Volvo developed a virtual pregnant dummy ‘Linda’ in her 38-week of pregnancy. In 2019, Volvo announced its E.V.A. (Equal Vehicle for All), an initiative which recognizes that physical dummies need to represent a greater variety of body types.

Figure 2.4: 3-point seatbelt (left) and “Linda” by Volvo (right). Source [9]



Caroline Criado Perez in her book [10] has drawn attention towards numerous existing technical domains where women seem not to be recognized as a potential user group. As an illustration, the author mentions about Tactile Situation Awareness System (TSAS) which is used by pilots to be always aware of their orientation with respect to ground level. This vest is fitted with 32 sensors which vibrate when pilots adjust their position. A review on TSAS emphasized that vibrations are best detected on hairy, bony skin and are more troublesome to detect on soft, fleshy areas of body. Now pondering upon an undeniable fact that US force comprises of 20 percent women; since women have breasts and particularly do not have any hairy chest, relying on TSAS can put their safety at risk while flying in comparison to their male counterparts.

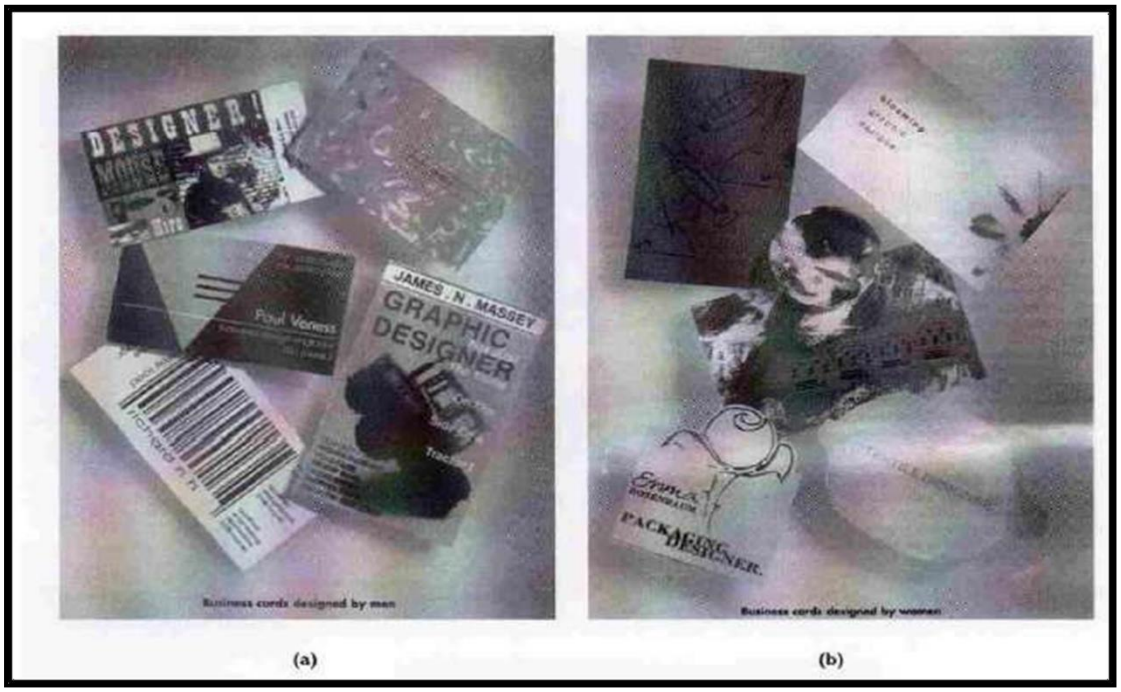
Figure 2.5: Tactile Situation Awareness System Vest. Source: [31]

2.1 Gloria Moss Studies

2.2.1 Business Cards

Publication from Gloria Moss [4], director of Product Psychology, involved quasi - experimental design to examine gender differences in consumer choices and preferences. In experiment 1, business cards collected from 144 female and 83 male designers were relatively compared across the attributes of their sizes and colors. For each collected card, the experimenter noted whether it was from a male/female designer, fell within standard size parameters and whether it was printed on a colored or white paper.

Figure 2.6: Representative business cards from a male designer (a) and a female designer (b) in experiment 1. Source: [4]



The observations drawn revealed that 54 percent of cards from males were within standard size in comparison to 38 percent of females' cards. Regarding the colors, 74 percent of the male designers used white card compared to 53 percent of female designers. It was concluded that females' choice of non-standard size and colored cards indicated that women prioritize aesthetics over practical considerations and form above function.

2.2.2 Two-Dimensional Designs

Conclusions drawn for 2-dimensional designs in [4] were further tested on three-dimensional designs by M.D.C. Stilma [5]. An experiment was performed involving 20 female and 44 male students where each participant had to look for 2 samples of a single product type; one being designed by a male designer and the other by a female designer. The products were then evaluated according to a questionnaire derived from previous studies of Gloria Moss. Significant score of 6 out of 21 proved that different characteristics exist in the design of products originating from designers of opposite gender. The limitation of the study is said to be the free choice of products that eventually resulted in a wide range of products being evaluated. One of the research questions in the paper from 2003 [6] was determining if the gender of information source interacts with the gender of the receiver in evaluating the perceived credibility of information on personal web sites. Both the websites differed only by the names 'Jeff Newmar' and 'Julie Newmar'. The story included in each site was identical and revolved around the harmful effects of radiation on pregnant women who fly in airplanes. During the results analysis, opposite gender credibility evaluations emerged to be higher than the same gender. The judgement of female's site was less favorable by female participants whereas males rated Julie's site the highest.

2.3 Bias and Stereotypes

Research from Stanford researchers [24] suggested that our evaluation or judgement of products is greatly impacted by gender stereotypes. One of the researchers, Soule in their work comments that customers show less willingness to purchase 'traditional' male products if they are manufactured by women. This can be explained with an illustration:

"*Imagine a scenario where you read the label of a craft beer and see that the brewer is Jane. So now realizing that this beer was made by a woman, will your perception of it change? Will you compare it with the taste of beer made by a man*" or

"*While buying cupcakes you are informed that they were baked by a man say 'John'. Will this impact your expectations. Do you contemplate whether John's cupcakes will be equally delicious as, say Mary's*"?

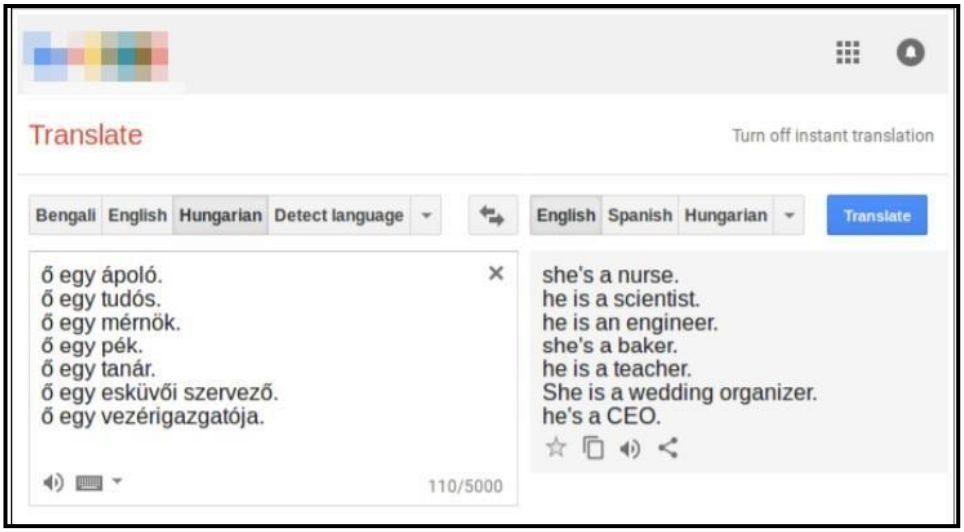
On assessing consumer's assessment of these products, the outcomes appeared quite striking. For craft beer, when consumers learnt that it was brewed by a woman, they remarked that they would make less payment for the beer. Also, regarding the taste and quality, the customers did not have high expectations. On the other hand, for the cupcakes, the difference was minor in consumers' attitudes towards producer's gender. From these observations, the researcher concluded that in traditional 'male' oriented markets, the goods made by women are subjected to penalization for no other reason except that these originated from a female.

Embry-Riddle Aeronautical University's research team worked on a design project [25] which involved studying the experiences of female pilots in safely fulfilling their flying missions. To facilitate continuous development of aviation industry, the need of diversity in aviation workforce has been recognized by Congressional leaders as per the 2005 report from United States Department of Education. Subsequently, as more females began to join this industry, the emerging concern was that the design of cockpits which typically catered to male physiology might not be accommodating the female physiology. Four subject matter experts, who were female pilots (with varying flying and aircraft experiences) were interviewed with the aim to optimize cockpit design for female pilots.

During the interview session, one pilot spoke about occasions when she had no choice except to sacrifice views either inside or outside of the aircraft which adversely affected her Situation Awareness (SA). Three pilots who were small in stature (height at or less than 5 feet 7 inches), expressed concerns regarding the distance between rudder pedals and seats. Each of these pilots had to adjust seat or rudder pedals to optime the information acquired inside or outside the aircraft as well as to reach other control inputs.

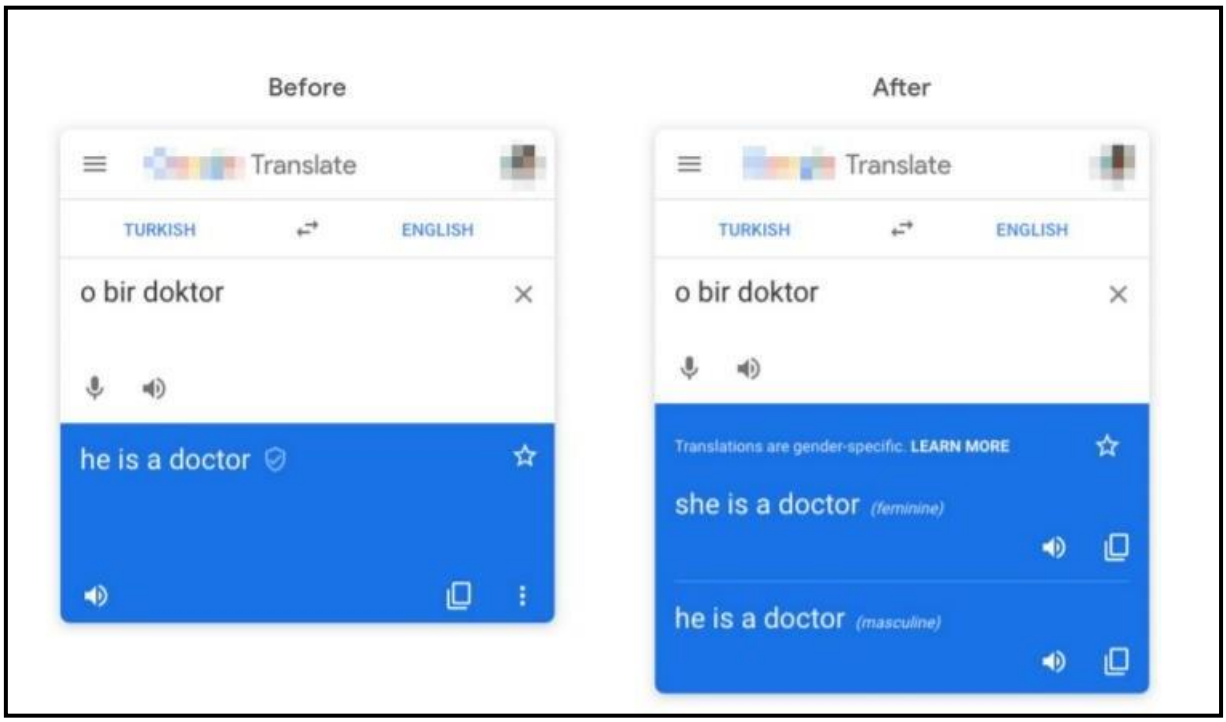
"Machine Bias" is the name given to phenomenon in which the trained statistical models begin to mirror the controversial societal assumptions especially gender and racial bias [26]. A significant majority of artificial intelligence tools are reported to be harmfully biased towards minority groups. An infamous incident when Apple's iPhone X failed to differentiate between two individuals of Asian ethnicity [27]. Another case of racial bias was observed with Google photos' image labelling algorithm which classified dark-skinned people as gorillas [28]. This 2018 research [26] provided evidence on gender bias in translation tools such as Google Translate. By translating sentences from gender neutral languages (Hungarian, Turkish etc.) into English, the translation tool produces outcomes which clearly indicates it being heavily biased with strong tendency towards male defaults. The dataset selected for this case study comprised of job positions to obtain a wider window into the nature of gender bias. Taking the example of a sentence in Hungarian language, "˝o egy ´apol´on˝" after translation to English yields "she is a nurse". The word apol´on˝ means nurse in English while "˝o" is a gender-neutral pronoun either he, she or it. Interestingly, the translation yields a male pronoun on replacing "nurse" with "engineer". The figure below is a screenshot of the translations of stereotypical gender roles from gender-neutral language, Hungarian to English.

Figure 2.7: Screenshot from Google Translate. Source [26]



When the research paper's pre-print version was released, it received an enormous media coverage. Thereby, in December 2018, the company's policies underwent modifications to feature masculine as well as feminine forms for Google's translations. The company released a statement where they acknowledged that the reason for taking this step is to promote fairness and make the machine learning translation tool debiased. Below is another screenshot of the GUI of Google Translate depicting the results before and after the release of new feature.

Figure 2.8: GUI of Google Translate before and after the release of new feature. Source [26]



2.4 Gender Effects in Human Robot Interaction (HRI)

The future predicted decades ago in relation to society's dependency on robots in domestic sphere might have or be on the verge of becoming true. During the situations such as worldwide pandemic outbreak, the entire world was shutdown with being enforced and confined to homes. At these times, the individuals who predominantly underwent a tide of social and emotional routine changes were senior citizens, children, university/college students, those living alone etc. With digitization and increased importance of internet, the international robotics engineers have been working to develop personal robots which are fully autonomous. These robots can act as personal assistants taking care of senior citizens by supporting everyday tasks, social robots in education sector as an alternative to traditional e-learning devices. With these advancements, it becomes crucial to extend the research in Human-Robot Interaction (HRI) to answer the questions such as:

1. What should be the appearance of personal robot assistants?
2. Should designers use and exploit gender stereotypes in the design of robots?
3. Should the robots be gender-neutral?

Eyssel and Hegel [29] through their experimental study investigated how the presence of visual cues affected the perception of anthropomorphic robots. In particular, they manipulated the hair length of robot and shape of lips. From the participants' responses, it was confirmed that the short- haired robot was perceived to be more masculine in comparison to the long-haired robot. Also, the data from their rating of each robot's traits from a list of 12 adjectives revealed that female robot was considered to be more communal whereas the male robot was perceived as more agentic. One drawback in this experiment was that it relied on stereotypical notions associated with gender roles for men and women. The respondents also rated on a 7-point Likert scale as to what extent they would utilize a specific robot for a typical male task vs female task. Examples of the tasks were, transporting goods, monitor technical devices, elderly care, prepare meals etc. Due to the pre-defined 'gender roles' into these categories (male vs female), the participants seem to have propagated the societal bias in their ratings. From their feedback, the research team noted that female robot was taken as more suitable for tasks related to household and care. Contrastingly, the male robot was stated to be more favorable for tasks related to repair, house guard. Another limitation in this study was its emphasis only on visual cues. Figures below show the headshots of male and female robot types which were presented on a computer screen.

Figure 2.9: Female Robot Type: long hair. Source [29]

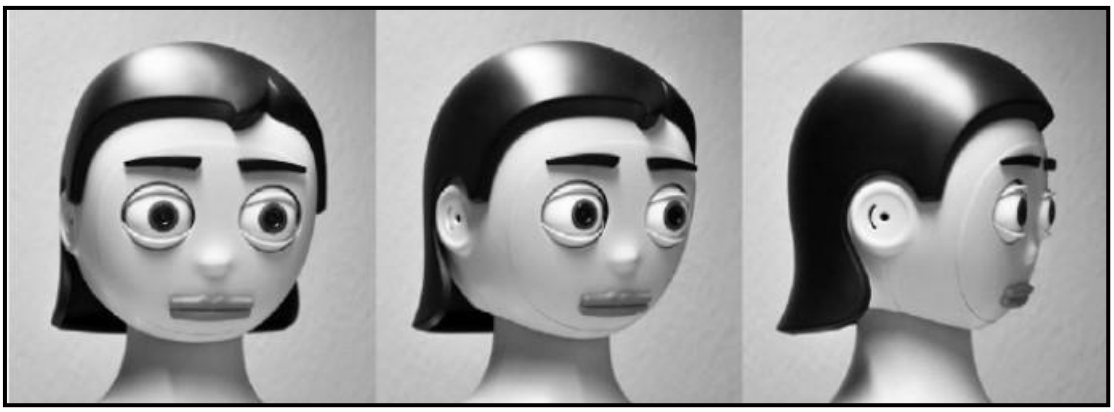
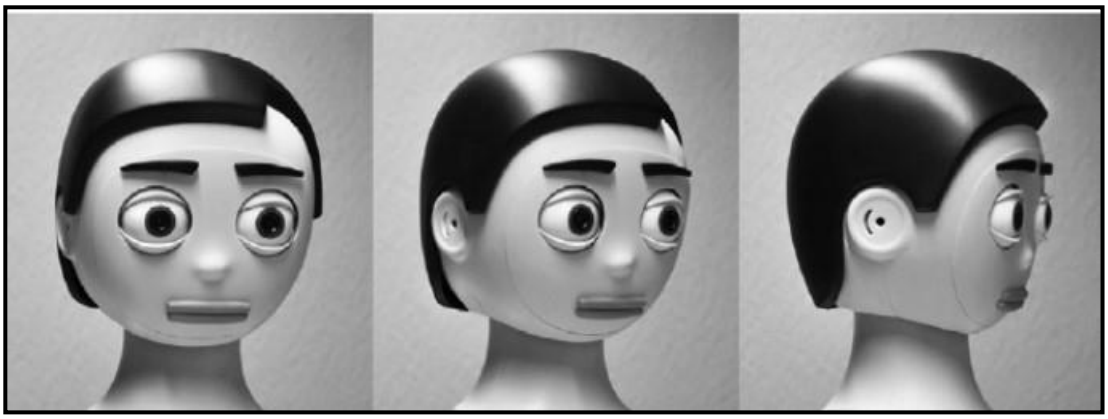


Figure 2.10: Male Robot Type: short hair. Source [29]



Human-Robot Interaction is not limited to adult only but also has been increasingly involved with children in the form of social robots for cognitive growth. This particular sector within HRI is known as child-Robot Interaction (cHRI). Sandygulova and O'Hare [32] conducted an experiment involving children aged between 5-12 years to identify whether a robot's synthesized voice evokes gender association in them. The research used a 58 cm tall NAO humanoid robot developed by Alderban robotics. Results revealed that the younger children preferred the robot of same gender as theirs whereas for a significant majority of older children, the gender of robot did not matter to them. Alderban robotics states that NAO and other company's robots are gender-neutral but refers to NAO as 'him' on their website.

Figure 2.11: Children interacting with NAO robot. Source [32]

Acknowledging the fact that a large proportion of Chinese population comprises of single women, the 2018 research [33] attempted to explore the design features of home robots to support single women in China. A thought-provoking section of results on the extent of humanoid/anthropomorphic features to be included in robots received both positive and negative comments from the participants. One of the respondents replied that since robots are not humans then why making them humanoid. One participant also expressed being offended on seeing all humanoid robots being portrayed with the imagery of female maids.

2.5 Gender Research in Website and Mobile App Development

Researchers Lopes and Vogel [36] from Trinity College, Dublin, through their experimental design attempted to explore any statistically significant difference in male and female participants' proposed requirements for a dating application. With the help of HCI research team, they developed a gender-neutral persona which served the purpose of application's potential user. The content and service of the dating app were presented as 'scenario'. 28 male and 28 female participants had to rate the personality traits that define the experience using dating application. From the results obtained, it was observed that male and female participants assigned male and female gender to gender-neutral persona. Women tended to assign the values of 'honesty' and 'honest personality' to the app, whereas men expressed to embed the app with features such as 'calm' communication and 'relationship mode'. 60 percent of the research participants said that it was difficult for them to establish requirements for an app to be used by someone else without considering their own preferences. Thus, the results supported the assumption that gender bias exists in the design process.

A final year project [37] submitted in Worcester Polytechnic University by one of its' undergraduate students was themed on designing gender neutral User Experience. The researcher aimed to introduce a website design which is pleasing to all users irrespective of gender. This gender-neutral UI was then tested against the other two designs, male biased and female biased. Researcher designed three versions of a book buying website: female biased with vibrant colors and emoticons, male biased with few colors and more professional language and a gender-neutral design containing nature imagery, rounder corners, and minimal language. Male and female participants performed same task across each version of the website which was to add books into cart and go to checkout page to review the cart and check out. The websites were rated on SUS (System Usability Scale) questionnaire after the participants completed task using each in the defined order. From the results, it became clear that female biased website was rated lowest by both the genders. Referring to the conclusions from Gloria Moss's work, the researcher commented that being a male he designed a website oriented towards female audience due to which females might not have rated it as highly as if it had been designed by a female.

Health apps are one of the most popular types of mobile apps used frequently by users of all genders. When Apple launched health-tracking app for its' devices [38], it allowed users to keep track of blood count, sugar count, step count etc. However, it did not get much appreciation from females as the app did include any feature such as period/ menstrual cycles tracking. This resulted in Apple re-designing the Apple health, thereby adding 'Reproductive Health' feature with iOS 9 update. With the recent updates, 'Reproductive Health' has been grouped under the term 'Cycle Tracking' enabling all the users who identify themselves as females to keep track of their cycles, get insights into the cycle's history, get updates about next cycle, keep record of symptoms along with fertility prediction. Johanna Levy from University of Granada, Spain conducted a qualitative study [39] to understand experiences and feedback of the users of menstrual tracking apps. The interviews conducted with 26 participants (aged above 18 years) brought to attention that some users were aware of gendered assumptions embedded in these apps' graphic designs and imagery. One of the participants changed the menstrual tracking app she had been using earlier due to the reason that it contained pink background with flowers. She personally did not like this aesthetic look of the app saying that these apps do not necessarily need to be pink just because the users are girls. Another respondent also moved towards a more gender-neutral menstrual app to show her support towards developers who have been designing apps for trans-people or menstruators who do not identify themselves as females.

Another category of apps which have been mainly marketed towards females are the pregnancy and parenthood apps. Research [40] depicts that comparatively fewer numbers of apps have been developed specifically for fathers or fathers-to-be. This study's primary focus was to do analysis of how fatherhood is portrayed in these apps. Findings revealed that the apps available in the market directed at women contained detailed information on fetal development while the apps available for fathers-to-be only rendered short information or quick tips. Moreover, the information for men also differed from that available in female oriented apps, for example, some offered the fathers advice on preparation of nursery and the best ways to manage finances on arrival of their baby.

Chapter 3: Methodology and Procedures

3.1 Deployment

The initial plan was to have a male and a female participant providing UI designs of a typical 'masculine' and 'feminine' themed apps. Both the apps would then be subject to usability testing by male and female respondents. After contemplating on this idea, it was realized that by doing so gender stereotypes in designs would be propagated by pre-defining the themes to be masculine or feminine. User of any gender can identify themselves with the theme of an app, hence labelling an application to be designed for 'men' or 'women' would have been a source of bias in the first place itself. This can be better understood by considering that a female can have interest in the games or apps which were traditionally used by men dominantly; similarly feminine or feminist do not only include or refer to those born female, but a man can be also feminist too and have taste towards feminine aesthetics. Designers too could introduce conscious or unconscious bias while adding visible attributes of color, images, and content.

In today's rapidly evolving society where acceptance and awareness of products, designs, self-identifying pronouns and even jobs have emerged on the greater side of gender-neutral spectrum, it would not be justified to 'label' an application's theme based on traditional/stereotypical gender roles. Hence, the second idea which eventually turned into final decision was to let a male and a female participant design the interface of 'Tourism' themed mobile app providing details on five tourist places in their respective hometown with the options to plan a virtual trip and save places for later visit. The selected theme being gender neutral was considered suitable for the current research.

3.2 Use Cases

After ruling out gender specific mobile app themes (as stated in 3.1), 'Tourism' theme was finalized for the study. A male and a female designer provided their individual UI designs based on five tourist destinations in their hometown. Travelling apps are said to be equally popular among users of both the genders nowadays. User flow diagrams and demonstration about the theme were provided to both the participants. As a part of pilot-testing, wireframes were obtained from the designers to determine if they possessed similar or nearly similar proficiency in UI designing. Developing working versions of both the apps was the researcher's task.

After the design and development (coding) of both the apps were completed and pilot- tested, the following task flow was adapted for the usability testing phase:

1. Download the first app (order pre-determined by researcher).
2. Sign up and login
3. Navigate through each screen of the app to get familiar with the interface. Select any 2 tourist places, see their details and plan on calendar for visiting.
4. Select another 2 places and save for later.
5. Logout from the app.
6. Rate the experience on SUS questionnaire.
7. Repeat the procedure for second app.
8. Rate the second app on SUS questionnaire.
9. Answer the post-experiment questions.

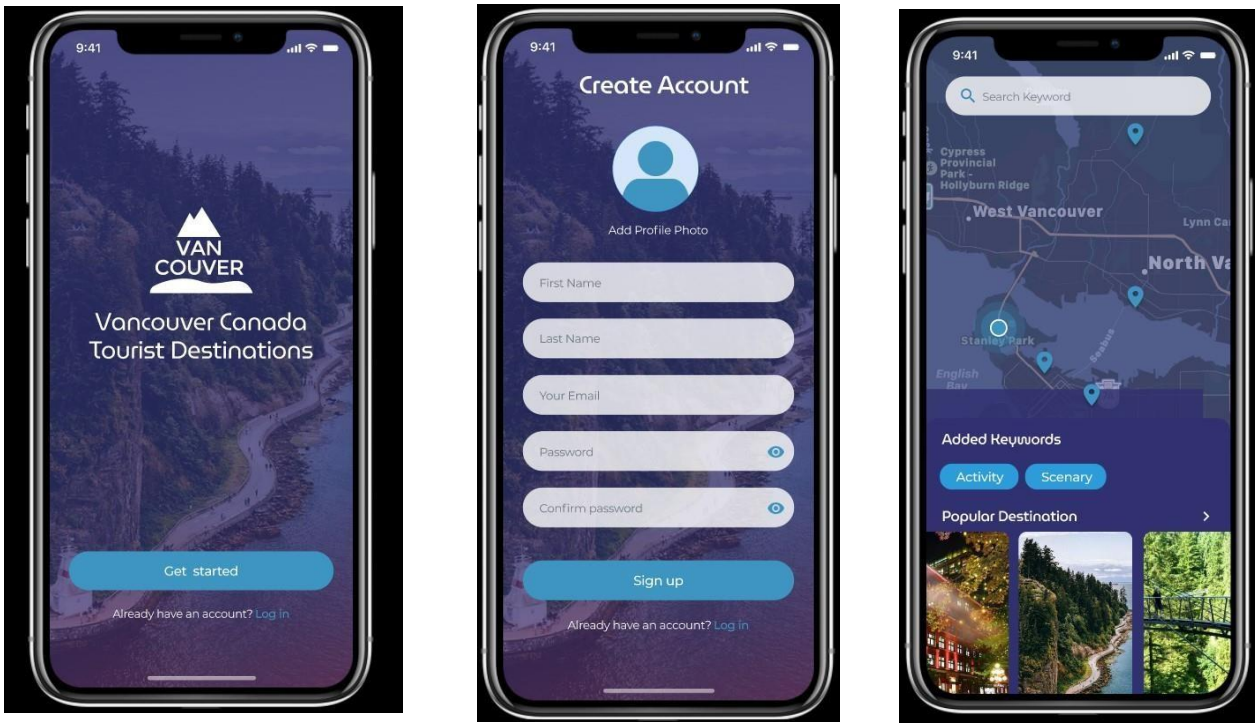
3.3 Experimental Flow

The experiment is divided into three stages involving task for participants recruited for UI designing, researcher and testers. The role of each has been explained below:

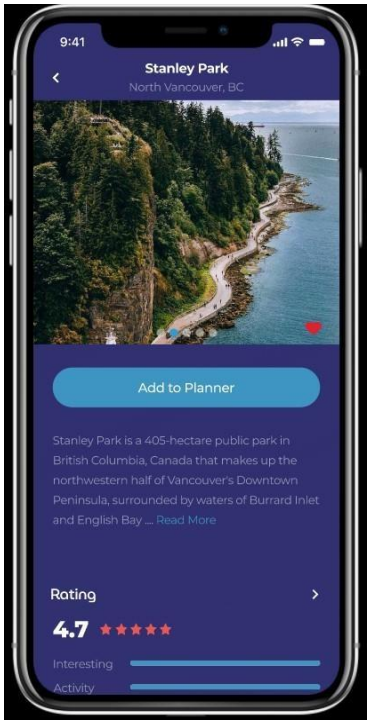
3.3.1 Designing

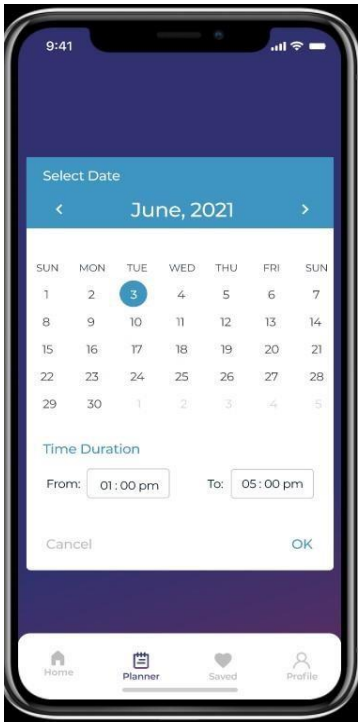
Using no-code/low-code platform 'Figma' both the designers (a male and a female) created UI designs of tourism themed mobile apps individually. The selection of colors, features, layout were their personal choice, however, mandatory features to be included in the app such as 'add to calendar/plan on calendar' and 'save for later' were laid out in the user flow diagrams. The UI design comprised of 5 tourist places in the respective hometowns of the designers on home screen, their corresponding details screen, separate screens for calendar and saved places. Also included was the user profile screen which displays information of the signed user such as name, email etc. The time allotted to each designer for submitting their designs was 1 week. The male designer's app was based on five tourist places in New Delhi, India while the female designer's app was centered on Vancouver tourist places.

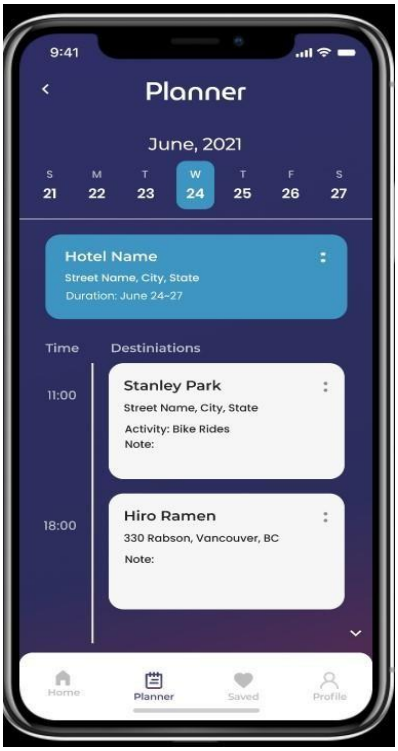
Figure 3.1: UI Design Submitted by Female Designer




Stanley Park
Vancouver, BC
Capilano Sus
North Vancouver,
Hotel Best Deals
Fairmont Hotel
$ 147.00 /per night
Pinnacle Hotel
$ 107.00 /per night
Sheraton Vancouver
$ 106.00 /per night
Georgian Court Hotel
$ 99.00 / per night
Seaside Hoteal
$ 105.00 /per night
Home
Planner
Saved
Profile


9:41
Stanley Park
North Vancouver, BC
Add to Planner
Stanley Park is a 405-hectare public park in British Columbia, Canada that makes up the northwestern half of Vancouver's Downtown Peninsula, surrounded by waters of Burrard Inlet and English Bay ... Read More
Rating
4.7
Interesting
Activity


9:41
Select Date
June, 2021
SUN MON TUE WED THU FRI SUN
Time Duration
From: 01 : 00 pm
To: 05 : 00 pm
Cancel
OK
Home
Planner
Saved
Profile


9:41
Planner
June, 2021
S M T W T F S
21 22 23 24 25 26 27
Hotel Name
Street Name, City, State
Duration: June 24-27
Time
Destiniations
11:00
Stanley Park
Street Name, City, State
Activity: Bike Rides
Note:
18:00
Hiro Ramen
330 Rabson, Vancouver, BC
Note:
Home
Planner
Saved
Profile


9:41
Saved
Stanley Park
Lion's Gate
Stanley Park
Bike Rental
Stanley Park
Vancouver Aquarium
Deep Cove Park
Kayaking

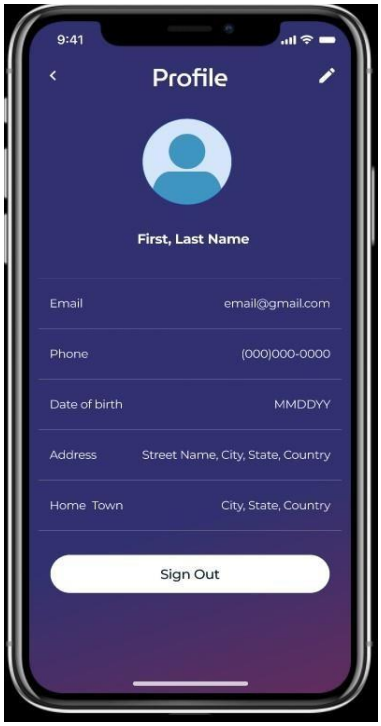
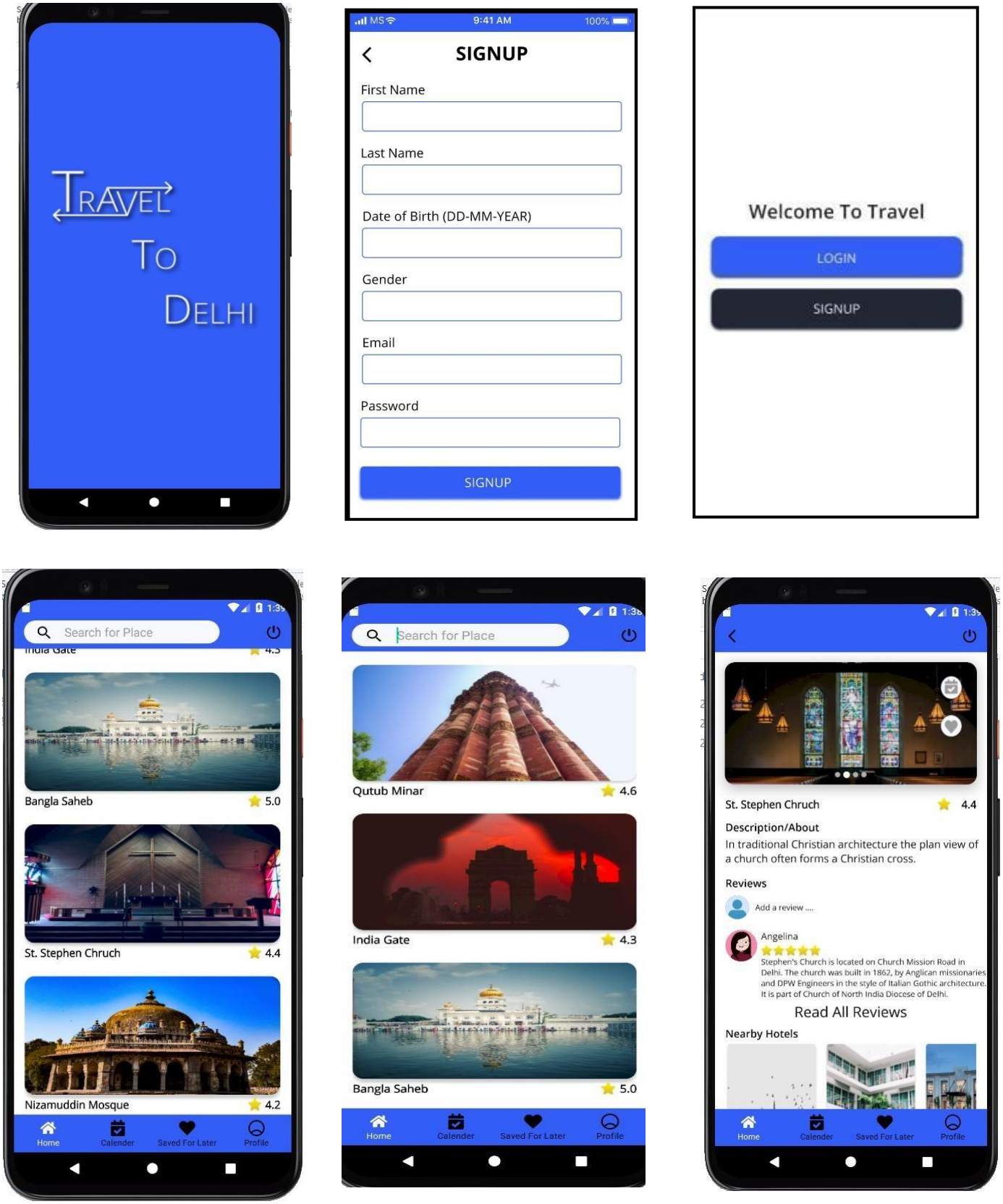
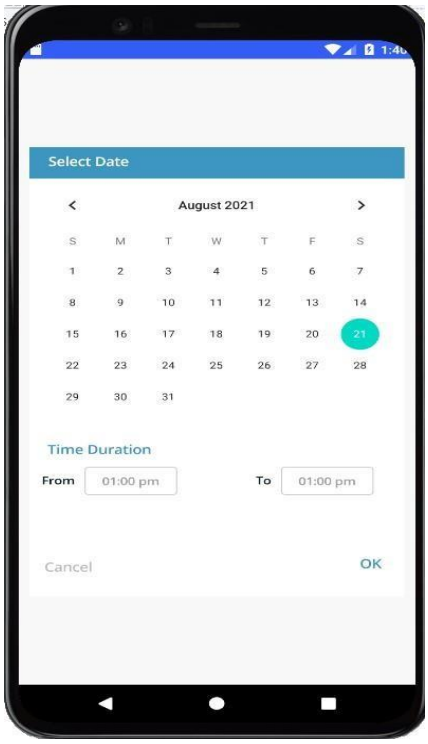
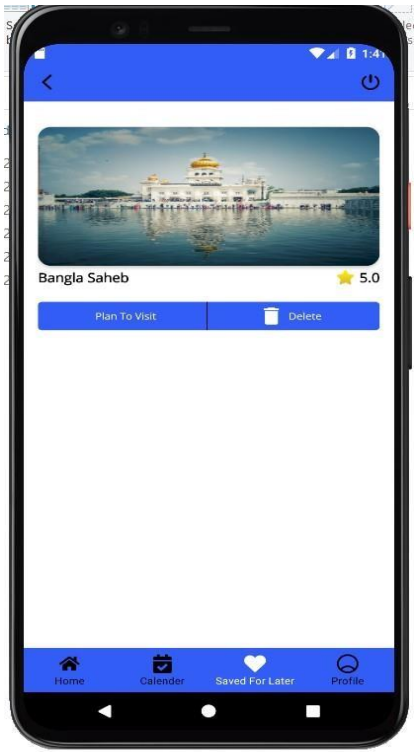

9:41
Profile
First, Last Name
Email
email@gmail.com
Phone
(000)000-0000
Date of birth
MMDDYY
Address
Street Name, City, State, Country
Home Town
City, State, Country
Sign Out

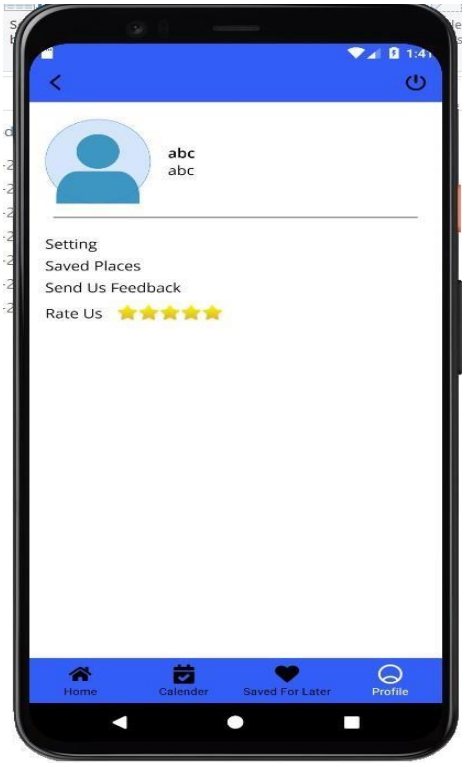
Figure 3.2: UI Design Submitted by Male Designer




St. Stephen Chruch
4.4
18-8-2021 3:40
Save For Later
Delete
+ Add new trip
Home
Calender
Saved For Later
Profile


Select Date
August 2021
S M T W T F S
Time Duration
From 01:00 pm
To 01:00 pm
Cancel
OK


Bangla Saheb
5.0
Plan To Visit
Delete
Home
Calender
Saved For Later
Profile


abc
abc
Setting
Saved Places
Send Us Feedback
Rate Us
Home
Calender
Saved For Later
Profile

An e-gift card worth Canadian dollars 50 was given to the female designer. The equivalent electronic money was transferred to the male participant in Indian Rupees in accordance with their preference.

3.3.2 Development

Upon receiving the UI designs, we converted these into functioning Android mobile applications. Coding has been done in JAVA using Android Studio. For the backend, 'Firebase' database is used. The working versions of both the applications were shown to their respective UI designers for obtaining a confirmation that the apps' designs reflected the original designs submitted by them. We conducted the pilot-testing of both the apps with 2 males and 2 females. Any technical limitations observed were noted and rectified before proceeding to the testing phase.

3.3.3 Testing

Thirty participants (15 males and 15 females) volunteered to participate in remote usability testing of mobile apps. All the participants owned an Android smartphone with O.S.7 or above. Approximately, a 30–40-minute zoom session was conducted with each participant. The session began with participants signing the consent form, answering demographic questionnaire, completing the assigned task in one app followed by answering SUS questions and then repeating the same procedure for second app.

The task was to first navigate through each screen of the app to get familiar with the interface. The participants then had to choose 2 places, read their details from the 'details screen' and plan on calendar for visiting assuming themselves to be an actual traveler. Participants then had to choose another 2 places and save these places for later visiting. These places would then show up in the screen called 'Saved Places'. The candidates had to look out for the icons or buttons designed to plan and save the places. As the participants completed the task, they would logout by locating the 'logout' button. However, if the task was left incomplete the logout button was coded not to let them come out of the app and display an alert message saying, for example, "*plan at least 2 places on calendar*". In this scenario, the participants would re-check the number of places they planned and saved by navigating to the corresponding screens displaying the places planned on calendar and saved

for later. On fulfilling the task criteria, the participants had to exit from the app by clicking on logout button. After task completion, participants rated the app on SUS questionnaire. The same procedure was then followed for second application. Once the tasks were completed in both the apps and their respective SUS questionnaires responded, post-experiment questions were shared with the participants. Testing was done over a period of 1 month in July 2021.

3.4 Participants Recruitment

For the designing task, the recruitment poster (attached in Appendix A) was shared on Facebook. The criteria followed for selecting participants was that both (a male and a female) should have nearly same experience in UI designing and/or from similar academic backgrounds. A male designer from South Asia and a female designer from North America were eventually selected for the UI designing task. Both the participants had minimum 5 months experience in UI designing as well as using Figma. Male participant, from Computer Engineering background, had been pursuing Google's professional UI designing certification program while the female participant, from interior designing background, had enrolled in UI design bootcamp. The wireframe designs obtained from them, as a part of pilot-testing, indicated their similar level of knowledge in designing prototypes.

Recruitment poster for testing was also shared on Facebook. 15 males and 15 females volunteered to participate following the criteria that they owned an Android smartphone with operating system either 7 or higher. The volunteers were distributed across various countries such as: Canada, United States, Turkey, Iran, India, and Pakistan. Among these 30 participants, there were also students of Laurentian University. By profession, the respondents mainly comprised of software developer, product designer, UI designer, architect, students, and professors. 16 participants belonged to the age group of 18-25 years, 12 belonged to 26-39 years age group while 2 were above 40.

3.5 Questionnaires

Following set of questionnaires (attached in Appendix C) are used in the study:

1. Demographic questionnaire for designers and testers.
2. System Usability Scale questionnaire.
3. Post-experiment questions.

Demographic questionnaire is designed to gather information about participant's age, gender, UI designing experience etc.

System Usability Scale (SUS) questionnaire is most commonly used in studies involving usability testing. It comprises of 10 questions which are rated on a scale of 1-5 where, 1: strongly disagree and 5: strongly agree. The responses collected are then converted to a score out of 100.

Original SUS Questionnaire:

1. I think I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Modified SUS Questionnaire:

1. I would recommend this application to a tourist travelling to Delhi/Vancouver.
2. I found this application to be more complicated than it should be.
3. I think planning a trip using this app is simple and easy.
4. While completing the assigned task, I needed support from researcher.
5. I found various functions in the app were well integrated and worked smoothly.
6. I think there are certain irregularities/inconsistencies in this app.
7. I would imagine that most users would learn to use this app very quickly.
8. Finding information or navigating in this app was cumbersome.
9. I felt confident using this app to complete the assigned task.
10. I think there are lot of things in this app that require improvement.

Questions included in the post-experiment questionnaire attempted to draw participants' observations and opinions on the gender aspects of the experiment. Four subjective questions are framed in a way that reveal whether participants noticed or observed the UI designs from gendered perspective. For example, the first question asks if they perceived the colors, layout and/or content in either of the two apps to be more masculine, feminine or gender neutral. Though participants might not have rated the SUS questionnaires based on these factors, their responses to these questions signify whether in general they attribute 'gender' to the aesthetic appearance of an interface. Questions have been formulated in a manner to identify any gender-stereotypes being prevalent in the participants' reasoning during their impression of a mobile interface.

3.6 Procedure

The designers were first asked for their consent to participate by signing the consent form (provided in Appendix B). This was followed by them answering the demographic questions. Demonstration regarding the details about UI designs was shared with them in their e-mails. Both the candidates were informed that the time limit to submit their individual designs was one week. Once the designs were received, a reward worth CAD 50 was shared with each.

Similarly, during testing, consent forms were signed by each candidate. Zoom sessions were held with one participant at a time. After signing consent forms, they answered demographic questions. We then shared the details on what procedure had to be followed during the entire session along with the task to be completed using each app. APK file of one app was shared with each test participant. As soon as they signed up and logged in, the researcher would no longer assist them in the task. Participants had to figure out themselves which button or icon had to be clicked to successfully complete the assigned task. This was followed by SUS questionnaire being rated for the app shared first. As soon as we received their responses, the same procedure was then followed for the second app. Half of the male and female participants tested the 'Vancouver' themed app first while the remaining half tested the 'New Delhi' themed app. Post-experiment questions were answered by the respondents after testing of both the applications was done successfully.

3.7 Hardware and Software Requirements

The hardware requirements imply the hardware needed on the testers' side for testing:

1. Android Smartphone with API 7 or above
2. 4 GB RAM
3. 2 GB of available disk space (minimum)
4. High speed Internet connection

The technical specification of software requirements on the experimenter's side include:

1. Any Operating system (Windows, Linux, MAC)
2. Android Studio
3. Android SDK (Android Software Development Kit)
4. Android API 5.1 or above
5. AVD (Android Virtual Device) Plugin for testing the applications on desktop.

3.8 Debriefing

Designers were debriefed about the gender aspects of the experimental study and the motive behind collecting designs from a male and a female designer. Both the designing candidates signed the debriefing form after reading the details. Any questions from their side were welcomed and answered. Debriefing forms are attached in Appendix D.

Similarly, testing participants were also debriefed at the end of the experiment about the gendered nature of the research. Prior to commencing the task, they were not informed of the gender of app designers so as not to propagate any bias during the time they tested and rated. Debriefing form was shared with them which laid out details on why deception was involved. The participants were informed of the gender of app designers and the way their responses will be used during results analysis. Participants signed the form after having read it carefully.

3.9 Ethics Approval

Since the experimental study involved human participants, an application for 'Research Involving Human Subjects' was submitted to Research Ethics Board at Laurentian University in March 2021. After necessary revisions, the application got the approval in April 2021. The approved Ethics certificate is provided in Appendix C.

3.10 Data Collection

The collection of questionnaire data has been done using REDCap, a secured data storage approved by Laurentian University. Forms and questionnaires are designed in REDCap and are attached in Appendix C and D. Data is securely stored in REDCap with password protection. Responses to SUS questionnaires, demographic questionnaire and post-experiment questions are stored anonymously.

3.11 Data Analysis

To calculate the final SUS score, the following method is applied:

1. Convert the scale into a number for each question

Strongly Disagree: 1point

Disagree:2 points

Neutral: 3 points

Agree: 4 points

Strongly Agree: 5 points

2. Calculate:

X = Sum of the points for all odd-numbered questions - 5

Y = 25 - Sum of the points for all even numbered questions

Final SUS Score = (X+Y) X 2.5

The total score is 100 and each question has a weight of 10.

Rationale behind odd-even questions selection: All the odd-numbered questions being in positive tone, if the response to these is 'strongly agree' then we must give these maximum points which are 10. In case, the response is 'strongly disagree', we must give minimum point which is 0. By subtracting 1 from each odd numbered question, we are ensuring that the minimum is 0. On multiplication by 2.5, we ensure that maximum is 10. On the other hand, all the even numbered questions are in negative tone. Thus, if the response is 'strongly agree', we wish to give them a minimum score of 0. By subtracting all the even numbered questions from 5, we ensure that the minimum score remains 0.

For data analysis we have made use of R language. First, we have performed 'descriptive statistics' to provide a highlight on the characteristics of data by taking these study variables,

FAPP.F: Female designed app evaluated by females

FAPP.M: Female designed app evaluated by males

MAPP.F: Male designed app evaluated by females

MAPP.M: Male designed app evaluated by males

Next, we checked whether our data followed normal distribution. This is followed by test of hypothesis:

1. Hypothesis, as indicated in 1.3, have been tested by applying paired sample t-test along with its non-parametric version, Wilcoxin test.

2. Additional computations have also been done on the Likert Scale data. Fixing the gender of app designer, we have applied tests of statistical significance across Likert scale ratings given by male respondents versus female respondents (it should be noted that primary hypothesis involved fixing 'gender' of the respondents). Since the sample in this case is different with respect to each other that is, male respondents' ratings are compared across female respondents' ratings, we have applied unpaired t-test and its non-parametric version: Mann-Whitney Test.

3. Finally, to identify any statistically significant interaction between apps and gender of testers, ANOVA test has been applied.

Chapter 4: Results

4.1 Descriptive Statistics

For a given variable, Standard deviation (SD) is known to be a plausible measure of variability in the data set on the variable. In fact, the range Mean ± Standard Deviation in the values of a parameter implies that nearly two-third of the observations would expectantly lie within this range.

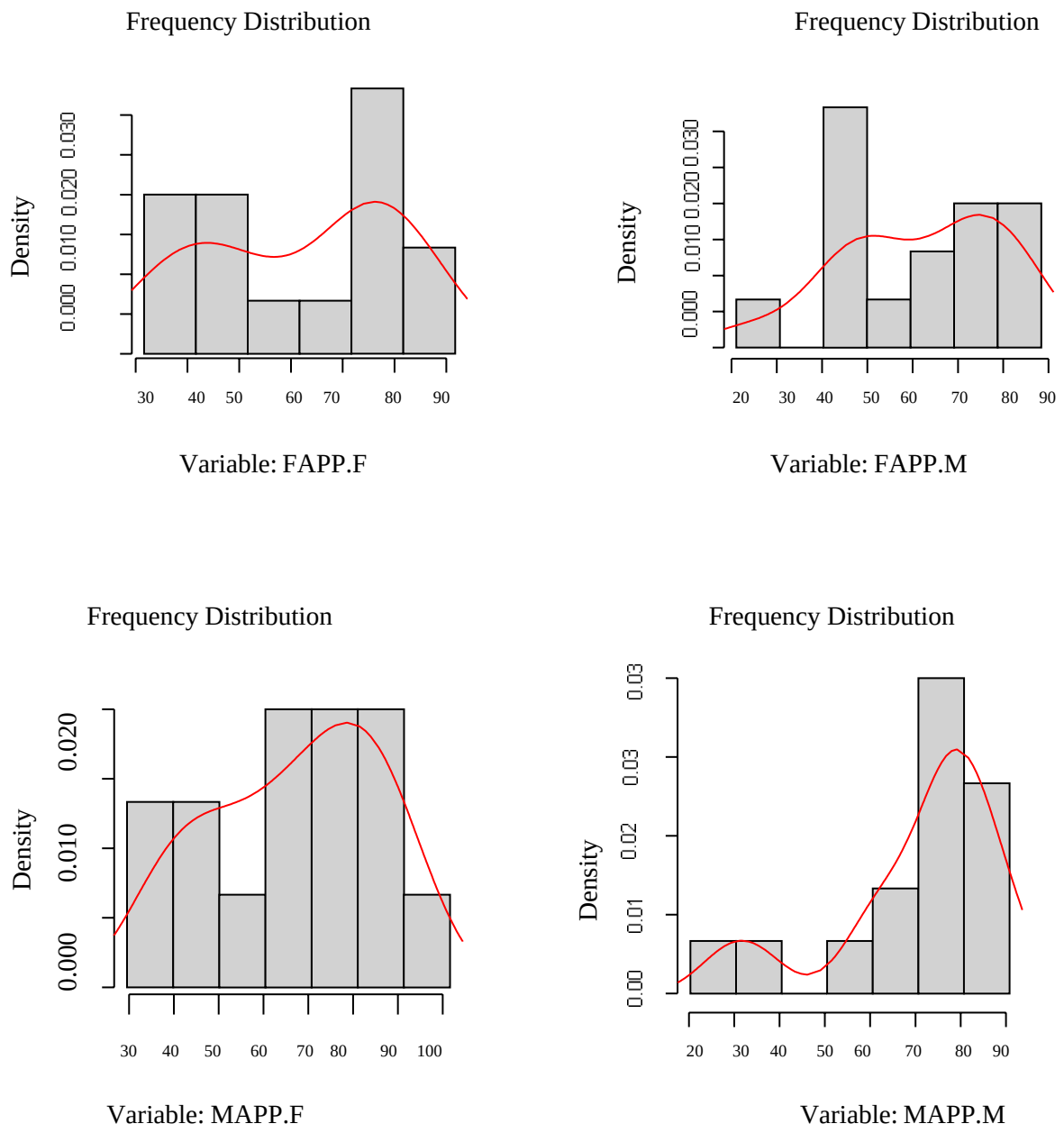
Ninety-five per cent confidence interval for the sample mean implies that if the experiment were repeated under the same set of conditions over and over again, then in 95 percent of the repetitions, sample mean would expectantly lie somewhere within this range.

The quantity γ_1 provides us with a measure of skewness (i.e., lack of symmetry) in the data set, whereas the quantity γ_2 provides us with a measure of kurtosis (i.e., relative concentration of the distribution in the middle) in the set.

Table 4.1: Descriptive Statistics in respect of the SUS Data on the Study Variables (N= 15)

Variable#	Mean ± SD	95% Confidence Interval	γ_1	γ_2
FAPP.F	60.50 ± 18.35	50.34 – 70.66	-0.25	-1.40
FAPP.M	62.17 ± 18.16	52.11 – 72.22	-0.46	-0.72
MAPP.F	66.50 ± 17.48	56.82 – 76.18	-0.29	-1.17
MAPP.M	69.83 ± 17.43	60.18 – 79.48	-1.27	0.65

Figure 4.1: Density Plots



4.2 Tests for Normality

As a pre-requisite for testing gender difference, if any, among the study variables, we have examined normality pattern in the data set. For this purpose, we have adopted Shapiro- Wilk's Test, the results from which have been presented in table below:

TABLE 4.2: Testing Normality in the Sample Data on Different Study Variables (in respect of Both Genders)

H0: Pattern in a given set of data is normal
H1: Pattern in the data set is non-normal

APP	Male Respondents			Female Respondents		
	Test Statistic (T)	p-Value	Remark	Test Statistic (T)	p-Value	Remark
Male	0.847	0.0158	*	0.938	0.3602	NS
Female	0.937	0.3447	NS	0.918	0.1779	NS

*: Significant at 5% probability level; NS: Non-significant.

Main implications from Table 4.2:

In respect of the Male App, value of the Shapiro-Wilk's test statistic computed from the data compiled on male respondents was 0.847. The statistics were associated with a fairly low value (= 0.0158) of p. Since the p-value was less than the critical limit of 0.05 (but more than 0.01); therefore, we were led to reject (at 5 per cent probability level) the null hypothesis of normality pattern of the sample data on SUS for the male app, as evaluated by male respondents.

However, value of the test statistic for the App as evaluated by female respondents turned out to be 0.938. Since p-value of the test statistic (= 0.3602) was larger than the critical limit of 0.05; therefore, the assumption of normality in the data set could not be rejected. In other words, we could say that the sampled data on Male App, as evaluated by female respondents followed a normal distribution.

On similar lines, the null hypothesis of normality pattern could not be rejected in respect of the Female App, as evaluated by both Male (T = 0.937; p = 0.3447) and Female (T = 0.918; p = 0.1779) respondents. Thus, in general, the pattern followed by the experimental data on Male and Female respondents on the two Apps followed a normality pattern. Data on the responses obtained from males regarding Male App was the lone exception wherein the normality pattern was detected to be at a doubt.

4.3 Tests of Hypothesis

Prevalence of normality pattern in the data sets suggested in favor of the application of t-test for testing statistical significance of gender difference between mean values of the study variables. However, we have opted for the application of Mann-Whitney U-test/ Wilcoxon's Signed Rank test along with unpaired/ paired t-test, to have a more robust picture about the comparisons. Findings from the application of the test procedures are given in Tables 4 .3and 4.4, as follows:

Null Hypotheses (H0):

1. For Male Testers: No statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.
2. For Female Testers: No statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

Alternate Hypotheses (H1):

1. For Male Testers: Statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.
2. For Female Testers: Statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

Before proceeding, we may clarify hereby that for evaluating the two Apps, we have used the same set of testers (either Males or Females). Accordingly, we have adopted ¬**Paired t- test and Wilcoxon's Signed rank test** (both of which are used for related samples).

TABLE 4.3: Results from the Application of Paired t-test and Wilcoxon's Rank Sum test as applied to the Gender-wise Values of the Two Apps (n1 = 15; n2 = 15)

Tester	App	Mean	SE_m	Mean Difference	Paired t-test		Wilcoxon's test	
					Statistic	p-value	Statistic	p-value
Male	Male	69.83	4.50	7.67	1.758^{NS}	0.1006	85.00^{NS}	0.1632
	Female	62.17	4.69					
Female	Male	66.50	4.51	6.00	1.102^{NS}	0.2889	59.50^{NS}	0.3444
	Female	60.50	4.74					

[#] At 14 *degrees of freedom;* [NS]: Non-significant

Main implications from Table 4. 3:

In respect of the Male testers, mean values of the SUS scores for the Male and Female Apps were 69.83 and 62.17, associated respectively with the values for standard errors of mean equaling 4.50 and 4.69. Thus, there was a difference of 7.67 between the two mean values. For the observed difference, value of the paired t-statistic (at 14 degrees of freedom) was computed to be 1.758, which was associated with a p-value of 0.1006. And value of Wilcoxon's test statistic was computed to be 85.00, which was associated with a p-value of 0.1632. Since, for both the test statistics, values of p were higher than the critical limit of 0.05; therefore, the observed difference between the two Apps in respect of Male testers was statistically non-significant.

On similar lines, the observed difference (= 6.00) between the mean SUS scores of the two Apps for Female testers, too, was statistically non-significant (p-values for the paired t and Wilcoxon's test statistics being 0.2889 and 0.3444, respectively)

Accordingly, the null hypothesis could not be rejected.

We shall now proceed on similar lines to test the equality of mean scores for Male and Female responses for each of the two Apps. Here, we may clarify that for a given APP, since the respondents (Males & Females) were different; therefore, we have made use of ***unpaired t-test*** **and** ***Mann-Whitney's U-test*** to provide answer to the question (Table 4.4).

Null Hypotheses (H_0):

1. For Male Designer App: No statistically significant difference exists in the mean SUS ratings given for Male Testers
   vis-à-vis Female testers.

2. For Feale Designer App: No statistically significant difference exists in the mean SUS ratings given for Male Testers
   vis-à-vis Female testers.

Alternate Hypotheses (H1):

1. For Male Designer App: Statistically significant difference exists in the mean SUS ratings given for Male Testers vis-à-vis Female testers.

2. For Female Designer App: Statistically significant difference exists in the mean SUS ratings given for Male Testers vis-à-vis Female testers.

Table 4.4: Results from the Application of Unpaired t-test and Mann-Whitney's U-test applied to the Gender-wise Mean Values of the Two Apps (n1 = 15; n2 = 15)

Design of the App	Tester	Mean	SE_m	Mean Difference	Unpaired t-test		M-W test	
					Statistic#	p-value	Statistic	p-value
Male	Male	69.83	4.50	3.33	0.505^{NS}	0.6174	123.50^{NS}	0.6624
	Female	66.50	4.51					
Female	Male	62.17	4.69	1.67	0.242^{NS}	0.8109	121.00^{NS}	0.7392
	Female	60.50	4.74					

Main Conclusions from the above Table:

For each of the two Apps, no statistical significant differences could be detected (through each of the two test procedures – Unpaired t-test and Mann-Whitney's test) between the mean SUS scores of male and female testers.

Furthermore, to study the presence (or otherwise) of the interaction effects between design of the App and gender of the testers, we have made use of **Two-way Analysis of Variance** technique. Design of the App was taken to represent the first factor, while gender of the tester as the second factor. Analytical findings from the technique have been given in brief as follows:

Null Hypotheses (H0):

3. No statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

4. No statistically significant difference exists in the mean SUS ratings given for Male Testers vis-à- vis Female testers.

5. No statistically significant interaction exists between Apps and Gender of the Testers

Alternative Hypotheses (H1):

5. Statistically significant difference exists in the mean SUS ratings given for male designer's app vis-à-vis female designers' app.

6. Statistically significant difference exists in the mean SUS ratings given for Male Testers vis-à-vis Female testers.

7. Statistically significant interaction exists between Apps and Gender of the Testers

TABLE 4.5: Mean Values of SUS Scores of the Two Apps as Evaluated by Male and Female Testers

Design of the App	Gender of the Tester		Mean
	Male	Female	
Male	69.83	66.50	68.17
Female	62.17	60.50	61.33
Mean	66.00	63.50	64.75

Interpretation of mean scores:

Each male and female tester rated the SUS questionnaires designed for both the apps. The data was further segregated according to type of app and gender of testers, which means that the Likert scale data was then split up into following categories:

1. SUS scores given for MAPP by female testers
2. SUS scores given for MAPP by male testers
3. SUS scores given for FAPP by female testers
4. SUS scores given for FAPP by male testers.

The formula mentioned in Section 3.11 was then applied on each category of data mentioned above to calculate "Final SUS" scores. The mean from each "Final SUS Scores" table was calculated, resulting in four mean SUS values. Thus, mean SUS scores is the average of 15 "Final SUS" scores calculated for female testers for both the categories of apps. Similarly, from 15 "Final SUS" scores calculated for MAPP and FAPP for male testers, two mean values were calculated respectively. These mean values indicate the average usability performance of MAPP and FAPP as perceived by male and female testers while performing the task.

From the interpretation based on the studies [34] [35], the SUS scores are associated with an adjective rating. According to this rating, the acceptable average SUS score is 68 which is considered as 'good' and any SUS score below 50 implies serious usability issues in the product/software. The mean SUS scores shown in Table 4.5 above, are in the range of 60-70 which means that these lie close to average SUS score of 68, that is, 50th percentile;

hence, they are acceptable. However, identifying whether the mean SUS scores obtained in this research were acceptable or not, is not the objective of this research; the objective being to find if male and female tester's mean SUS scores had any statistically significant difference.

TABLE 4.6: Two-way Analysis of Variance (ANOVA) Table for Testing Statistical Significance of Difference between Designs of the Apps, Gender of the Testers and Their Interaction

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F-Value	p-Value	Remark
Design of the App	1	700.42	700.42	2.049	0.1578	NS
Gender of the Tester	1	93.75	93.75	0.274	0.6025	NS
Design × Gender Interaction	1	10.42	10.42	0.030	0.8620	NS
Residuals	56	19141.67	341.82	-	-	-
Total	59	19946.25	-	-	-	-

Main Conclusions from Table 4.6:

For testing the statistical significance of difference between the two Apps, value of the F-statistic (at 1 and 56 degrees of freedom) was 2.049. The F-value was statistically non-significant because its p-value (= 0.1578) was more than the critical value of 0.05. Thus, in respect of Design of the Apps, the null hypothesis could not be rejected, thereby implying that there did not exist any statistically significant difference between the mean SUS score of the two Apps.

On similar lines, the null hypothesis in respect of gender of the testers, too, could not be rejected (F-value = 0.274; its p-value = 0.6025), thus implying that there did not exist any statistically significant difference between the mean SUS score of the two genders.

Furthermore, the null hypothesis for the interaction effect also could not be rejected (F-value = 0.030; p-value = 0.8620). This means that there did not exist any interaction effect between design of the apps and gender of the testers. In other words, relative comparison of the two apps were similar whether these were evaluated by male or by female testers, and vice versa.

For the sake of convenience, mean values of SUS Scores of the two Apps, as evaluated by Male and Female testers, have been displayed in Figures below:

Figure 4.2: Grouped Bar Chart for Gender Within the Apps in Respect of SUS Scores

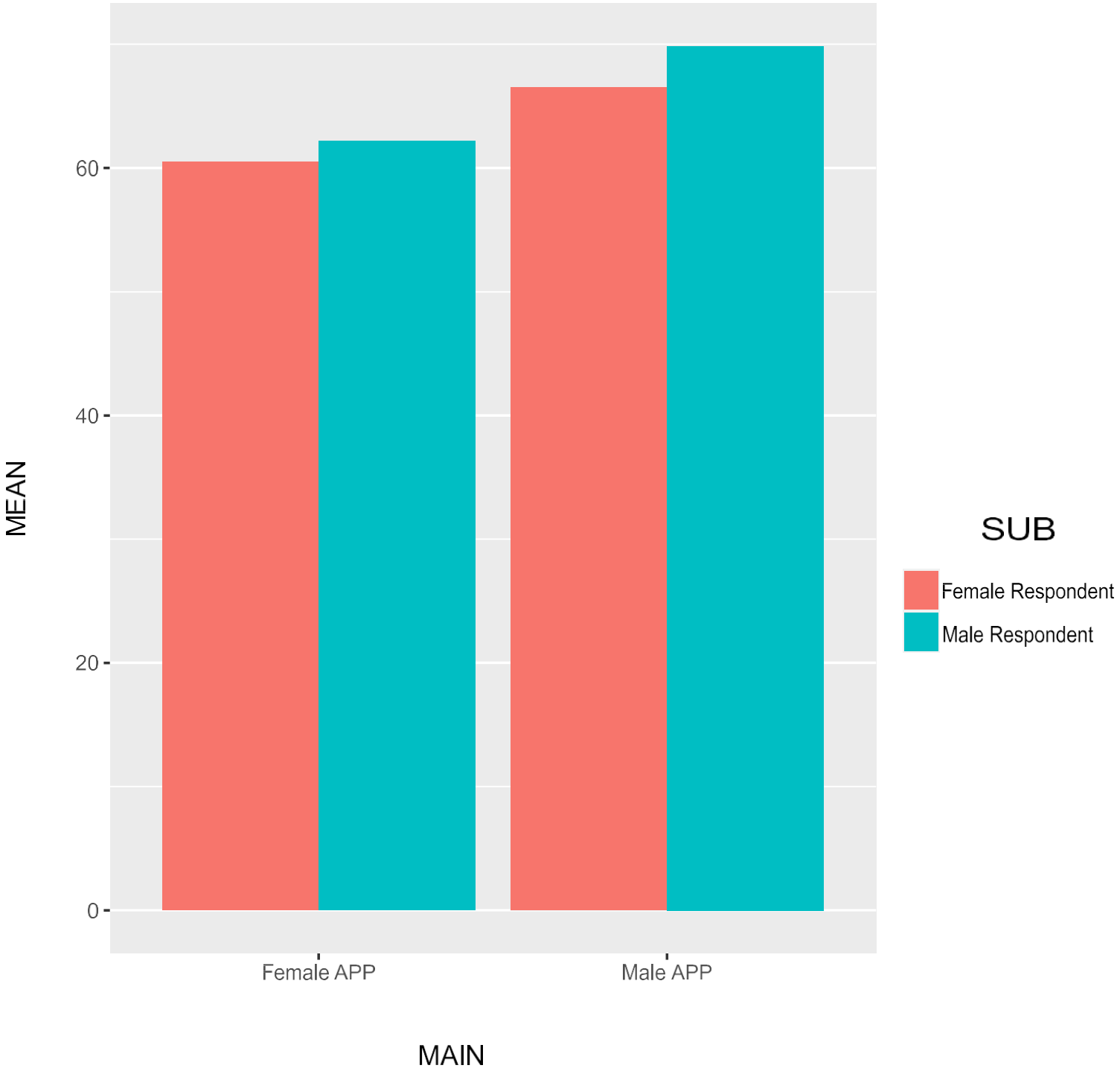
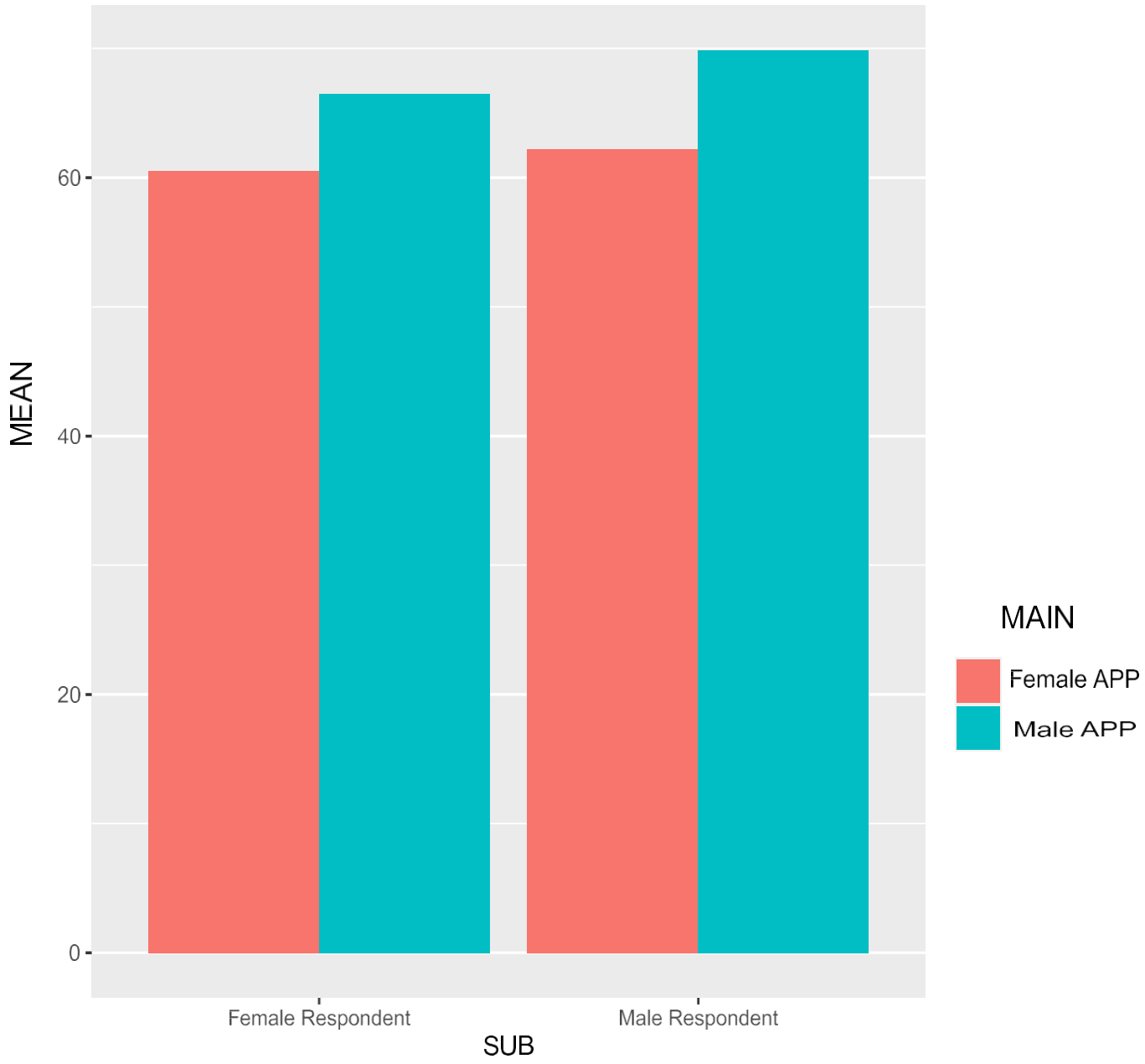


Figure 4.3: Grouped Bar Chart for Apps Within the Gender in Respect of SUS Scores



4.4 Qualitative Results

This section includes responses to post-experiment questions obtained from research participants.

Question 1: Do you find the elements/colors/theme used in either of the two apps tobe feminine, masculine or gender-neutral? 19 participants gave direct answers that they found both the apps as gender neutral.

Participant 1 (female) found colors in FAPP to be more feminine, however said that both the apps are gender neutral in terms of content.

Participant 6 (female) said that the elements of MAPP are feminine while FAPP is gender neutral.

Participant 17 (female) said that both the apps appeared feminine to them while participant 10 (female) commented that both the apps are kind of masculine.

Participant 16 (male) mentioned that both the apps are generally suited towards a typical businessman aesthetic. To participant 18 (male), FAPP's look appeared masculine but MAPP came across as gender neutral.

Question 2: Which app do you think was designed by female and which one by male designer? Support your answer with reasoning.

Below are the answers from participants who correctly guessed the gender of both the app's designer along with the statements of reasoning provided by them in support of their response.

Participant 12 (female): "*Delhi app is designed by a male designer as it is more straightforward and the interface is very basic. Vancouver app is designed by a female as it had more navigation options, and I felt it was more in detail. The interface was beautifully designed giving a vibe of summer"*

Participant 5 (male): "*Delhi app did not have many pictures or colors and focused mostly on functionality while the Vancouver app had vibrant colors along with functionality*."

Participant 6 (female): "*Vancouver app was designed by female as the colors and theme were more vibrant and diverse. Delhi app was designed by male as it was simple and easy."*

Participant 7 (male): "*In my opinion, Delhi app was designed by a male designer while the Vancouver app was designed by a female designer. My answer is based on color schemes used in both the apps."*

Participant 19 (female): "*I feel like the Vancouver app was designed by a female due to it being more aesthetically pleasing. The New Delhi app was designed by a male due to the very simplistic approach on the design."*

Participant 20 (female): "*My guess would be that the Vancouver app was designed by a female designer because it was slightly more aesthetically pleasing: graphics on the login page, map to see destinations and the destinations descriptions was slightly easier to under- stand/less cumbersome, whereas the New Delhi app was more straightforward. It had less colors, very plain so I assume this one was from a male designer."*

Participant 29 (female): "*Delhi tourism is designed by a male and Vancouver tourism by a female because this is very colorful and females like colorful things."*

Incorrect responses to questions 2 are provided below:

Participant 1 (female): "*The New Delhi app was more aesthetic so it could have been designed by a female and the Vancouver app by a male."*

Participant 30 (female):""*I think the New Delhi app is designed by female designer be- because it looks better than the Vancouver one."*

Participant 4 (male): "*I would say that the Delhi app was designed by a female designer as it had more bright colors compared to the Vancouver app which had dark colors. More fancy icons on the Delhi app could be another reason too."*

Participant 18 (male): "*I think the Vancouver app is designed by a male. My reasoning for this is that it was significantly of inferior quality to the Delhi app. Less polish and overall aesthetic consideration was taken. This would be more consistent to 'get it done' attitude. I have experienced men use [this attitude] more than women. In my experience women take more consideration to make something aesthetically beautiful than men. The Delhi app had*

nice, rounded corners and looked very modern. Vancouver app was boxy and looked like an app developer 10 years ago."

Participant 26 (male): "*I think the Vancouver app was designed by male and Delhi app by a female. Reason behind this is that I had to disable dark mode of my device to enter basic information while sign up so this peculiarity seems to be left out by a female's mind."*

Question 3: While using an every-day product/device or website/interface, did you feel that the design team targeted only one particular gender.

Participant 1 (female): "*Many shopping brands are gender focused. For example, SHEIN is focused towards female audience, while some apps like Reddit and Twitter are gender neutral."*

Participant 19 (female): "*I don't find that most of the apps I use are targeted to a specific gender with the exception of one single app that is for female reproductive health.*"

Participant 11 (male): "*No I disagree. If we move on from the presumed association of certain aspects of fields for some gender over the other, we can see that all apps are essentially for both genders. Unless the apps had to do something with the biological aspects of males and females."*

One of the questions in demographic questionnaire for testers asked them about the apps most frequently used by them on daily basis. From their responses it was noticed that Instagram ranked as the most used app in everyday life by 13 participants followed by Whatsapp (used daily by 11 participants) and Amazon (used by 9 participants).

Chapter 5: Conclusion

The results obtained on the application of paired T-test, unpaired T-test, their respective non-parametric tests namely Wilcoxon's test and Mann Whitney test and ANOVA test, turned out to be statistically non-significant. The p-values associated with paired t-statistic and Wilcoxon's test statistic were 0.1006 and 0.1632 respectively when computed for male testers' rating of MAPP and FAPP. Similarly, for female testers, p-value for paired t-test was 0.2889 and p- value for Wilcoxon's test was 0.3447. Each p value being greater than the threshold or critical limit of 0.05, the null hypothesis could not be rejected. This implies that male and female testers did not rate the perceived credibility while completing the assigned task using both the apps with any statistically significant difference. In other words, it can be phrased that the UI designers did not target a particular gender while designing the interface. No statistically significant difference observed in male and female testers' rating of both the apps is an indication that neither of the gender group favored any specific app over the other.

Results from unpaired t-test and Mann Whitney test also revealed no statistical significance. As a result of this, null hypothesis for this case too could not be rejected which implied that for male designer's app and female designer's app no statistically significant difference exists in the mean SUS ratings given by male testers vis-à-vis female testers. Similarly, while attempting to detect any interaction effect between the apps and gender of testers using two-way ANOVA, null hypothesis could not be rejected.

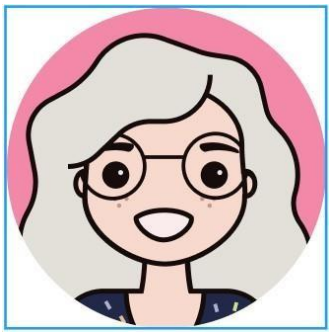
The most noticeable reason for results being statistically non-significant can be attributed to small sample size. Remote testing, online recruitment, and travel restrictions in accordance with COVID policies have been major barriers in recruiting a greater number of participants for testing. Even though the recruited participants belonged to diverse educational backgrounds, resided in different parts of the world, having a larger sample of at least 50 males and 50 females could have provided a better data set to test for statistical significance. Moreover, from the normality pattern and liker scale data, we can notice that there is a large variation in the responses given by participants. For example, when observing the Likert scale scores for female designer's app given by female testers, the minimum SUS score is 30 while the highest goes up to 87.5. Smaller the variation, better had been the results.

It must be mentioned that there were no prior expectations of what the SUS scores would be due to the unpredictable number of testing participants. As the experiment progressed, more participants were eventually recruited for testing, but SUS scores for each participant totally relied on their experience using those apps in accomplishing the task.

The present literature on gender inclusion in technology has no official difference stated between the terms gender inclusive vs gender neutral. For the works done on gender inclusive designs, the design team meant to design for ‘all ‘groups; similarly, when designing something as 'gender-neutral' it implies that the design does not reflect bias towards any gender or orientation towards specific gender groups. Therefore, both the terms essentially point in the same direction, though they can be interpreted differently. When any design is built for all groups, it means that it is inclusive. Hence, once the design is built in a manner that it should be inclusive, that is an initiative to obtain gender neutrality in its appearance and/or functionality; or we can say that ‘gender-neutrality' is associated with UI and 'gender-inclusiveness' with UX.

One segment of the research which has been excluded from the main objective is worth mentioning here. The UI designers also submitted an avatar in addition to their designs which would represent a virtual travel guide for their respective apps. It is fascinating to learn that both the designers chose to design their avatar with its gender matching as of their own. Male designer named his avatar, John while the female designer named her avatar Heidi. The avatars have been excluded from the main research because of the two main reasons. First, had the avatars been included then the participants might form the impression of the designer’s gender from the persona’s perceived gender. The second reason is that developing a dynamic, functioning virtual avatar with facial cues and impressions as well as VUI is a detailed topic of research in itself. Merging the current research with studying the influence of avatar gender would have diverted from the main objective of exploring the perceived credibility of interface designs to the immersive experience provided by virtual avatars.

Figure 5.1: Male Designer’s Persona, John (left) and Female Designer’s Persona, Heidi (right)



Chapter 6: Future Research Direction

This research mainly identified differences from the perspective of gender binary, that is, male and female. Looking at the west side of the world especially North America, where society is openly accepting and welcoming the individuals coming out as either non-binary or identifying themselves as members of other gender groups, the future research in gender studies related to design and development of interfaces should include the UI designs from these group members too. One of the reasons for the non-involvement of other gender in this study is that in developing countries such as India, the awareness regarding non-binary, other gender groups is still at its infancy. Since the test participants were recruited from North India as well as Pakistan, obtaining their feedback on the designs submitted by a designer who does not confirm to the gender binary of male and female, would be unfair due to limited number of 'open-talks' on this subject in these developing countries. Therefore, gender studies researchers and designers might integrate the involvement of all gender groups to make the future research and designs more inclusive.

Secondly, the research explored the perceived credibility of mobile applications with task limited to planning and saving the tourist destinations. To make the participation more intuitive, in future the experiment can involve understanding the participants' interaction with the interface through eye- tracking and conducting the experiment in HCI lab where the researcher is also aware of their activities on the interface to obtain an enhanced picture. Additionally, since the avatar designs submitted by the UI designers were excluded from main research objective and results, the study can be extended to include a virtual avatar acting as a tourist guide for each app with three different personas namely, male, female, and gender neutral. Hence, the behavior of participants can be studied while interacting with each kind of avatar along with their choice of avatar.

Chapter 7: Appendices

7.1 Appendix A: Recruitment Posters

Figure 7.1: Recruitment Poster for Designers

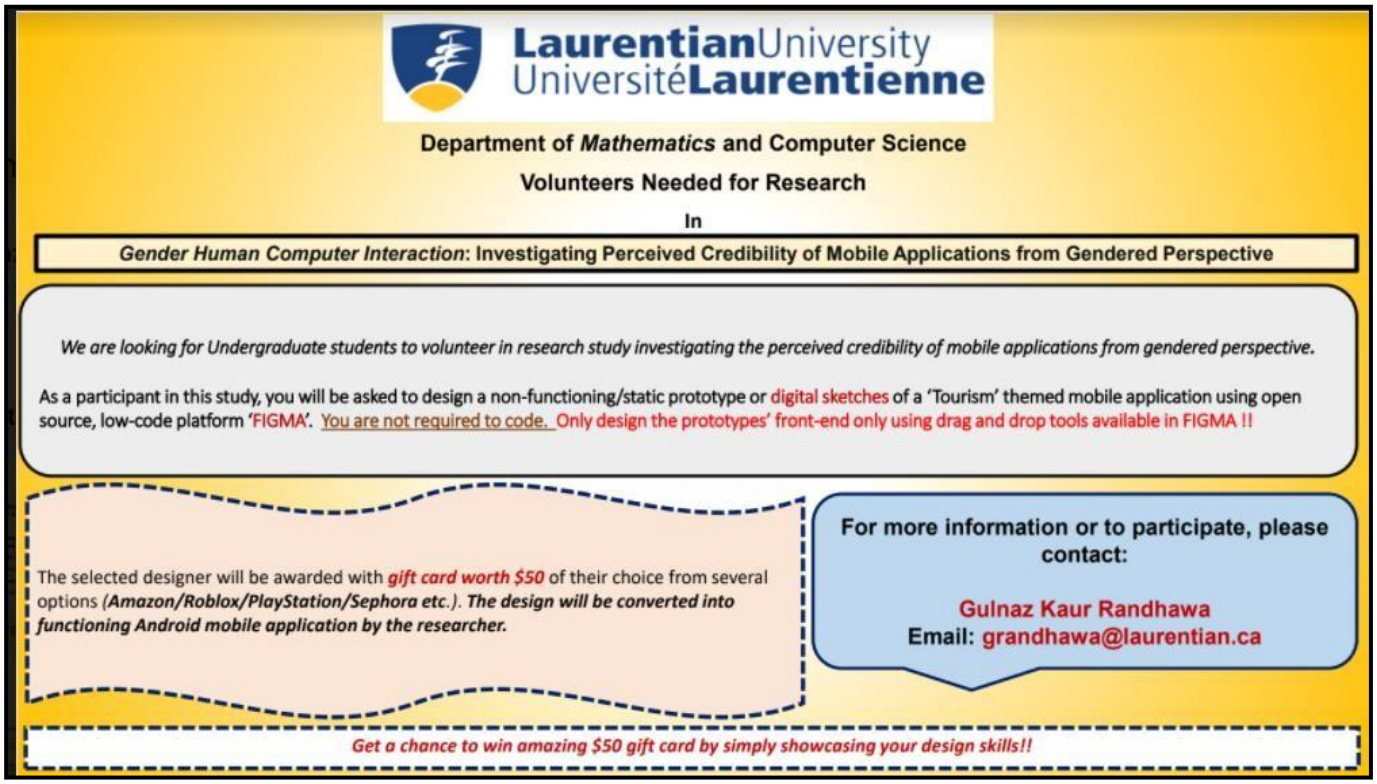
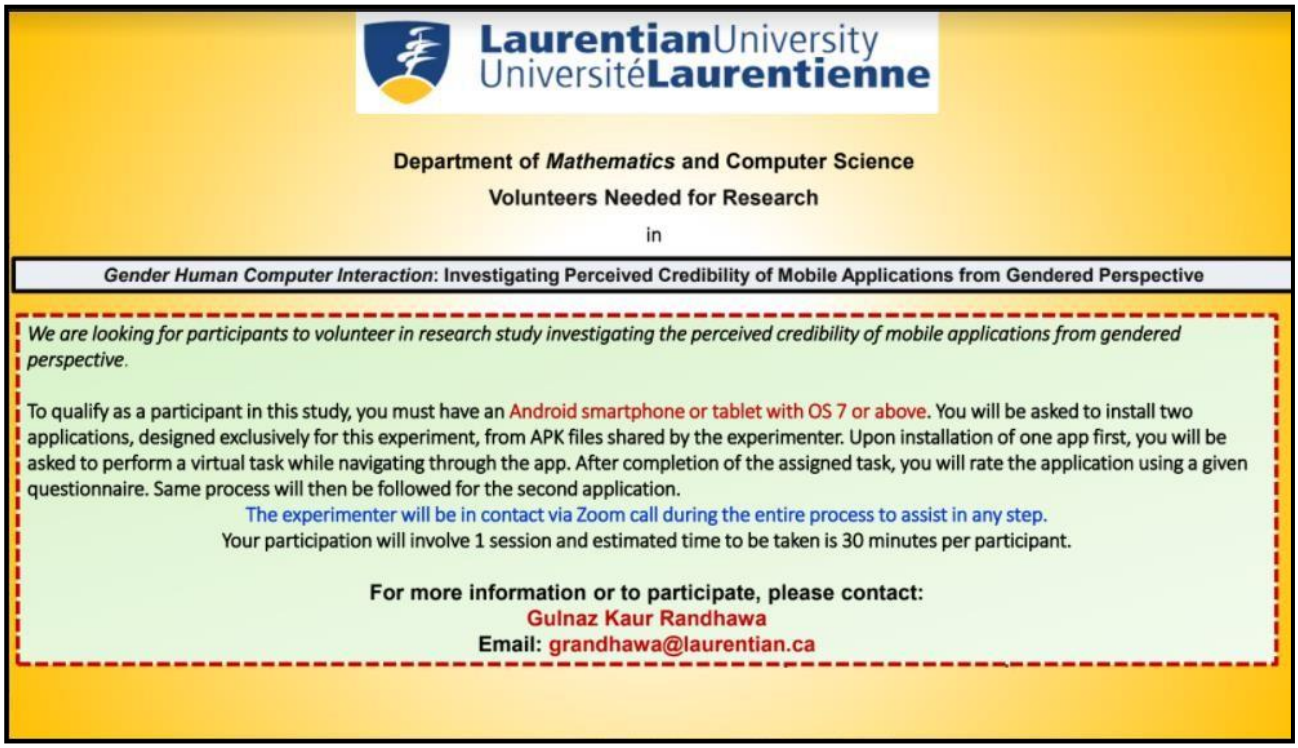


Figure 7.2: Recruitment Poster for Testers



LaurentianUniversity
UniversitéLaurentienne

Department of *Mathematics* and Computer Science

Volunteers Needed for Research

in

***Gender Human Computer Interaction*: Investigating Perceived Credibility of Mobile Applications from Gendered Perspective**

We are looking for participants to volunteer in research study investigating the perceived credibility of mobile applications from gendered perspective.

To qualify as a participant in this study, you must have an Android smartphone or tablet with OS 7 or above. You will be asked to install two applications, designed exclusively for this experiment, from APK files shared by the experimenter. Upon installation of one app first, you will be asked to perform a virtual task while navigating through the app. After completion of the assigned task, you will rate the application using a given questionnaire. Same process will then be followed for the second application.

The experimenter will be in contact via Zoom call during the entire process to assist in any step.

Your participation will involve 1 session and estimated time to be taken is 30 minutes per participant.

For more information or to participate, please contact:

Gulnaz Kaur Randhawa

Email: grandhawa@laurentian.ca

7.2 Appendix B: Consent Forms

7.2.1 Informed Consent for Designers

Figure 7.3: Consent form for Designers

LaurentianUniversity
UniversitéLaurentienne

INFORMED CONSENT TO PARTICIPATE IN A RESEARCH STUDY AS A MOBILE APP PROTOTYPE DESIGNER

Full Study Title: Gender Human Computer Interaction – Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective

Principal Investigator:

Gulnaz Kaur Randhawa Email: grandhawa@laurentian.ca

Supervisor:

Dr. Ratvinder Grewal Email: rgrewal@cs.laurentian.ca

INTRODUCTION

You are being asked to consider participating in this research study as a mobile app interface designer. Your participation will contribute towards first stage in the commencement of this experiment, that is, obtaining prototypes of mobile application. This experimental research involves participant designing a digital sketch or prototype of a mobile application. The objective is to convert into functioning Android applications by the researcher which will then be used for Usability Testing by recruiting members of Laurentian University.

WHAT WILL HAPPEN DURING THIS STUDY?

As an application interface designer during this stage 1 of the study, participant will be asked to get familiar with online available open-source, low-code/no-code platform 'Figma'. Using drag and drop tools available in Figma, participant will have to design a non-functioning prototype/digital sketch of a mobile application themed on 'Tourism'. Upon completing the design, participant will submit their work to the researcher through email. The researcher will convert the prototype into functioning Android mobile app by the researcher and the designer will be awarded with e-gift card. Finally, the mobile application will be remotely tested by another group of testers recruited online from within Laurentian University.

WHY IS THIS STUDY BEING DONE?

The purpose of asking participant to design the prototype is to get an insight into their design approach and criteria followed. Researcher will convert the design into working Android application. The functioning Android app will be tested by members of Laurentian University on System Usability Scale questionnaire.

LaurentianUniversity
UniversitéLaurentienne

INFORMED CONSENT

The purpose of this form is to obtain your consent to participate in a research study conducted by Gulnaz Kaur Randhawa under the supervision of Dr. Ratvinder Grewal from the Department of Mathematics and Computer Science at Laurentian University. Please feel free to ask the researcher any questions related to this study such as its purpose, your role as a participant, potential risks and benefits etc. After the researcher answers all the questions, you may decide to continue participating or withdraw. You have full right to decide whether you want to participate or not and can even stop your participation at any time. If you decide to stop participating, your data collected till that point will be permanently deleted.

WHAT ARE THE POTENTIAL BENEFITS?

Participant involved in the design phase may derive enjoyment or learn to use the online available platform 'Figma' for designing digital prototypes of mobile application interfaces. This will also allow them to improvise their designing skills.

WHAT ARE THE POTENTIAL RESEARCH STUDY RISKS?

Participant will not be exposed to any medical substances or hazardous chemicals. There is no risk associated with personal data as multiple measures are taken to ensure security. No personally identifiable information will be collected from participants apart from name and email address. The forms that do collect this information will be stored safely in REDCAP's storage which is protected with password only known to the researcher.

ARE PARTICIPANTS PAID TO PARTICIPATE IN THE STUDY?

The participant who get selected to design prototype for the experiment will be offered an e-gift card worth $50 of their choice from the options: Amazon, Play Station, Roblox, Sephora, etc. This will be done to thank them for their participation and for putting efforts into designing prototypes in accordance with the design guidelines

HOW WILL MY INFORMATION BE KEPT CONFIDENTIAL?

All the data from questionnaires will be stored securely in REDCAP's storage which is considered to be the most secured storage space as mentioned in one of the Laurentian University's IT webpage. The data in REDCAP can only be accessed by the researcher using LU login and password which is only known to the researcher. The data will also be encrypted in the researcher's personal computer and the personal computer's password is only known to the researcher. All the data including database and transcripts will be destroyed five years after the completion of this project. Any other information that is collected, used and disclosed for this study will be handled with confidentiality. Any information that we obtain from you which identify you will not be published or shared with anyone else.

INFORMATION ABOUT THE STUDY RESULTS

Participant has the right to know the study's results. If you wish to be notified about the results of this experimental study, provide us the email address below at which you should be contacted:

Email:_____

WHAT ARE THE RIGHTS OF THE PARTICIPANTS IN THE RESEARCH STUDY?

Any time during the research study participation, you can withdraw your consent to participate without any fear of consequence. You have right to receive all the necessary information that will help to decide to participate in this study. You have right to ask question about this research study and to have them answered to your satisfaction before you make any decision. Throughout the study, you have right to ask questions and receive answers. You have right to leave at any time during filling out questionnaire forms.

Issues and ethical concerns held by the participant regarding this study can be addressed to the Ethics Department at Laurentian University:

Research Ethics Officer

Office of Research Services

Telephone: 705-675-1151 ext. 3681 or 2436

Toll-free: 1-800-461-4030

Email: ethics@laurentian.ca

Any questions and concerns related to this study can be directed to:

Gulnaz Kaur Randhawa: grandhawa@laurentian.ca

Dr. Ratvinder Grewal: rgrewal@cs.laurentian.ca

LaurentianUniversity
UniversitéLaurentienne

DOCUMENTATION OF THE INFORMED CONSENT

You will be given a copy of this informed consent after it has been signed by you and the study staff.

Full Study Title: Gender Human Computer Interaction – Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective.

Name of Participant: _____

Participant: By signing this form, I confirm that:

- This research study has been fully explained to me and all of my questions have been answered to my satisfaction.
- I understand the requirements for participating in this research study.
- I have been informed of the rights of the research participants.
- I have been informed of the risks and benefits of participating in this research study.
- I have read each page of this form.
- I have agreed to participate in this research study

_____	_____	_____
Name of Participant	Signature	Date
_____	_____	_____
Name of Person Administering	Signature	Date

Thank you for participating in this research study.

7.2.2 Informed Consent for Testers

Figure 7.4: Consent form for Testers

LaurentianUniversity
UniversitéLaurentienne

INFORMED CONSENT TO PARTICIPATE IN A RESEARCH STUDY AS A MOBILE APP TESTER

Full Study Title: Gender Human Computer Interaction – Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective

Principal Investigator:

Gulnaz Kaur Randhawa Email: grandhawa@laurentian.ca

Supervisor:

Dr. Ratvinder Grewal Email: rgrewal@cs.laurentian.ca

INTRODUCTION

You are being asked to consider participating in this research study because respondents are an essential part of any prototype testing. It is vital to get opinions from the users in order to make any design as usable as possible. Your participation will bring to knowledge of the researcher how easy or difficult it was for you to get familiar with the user interface during your first contact with the prototype.

This experimental research involves participants testing two Android mobile applications based on 'Travel' theme. The idea is to identify which applicationscores better in terms of User-Experience derived.

WHAT WILL HAPPEN DURING THIS STUDY?

The purpose of this study is to test two prototypes of mobile applications, designed for this study by two different individuals. As a participant, you will be asked to install one application first from the APK file shared by the researcher. After downloading the app, you will be given a task to complete using the app whose demonstration will be provided by the experimenter at the beginning. Upon completion of the assigned task, you will be asked to fill out a questionnaire form to give us an overview about your experience using the app. The same procedure will then be followed for the second application. After both the applications are tested and their respective questionnaires are submitted, the experimenter will ask open-ended questions related to the nature of the study.

WHY IS THIS STUDY BEING DONE?

This study is being conducted to get an insight into which mobile application ranks better on System Usability Scale. The idea is to compare the designs of User Interface designs originating from two different designers. By doing so, it will become clear as what aspects in one mobile User Interface are more appealing to users as compared to the other design. This is an attempt to explore whether or not designers aim to make their designs usable to a specific audience or they design in a manner that User Experience derived is nearly same for anyone using it.

LaurentianUniversity
UniversitéLaurentienne

INFORMED CONSENT

The purpose of this form is to obtain your consent to participate in a research study conducted by Gulnaz Kaur Randhawa under the supervision of Dr. Ratvinder Grewal from the Department of Mathematics and Computer Science at Laurentian University. Please feel free to ask the researcher any questions related to this study such as its purpose, your role as a participant, potential risks and benefits etc. After the researcher answers all the questions, you may decide to continue participating or withdraw. You have full right to decide whether you want to participate or not and can even stop your participation at any time. If you decide to stop participating, your data collected till that point will be permanently deleted.

WHAT ARE THE POTENTIAL BENEFITS?

Benefits to participants may include increased awareness about how UI testing takes place by interacting with interfaces and how two different applications based on similar theme are tested as well as compared.

WHAT ARE THE POTENTIAL RSEARCH STUDY RISKS?

Participants will not be exposed to any medical substances or hazardous chemicals. There is no risk associated with personal data as multiple measures are taken to ensure security. No personally identifiable information will be collected from participants apart from name and email address. The forms that do collect this information will be stored safely in REDCAP's storage which is protected with password only known to the researcher.

ARE PARTICIPANTS PAID TO PARTICIPATE IN THE STUDY?

This study is a voluntary participation; therefore, you will not be paid for participating.

HOW WILL MY INFORMATION BE KEPT CONFIDENTIAL?

All the data from questionnaires will be stored securely in REDCAP's storage which is considered to be the most secured storage space as mentioned in one of the Laurentian University's IT webpage. The data in REDCAP can only be accessed by the researcher using LU login and password which is only known to the researcher. The data will also be encrypted in the researcher's personal computer and the personal computer's password is only known to the researcher. All the data including database and transcripts will be destroyed five years after the completion of this project. Any other information that is collected, used and disclosed for this study will be handled with confidentiality. Any information that we obtain from you which identify you will not be published or shared with anyone else.

INFORMATION ABOUT THE STUDY RESULTS

Participants have right to know the study's results. If you wish to be notified about the results of this experimental study, provide us the email address below at which you should be contacted:

LaurentianUniversity
UniversitéLaurentienne

Email:_____

WHAT ARE THE RIGHTS OF THE PARTICIPANTS IN THE RESEARCH STUDY?

Any time during the research study participation, you can withdraw your consent to participate without any fear of consequence. You have right to receive all the necessary information that will help to decide to participate in this study. You have right to ask question about this research study and to have them answered to your satisfaction before you make any decision. Throughout the study, you have right to ask questions and receive answers. You have right to leave at any time during filling out questionnaire forms.

Issues and ethical concerns held by the participant regarding this study can be addressed to the Ethics Department at Laurentian University:

Research Ethics Officer

Office of Research Services

Telephone: 705-675-1151 ext. 3681 or 2436

Toll-free: 1-800-461-4030

Email: ethics@laurentian.ca

Any questions and concerns related to this study can be directed to:

Gulnaz Kaur Randhawa: grandhawa@laurentian.ca

Dr. Ratvinder Grewal: rgrewal@cs.laurentian.ca

DOCUMENTATION OF THE INFORMED CONSENT

You will be given a copy of this informed consent after it has been signed by you and the study staff.

Full Study Title: Gender Human Computer Interaction – Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective.

Name of Participant: _____

Participant: By signing this form, I confirm that:

- This research study has been fully explained to me and all of my questions have been answered to my satisfaction.
- I understand the requirements for participating in this research study.

LaurentianUniversity
UniversitéLaurentienne

- I have been informed of the rights of the research participants.
- I have been informed of the risks and benefits of participating in this research study.
- I have read each page of this form.
- I have agreed to participate in this research study

Name of Participant	Signature	Date
Name of Person Administering	Signature	Date

Thank you for participating in this research study.

7.3 Appendix C: Questionnaires

7.3.1 Demographic Questionnaire for Designer

Figure 7.5 Demographic Questionnaire for Designers

LaurentianUniversity
UniversitéLaurentienne

Demographics Questionnaire for Designers

1. How do you identify yourself?

   Male Female Non-binary

2. Program of Study

3. Do you have prior experience in UI designing?

   Yes No

4. If yes, which platform have you used before for designing? (Type N/A if answer to previous question is No)

5. Are you familiar with low code, no-code platform called 'Figma'?

   Yes No

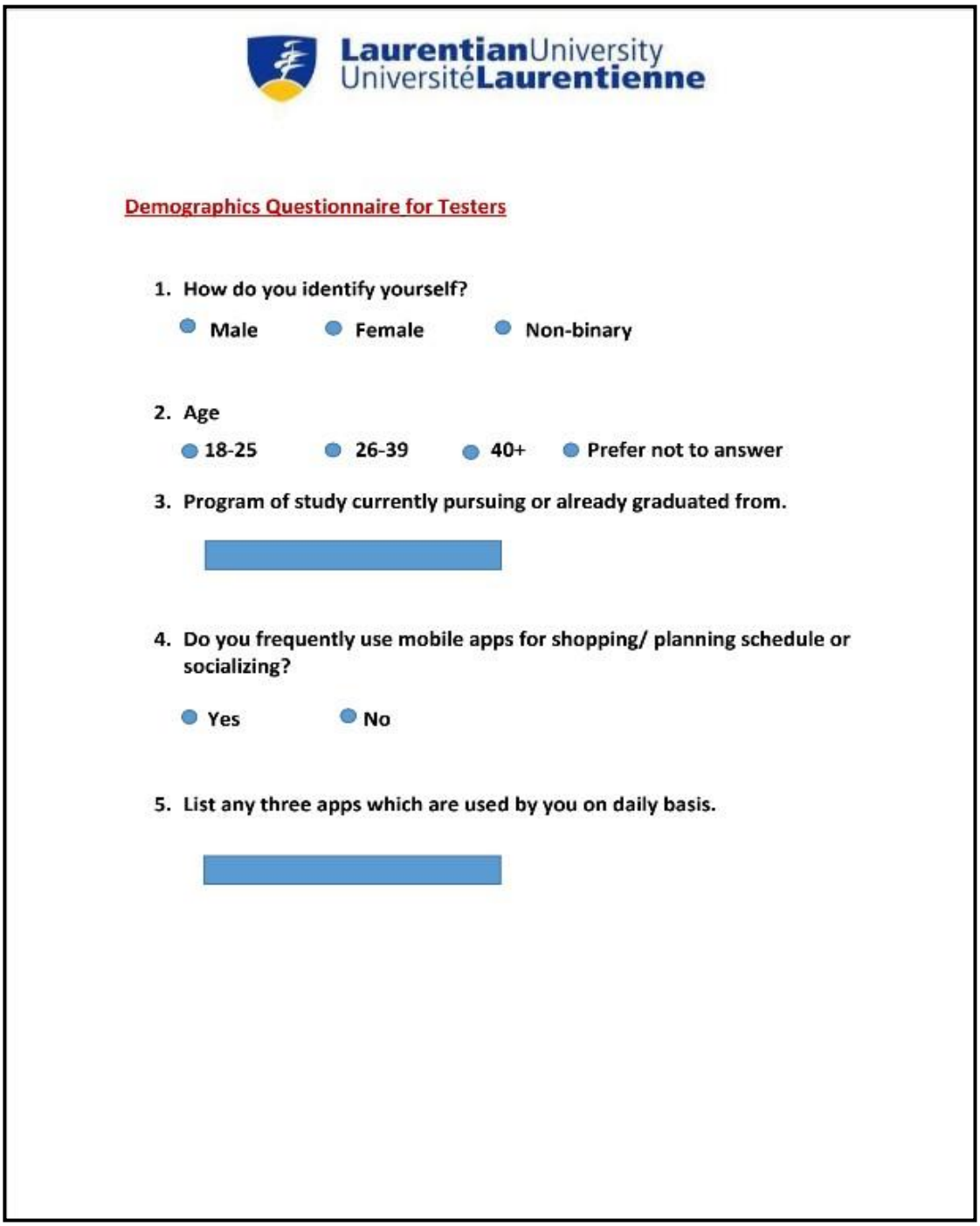
6. How would you rate your designing skills?

   1 2 3 4 5

1: Worst
2: Poor
3: Average
4: Good
5: Excellent

7.3.2 Demographic Questionnaire for Testers

Figure 7.6 Demographic Questionnaire for Testers



LaurentianUniversity
UniversitéLaurentienne

Demographics Questionnaire for Testers

1. How do you identify yourself?
   - Male
   - Female
   - Non-binary
2. Age
   - 18-25
   - 26-39
   - 40+
   - Prefer not to answer
3. Program of study currently pursuing or already graduated from.
4. Do you frequently use mobile apps for shopping/ planning schedule or socializing?
   - Yes
   - No
5. List any three apps which are used by you on daily basis.

7.3.3 System Usability Scale Questionnaire

Figure 7.7 System Usability Scale Questionnaire

LaurentianUniversity
UniversitéLaurentienne

System Usability Scale (SUS) Survey Questionnaire

Please complete the survey below.

Thank you!

1) How do you identify yourself?

○ Male ○ Female ○ Non-Binary

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
2)	I would recommend this application to a tourist travelling to Vancouver.	○	○	○	○	○
3)	I found this application to be more complicated than it should be.	○	○	○	○	○
4)	I think planning a trip using this app is simple and easy.	○	○	○	○	○
5)	While completing the assigned task, I needed support from the researcher.	○	○	○	○	○
6)	I found various functions in the app were well integrated and worked smoothly.	○	○	○	○	○
7)	I think there are certain irregularities/inconsistencies in this app.	○	○	○	○	○
8)	I would imagine that most users will learn to use this app very quickly.	○	○	○	○	○
9)	Finding information or navigating in this app was cumbersome.	○	○	○	○	○
10)	I felt confident using this app to complete the assigned task.	○	○	○	○	○
11)	I think there are lot of things in this app that require improvement.	○	○	○	○	○

7.3.4 Post-Experiment Questions

Figure 7.8 Post-Experiment Questionnaire

REDCap

LaurentianUniversity
UniversitéLaurentienne

Post-Experiment Questionnaire

Please answer the following questions in few words.

1)	How do you identify yourself?	○ Male ○ Female ○ Non-binary
2)	Did you find the elements/colors/theme used in either of the two applications tested to be more masculine or feminine or gender-neutral?	_____
3)	In this experiment, one of the apps was designed by a male designer and the other by a female designer. After testing and navigating through both the apps, which do you think was designed by a male designer and which one by a female designer? Please give reasons for your answer.	_____
4)	While using an every-day product/device or website/interface, did you feel that the design team targeted only one particular gender.	_____

7.4 Appendix D: Debriefing Forms

7.4.1 Debriefing Form for Designers

Figure 7.9 Debriefing form for Designers

LaurentianUniversity
UniversitéLaurentienne

Debriefing Form for Participants Involved in Designing

Study Title: Gender Human-Computer Interaction – Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective

Faculty Supervisor: Dr. Ratvinder Grewal

Student Investigator: Gulnaz Kaur Randhawa

Thank you for spending your valuable time designing mobile application prototype for our research. Your participation is greatly appreciated by the research team.

In the beginning of the research, you were informed that the purpose of this experimental study is to obtain designs from you which will then be converted into functioning Android applications. These apps will then be subject to Usability Testing during testing stage by members of Laurentian University. However, details about the experiment and the study's actual research objective were left out on purpose. This means that only partial information was shared with you earlier. Some studies on Usability Testing involve deception in which the participants are informed about the actual purpose only after completion.

In this research we aim to determine whether User Experience derived while using mobile applications is influenced by the gender of its designer. Precisely, the objective is to identify if UX derived from using mobile apps by male and female users is different for an app designed by a similar gender designer and opposite gender designer. This shall bring to notice any synthesis between gender of designer and gender of the user while rating perceived credibility of mobile interfaces. The research results will also highlight if the aesthetic elements used in designs by male and female designer are perceived to be masculine, feminine or gender neutral by the testers. Also, the research aims to identify whether through the particular choice of colors and layouts, users of a particular gender feel marginalized or left out as the potential user group.

To conduct this experiment, two participants (one who identifies as male and the other who identifies as female) were recruited to design mobile application prototypes. Participants were given same demonstration and briefing about the experiment. Both the participants were asked to design a prototype for a 'Tourism' themed mobile application using Figma and given a time of 1 week to submit their designs. They were told that their designs will then be converted into functioning mobile apps by the researcher after coding. The functioning versions would be tested by members of Laurentian University during testing stage. Both the participants were awarded e-gift cards worth $50 of their choice.

Both the participants were not told that their prototype design will be compared with the other designer's prototype during testing in terms of whether the applications appear to have masculine, feminine or gender neutral elements and which

LaurentianUniversity
UniversitéLaurentienne

applications obtains higher SUS scores by users belonging to male and female gender groups.

During testing, each participant will be asked to download one designer's mobile application first. The experimenter will share its APK file with the participant. After installation, participants will have to navigate through the app and complete the assigned task. On completion of the task, participant will rate the experience on System Usability Scale questionnaire. Next, same procedure will be followed for the second application. Participants will not be told about the gender of the applications designer during the testing phase. The order for testing the applications will be different for the participants to prevent any bias. When testing phase will be over, results from SUS questionnaire will be analyzed and compared across both the applications. The results will be segregated in terms of gender of application's designer and gender of tester. At the time of analyzing results, the SUS questionnaire scores of both the applications will be calculated for each participant. Further, the results will be categorized in terms of gender of the application designer and gender of the application tester. Researcher will graphically plot the results depicting whether application designed by male designer and the one designed by female designer is perceived differently by participants of both the gender groups on the basis of SUS scores of each app.

We did not reveal the details about the experiment to participants involved in designing at the beginning because we did not want any conscious or unconscious bias being introduced in their respective designs during the choice of colors and other visible attributes. This was done to ensure that participants make their design decisions freely without being influenced by the research objective or designing under pressure that their design will be compared with another participant's design.

We apologize for not providing entire details about the research objective in the beginning. This debriefing form is, hence, to provide you with each and every detail about the true nature of the research. Therefore, we hope that you understand the reason for the use of deception.

The data and information you provided in consent form and demographic questionnaire will be securely stored on experimenter's personal password protected computer with its password only known to experimenter. The data and information will not be shared with anyone outside the research team. The electronic data will be deleted after 5 years. You may choose to withdraw the information and data shared provided anytime by contacting the researcher without any penalty.

Since the actual research objective was not revealed at first and now this debriefing form provides all the details about this study, we request you not to disclose the research objective with anyone. Discussing this study's research objective and procedures to be followed could affect the results to be obtained during testing.

This study has received Ethics clearance through Laurentian University Research Ethics Board. If you have any questions for the Research Committee, you may contact Ethics Department at Laurentian University:

LaurentianUniversity
UniversitéLaurentienne

Research Ethics Officer
Office of Research Services
Telephone: 705-675-1151 ext. 3681 or 2436
Toll free: 1-800-461-4030
Email: ethics@laurentian.ca

Any questions related to the study can be directed to:

Gulnaz kaur Randhawa: grandhawa@laurentian.ca
(Principle Investigator)

Dr. Ratvinder Grewal: rgrewal@cs.laurentian.ca
(Faculty Supervisor)

Your signature below indicates that:

- You have been debriefed about the true nature of this research.
- You give permission to the researcher to use the data and information provided by you for using in this research.
- All your questions have been answered
- You understand the reason for using deception.
- You have read each page of this form.

_____	_____	_____
Name of Participant	Signature	Date
_____	_____	_____
Name of Researcher	Signature	Date

Thank you for participating in this research study.

7.4.2 Debriefing form for Testers

Figure 7.10 Debriefing form for Testers

LaurentianUniversity
UniversitéLaurentienne

Debriefing Form for Participants Involved in Testing

Study Title: Gender Human-Computer Interaction – Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective

Faculty Supervisor: Dr. Ratvinder Grewal

Student Investigator: Gulnaz Kaur Randhawa

Thank you for spending your valuable time testing the mobile applications for our research. Your participation is greatly appreciated by the research team.

In the beginning, you were informed that the purpose of this testing stage is to compare the mobile applications designed by two different designers. The objective being to identify which application ranks better on System Usability Scale. However, details about the experiment and the study's actual research objective were left out on purpose. This means that only partial information was shared with you earlier. Some studies on Usability Testing involve deception in which the participants are informed about the actual purpose only after completion.

In this research we aim to determine whether User Experience derived while using mobile applications is influenced by the gender of its designer. Precisely, the objective is to identify if UX derived from using mobile apps by male and female users is different for an app designed by a similar gender designer and opposite gender designer. This shall bring to notice any synthesis between gender of designer and gender of the user while rating perceived credibility of mobile interfaces. The research results will also highlight if the aesthetic elements used in designs by male and female designer are perceived to be masculine, feminine or gender neutral by the testers. Also, the research aims to identify whether through the particular choice of colors and layouts, users of a particular gender feel marginalized or left out as the potential user group.

To conduct this experimental research, two UI designers were recruited. One participant identified as male and the other as female, both were selected to design mobile app prototypes. The participants were selectedfollowing the criteria that both should be from similar academic background and withsame level of familiarity using FIGMA, the platform used for designing prototypes.
Both of them were given demonstration about the experiment and application's theme. Participants submitted their respective prototype designs to the researcher. These designs were then converted into functioning Android mobile applications by the researcher after coding in JAVA. During the testing stage, each participant installed both the applications, navigated through their screens, completed the assigned task and then rated their experience on SUS questionnaire. This process was done one application at a time and then repeated for the other. After answering SUS questionnaires, participants answered open ended questions at the end of experiment.

LaurentianUniversity
UniversitéLaurentienne

The participants were not told that the aim of this research study is to find any synthesis between gender of interface designer and gender of the user. They were not made aware of the fact that this study is gender based. Participants were not told about the gender of the application designer and were acknowledged about this only after their participation in testing phase was successfully done.

At the time of analyzing results, the SUS questionnaire scores of both the applications will be calculated for each participant. Further, the results will be categorized in terms of gender of the application designer and gender of the application tester. Researcher will graphically plot the results depicting whether application designed by male designer and the one designed by female designer is perceived differently by participants of both the gender groups on the basis of SUS scores of each app.

We did not reveal about the gendered aspect of the research as well as the gender of the respective applications' designer to prevent any conscious or unconscious bias being propagated at the time of testing. This was done to ensure that participants navigated through each app and rated their experience without being influenced by the gender of the designer and objective of the experiment.

We apologize for not providing entire details about the research objective in the beginning. This debriefing form is, hence, to provide you with each and every detail about the true nature of the research. Therefore, we hope that you understand the reason for the use of deception.

The data and information you provided will be securely stored on experimenter's personal password protected computer with its password only known to experimenter. The data and information will not be shared with anyone outside the research team. The electronic data will be deleted after 5 years. You may choose to withdraw the information and data shared provided anytime by contacting the researcher without any penalty.

Since the actual research objective was not revealed at first and now this debriefing form provides all the details about this study, we request you not to disclose the research objective with anyone. Discussing this study's research objective and procedures to be followed could affect the results to be obtained during testing with other participants.

This study has received Ethics clearance through Laurentian University Research Ethics Board. If you have any questions for the Research Committee, you may contact Ethics Department at Laurentian University:
Research Ethics Officer
Office of Research Services
Telephone: 705-675-1151 ext. 3681 or 2436
Toll free: 1-800-461-4030
Email: ethics@laurentian.ca

LaurentianUniversity
UniversitéLaurentienne

Any questions related to the study can be directed to:

Gulnaz Kaur Randhawa:
grandhawa@laurentian.ca(Principle Investigator)

Dr. Ratvinder Grewal: rgrewal@cs.laurentian.ca
(Faculty Supervisor)

Your signature below indicates that:

- You have been debriefed about the true nature of this research.
- You give permission to the researcher to use the data and information provided by you for using in this research.
- All your questions have been answered
- You understand the reason for using deception.
- You have read each page of this form.

Name of Participant	Signature	Date
Name of Researcher	Signature	Date

Thank you for participating in this research study.

7.5 Appendix E: Ethics Approval Certificate

Figure 7.11 Ethics Approval Certificate

LaurentianUniversity
UniversitéLaurentienne

APPROVAL FOR CONDUCTING RESEARCH INVOLVING HUMAN SUBJECTS
Research Ethics Board – Laurentian University

This letter confirms that the research project identified below has successfully passed the ethics review by the Laurentian University Research Ethics Board (REB). Your ethics approval date, other milestone dates, and any special conditions for your project are indicated below.

TYPE OF APPROVAL / New X / Modifications to project / Time extension	
Name of Principal Investigator and school/department	Gulnaz Randhawa, Math and Computer Sciences, supervisor Ratvinder Grewal
Title of Project	Gender Human-Computer Interaction: Investigating the Perceived Credibility of Mobile Applications from Gendered Perspective
REB file number	6020990
Date of original approval of project	April 28, 2021
Date of approval of project modifications or extension *(if applicable)*	
Final/Interim report due on: *(You may request an extension)*	April 28, 2022
Conditions placed on project	

During the course of your research, no deviations from, or changes to, the protocol, recruitment or consent forms may be initiated without prior written approval from the REB. If you wish to modify your research project, please refer to the Research Ethics website to complete the appropriate REB form.

All projects must submit a report to REB at least once per year. If involvement with human participants continues for longer than one year (e.g. you have not completed the objectives of the study and have not yet terminated contact with the participants, except for feedback of final results to participants), you must request an extension using the appropriate LU REB form. In all cases, please ensure that your research complies with Tri-Council Policy Statement (TCPS). Also please quote your REB file number on all future correspondence with the REB office.

Congratulations and best wishes in conducting your research.

Rosanna Langer

Rosanna Langer, PHD, Chair, *Laurentian University Research Ethics Board*

References

[1] Huff, C., & Cooper, J. (1987). Sex bias in educational software: The effect of designers' stereotypes on the software they design 1. *Journal of Applied Social Psychology*, *17*(6), 519-532.

[2] Van Oost, E. (2003). Materialized gender: how shavers configure the users' femininity and masculinity. In How Users Matter: The Co-construction of Users and Technology (pp. 193–208).

[3] (Boyle, 2013) Boyle, R., 2013. THE BENEFITS OF REDUCED-SIZE PIANO KEYBOARDS FOR SMALLER HANDED PIANISTS: AN EXPLORATION OF BIOMECHANICAL AND PHYSIOLOGICAL FACTORS.

[4] (Moss and Colman, 2001) Moss, G., Colman, A.M., 2001. Choices and preferences: Experiments on gender differences. J. Brand Manag. 9, 89–98. https://doi.org/10.1057/palgrave.bm.2540057

[5] (Stilma, 2008) Stilma, M., 2008. THE INLFUENCE OF THE DESIGNER'S GENDER [WWW Document]. undefined. URL /paper/THE-INLFUENCE-OF-THE-DESIGNER%27S-GENDER Stilma/04f84903e421acdbe0924baeeec9fa331a087679 (accessed 1.26.21).

[6] (Flanagin and Metzger, 2003) Flanagin, A.J., Metzger, M.J., 2003. The perceived credibility of personal Web page information as influenced by the sex of the source. Comput. Hum. Behav. 19, 683–701. https://doi.org/10.1016/S0747-5632(03)00021-9

[7] Simone Stumpf; Anicia Peters; Shaowen Bardzell; Margaret Burnett; Daniela Busse; Jessica Cauchard; Elizabeth Churchill, Gender-Inclusive HCI Research and Design: A Conceptual Review, now, 2020

[8] ("Global LinkedIn user distribution by gender 2021," n.d.) Global LinkedIn user distribution by gender 2021 [WWW Document], n.d. Statista. URL https://www.statista.com/statistics/933964/distribution-of-users-on-linkedin-worldwide-gender/ (accessed 2.23.21).

[9] ("Inclusive Crash Test Dummies: Analyzing Reference Models | Gendered Innovations," n.d.) Inclusive Crash Test Dummies: Analyzing Reference Models | Gendered Innovations [WWW Document], n.d. URL https://genderedinnovations.stanford.edu/case-studies/crash.html (accessed 2.23.21).

[10] Perez, C. C. (2019). Invisible Women: Data Bias in a World Designed for Men (First Printing ed.). Harry N. Abrams.

[11] ("Women in Tech," n.d.) Women in Tech: The Facts (2016 Update) [WWW Document], n.d. . Natl. Cent. Women Inf. Technol. URL https://www.ncwit.org/resources/women-tech-facts-2016-update (accessed 2.23.21).

[12] Schwartzman, R., & Decker, M. (2008). A car of her own: Volvo's "Your Concept Car" as a vehicle for feminism? Studies in Popular Culture, 30(2), 100–118.

[13] (rmaanushi_joshi (she/her), 2021). Gender-inclusive design is the only way [WWW Document]. Medium. URL https://uxdesign.cc/gender-inclusive-design-is-the-only-way-968494d5afc2 (accessed 9.9.21).

[14] Coleman, R.: The case for inclusive design - an overview. In: 12th Triennial Congress, International Ergonomics Association and the Human Factors Association of Canada, Toronto, Canada (1994)

[15] Clarkson, P.J., Coleman, R., Keates, S., Lebbon, C. (eds.): Inclusive Design: Design for the Whole Population. Springer, London (2003). https://doi.org/10.1007/978-1-4471-0001-0

[16] (Zhu et al., 2020) Zhu, H., Gruber, T., Dong, H., 2020. Value and Values in Inclusive Design. pp. 109– 122. https://doi.org/10.1007/978-3-030-50252-2_9

[17] ("Inclusive Design Examples," n.d.) Inclusive Design Examples [WWW Document], n.d. URL https://webcache.googleusercontent.com/search?q=cache:7IZOoz0xshAJ:https://www.diversitybestpractices.com/sites/diversitybestpractices.com/files/attachments/2020/11/inclusive_design_examples.pdf+&cd= 13&hl=en&ct=clnk&gl=ca (accessed 9.9.21).

[18] Yatim, M.H.M., Nacke, L., & Masuch, M. (2006). Improving game design by understanding the gender differences: The cognitive approach. In Proceedings of the International Conference on Gender in Educational Games and Gender Sensitive Approaches to E-Learning 2006. Donau University Krems: Austria.

[19] THOMAS, Angela. & WALKERDINE, Valerie. (2000). Girls and computer games. In proceedings of 4th European Feminist Research Conference: Body gender subjectivity crossing disciplinary and institutional borders. Bologna, Italy.

[20] ROE, Keith & MUIJS, Daniel: Children and Computer Games: A profile of the heavy user, p. 181–200.
In: Journal for European Journal of Communication (1998), Vol. 13. No. 2

[21] Jones, C. (1997, December 30, 1997). Lara Croft, Female Enemy Number One? The Mining Company Guide.

[22] Girl Trouble. (1998, January, 1998). Next Generation, January, 98-102.

[23] Cassell, J. (2002). Genderizing HCI. In J. Jacko & A. Sears (Eds.), The Handbook of Human-Computer Interaction (pp. 402–411). Mahwah, NJ: Lawrence Erlbaum. Retrieved July 25, 2006 http://www.soc.northwestern.edu/justine/publications/gender.hci.just.pdf

[24] (Doris et al., 2019) Doris, Á., Doris, Á., Doris, Á., 2019. How gender bias impacts what we think of products made by women [WWW Document]. Fast Co. URL https://www.fastcompany.com/90341254/how- gender-bias-impacts-what-we-think-of-products-made-by-women (accessed 9.9.21).

[25] (Archer et al., 2020) Archer, S., Brewer, H., Fain, S., Smith, S., 2020. Gender Differences and Cockpit Design. Stud. Works.

[26] (Prates et al., 2020) Prates, M.O.R., Avelar, P.H., Lamb, L.C., 2020. Assessing gender bias in machine translation: a case study with Google Translate. Neural Comput. Appl. 32, 6363–6381. https://doi.org/10.1007/s00521-019-04144-6

[27] Papenfuss M (2017) Woman in China says colleague's face was able to unlock her iPhone X. http://www.huffpostbrasil.com/entry/iphone-face-recognition-double_us_5a332cbce4b0ff955ad17d50

[28] Garcia M (2016) Racist in the machine: the disturbing implications of algorithmic bias. World Policy J 33(4):111–117

[29] (Eyssel and Hegel, 2012) Eyssel, F., Hegel, F., 2012. (S)he's Got the Look: Gender Stereotyping of Robots1. J. Appl. Soc. Psychol. 42, 2213–2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

[30] Brewer, J., Bassoli, A.: Reflections of Gender, Reflections on Gender: Designing Ubiquitous Computing Technologies. In: Gender & Interaction: Real and Virtual Women in a Male World, Workshop at AVI, pp. 9–12 (2006)

[31] Raj, Anil & Beach, Jameson & Stuart, Elizabeth & Vassiliades, Lauren. (2015). Multimodal and multisensory displays for perceptual tasks. 10.13140/RG.2.1.2075.1769.

[32] Sandygulova, A., O'Hare, G.M.P. Age- and Gender-Based Differences in Children's Interactions with a Gender-Matching Robot. *Int J of Soc Robotics* **10,** 687–700 (2018). https://doi.org/10.1007/s12369-018- 0472-9

[33] G. Gao, Y. Zhang, Y. Bu and P. C. Shih, "Design Exploration on Robots to Support Single Women at Home in China," *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2018, pp. 1-6, doi: 10.1109/ROMAN.2018.8525694.

[34] García-Peñalvo, F. J., García-Holgado, A., Vázquez-Ingelmo, A., & Seoane-Pardo, A. M. (2018, July). Usability test of WYRED Platform. In *International Conference on Learning and Collaboration Technologies* (pp. 73-84). Springer, Cham.

[35] Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: Adding an adjective rating scale. J. Usability Stud. **4**, 114–123 (2009)

[36] Lopes, M. R., & Vogel, C. (2020, September). Gender Effects in Mobile Application Development. In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)* (pp. 1-6). IEEE.

[37] Barth, D. (2012). Designing the gender-Neutral user experience. *Worcester Polytechnic Institute*.

[38] Orchard, R. (n.d.). *Apple's cycle tracking: A personal review*. MacStories. Retrieved December 8, 2021, from https://www.macstories.net/stories/apples-cycle-tracking-a-personal-review/.

[39] Levy, J. (2018, May). 'It's your period and therefore it has to be pink and you are a girl' users' experiences of (de-) gendered menstrual app design. In *Proceedings of the 4th Conference on Gender & IT* (pp. 63-65).

[40] Thomas, G. M., Lupton, D., & Pedersen, S. (2018). 'The appy for a happy pappy': expectant fatherhood and pregnancy apps. *Journal of Gender Studies*, *27*(7), 759-770