GAYE AŞKIN

AUTONOMY AND ROBOTONOMY

BILKENT UNIVERSITY 2023

AUTONOMY AND ROBOTONOMY: THE ROLE OF GENDER

A Master's Thesis

by
GAYE AŞKIN

Department of
Psychology
İhsan Doğramacı Bilkent University
Ankara

June 2023

AUTONOMY AND ROBOTONOMY: THE ROLE OF GENDER

The Graduate School of Economics and Social Sciences of
İhsan Doğramacı Bilkent University

by

GAYE AŞKIN

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS IN PSYCHOLOGY

THE DEPARTMENT OF
PSYCHOLOGY
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

June 2023

AUTONOMY AND ROBOTONOMY: THE ROLE OF GENDER
By Gaye Aşkın

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Psychology.

…………………………………
Burcu Ayşen Ürgen
Advisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Psychology.

…………………………………
Seçil Gönültaş
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Psychology.

…………………………………
Cengiz Acartürk
Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

…………………………………
Refet Soykan Gürkaynak
Director

ABSTRACT

AUTONOMY AND ROBOTONOMY: THE ROLE OF GENDER

Gaye Aşkın

M.A.in Psychology

Advisor: Dr. Burcu Ayşen Ürgen

June 2023

This study examines how gender affects people's perception of autonomy in social robots. It investigates whether an agent's name (masculine/feminine/neutral/ technical/no-name) and type (human/robot) can impact agency, personification, competency, and gender evaluations and whether these evaluations are at an unconscious level. The study consists of a pre-study and a main study with 150 participants. Participants watched 4-second 18 videos in the main study and answered autonomy-related questions. They also completed three implicit association tasks, adapted according to the study's themes, autonomy, agents, and gender. ANOVA analysis was conducted to analyze the agent, naming, and gender effects for the explicit part. The d-score was calculated for all participants in the implicit part, and an ANOVA analysis was conducted. Regression analysis was conducted to determine gender attribution in the pre-study and main studies. Correlation analyses were also conducted to determine if explicit-implicit parts were correlated. Lastly, a thematic analysis was conducted for qualitative inputs in the explicit part and categorized into nine themes. The study found the main effect of action and agent in agency-level attribution. In competency and gender attribution, agents had no main effect in none of the name conditions. In the implicit part, women and men participants differed in men-independent/women-dependent association IAT-1. The other two IATs, women and men participants, responded similarly. The study suggests that name manipulation does not affect people's autonomy perception, but rather agent types and actions characteristics affect them. Furthermore, people's implicit and explicit answers do not predict each other.

Key Words: Autonomy; Name; Gender attribution Social-robots; Human-Robot Interaction

ÖZET

OTONOMİ VE ROBOTONOMİ: TOPLUMSAL CİNSİYETİN ROLÜ

Gaye Aşkın

Psikoloji Yüksek Lisans

Tez Danışmanı: Dr. Burcu Ayşen Ürgen

Haziran 2023

Bu çalışma, sosyal robotlarda toplumsal cinsiyetin insanların özerklik algısını nasıl etkilediğini incelemektedir. Bir varlığın adının (erkeksi / kadınsı / nötr / teknik / isimsiz) ve türünün (insan/robot) varlık türünün, eylemsellik düzeyini, insani özellikler atfedilmesini, yetkinliğini ve cinsiyet değerlendirmelerini etkileyip etkilemediğini ve bu değerlendirmelerin bilinç dışı düzeyde de gerçekleşip gerçekleşmediğini araştırmaktadır. Çalışma, 150 katılımcının yer aldığı bir ön çalışmadan ve bir ana çalışmadan oluşmaktadır. Ana çalışmada katılımcılar 4 saniyelik 18 video izlemektedirler ve otonomi ile ilgili soruları cevaplamaları gerekmektedir. Ayrıca çalışmanın temalarından olan: özerklik, varlık türleri ve toplumsal cinsiyet kategorilerine göre uyarlanmış üç örtük ilişkilendirme çalışmasını da tamamlarlar. Açık kısım için varlık türünün, adlandırma ve cinsiyet etkilerini analiz etmek için ANOVA analizi yapılmıştır. Örtük kısımda, tüm katılımcılar için d-skoru hesaplanıp sonrasında gruplar arası karşılaştırma yapmak için bir ANOVA analizi yapılmıştır. Ön çalışma ve ana çalışmalarda cinsiyet atfını belirlemek için regresyon analizi yapılmıştır. Açık-örtük bölümlerin ilişkili olup olmadığını belirlemek için korelasyon analizleri de yapılmıştır. Son olarak açık kısımda nitel girdiler için tematik bir analiz yapılmış ve dokuz temaya ayrılmıştır. Çalışma, eylemsellik ve varlığın temel etkisinde varlık düzeyinde anlamlı sonuçlar vermiştir. Yetkinlik ve cinsiyet atfında, varlıkların tüm isim koşullarında ana etkisi bulunmamıştır. Örtük kısımda, kadın ve erkek katılımcılar da otonomi ve cinsiyet IAT-1'de farklılık göstermiştir. Diğer iki IAT için, kadın ve erkek katılımcılar da benzer şekilde yanıt vermiştir. Çalışma, isim manipülasyonunun insanların özerklik algısını etkilemediğini, aksine varlık türleri ve eylem özelliklerinin onları etkilediğini göstermiştir. Ayrıca insanların örtük ve açık cevapları birbirini yordamamaktadır.

Anahtar Kelimeler: Özerklik; İsim, Cinsiyet atfı Sosyal robotlar; İnsan-Robot Etkileşimi

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Asst. Burcu Ayşen Ürgen, for her invaluable support of my interest and for providing me an opportunity to conduct research and make my interest to be hearable from my undergraduate year to my master's program. Her expertise and encouragement helped me to complete this research and author this thesis. I have the great privilege of having Asst. Seçil Gönültaş and Assoc. Cengiz Acartürk who are members of my thesis committee and provided me with valuable feedback and suggestions. Their insights and guidance were invaluable in shaping my research and authoring this thesis.

I would like to express my special thanks to Hilal Taşkan (or prospective Dr. Taşkan), who always supports me, values my effort, and encourages me with her friendship, care, and being my family here. I am sure she will always be beside me. I am grateful to dearest Deniz Satıcı, who always believes in me profoundly and makes me feel her love and support in every sphere of my life. I would like to sincerely thank Ismail H. Bayer and H. Papatya Yeşilbaş, who put so much effort into data collection, which helped me go a step further in my thesis. I would like to thank Tuğçe Pekçetin, Şeyda Evsen, Tuvana Karaduman, Gökçe Şahin¸Onur Sunay, and Nihal Gökçe for their listening with patience and sincere sharing. I sincerely thank Deniz Erdil who supports me during preparation of one of my stimuli. Thank you very much, İmge Saltık, who is always my co-author in research and life. I am sure we will have (in-depth) conversations and (quant and qual) analyses of our professional and personal lives where we will not understand how time passes.

I am grateful to my family for their endless support and love. My lovely mother, Dilek Aşkın, always encourages me wholeheartedly; my dearest father is a lifelong advisor who always supports me tenaciously. My brother Dr. Anıl Aşkın and my sister Dr. Fulya Pınar, my deepest thanks to them for showing me other aspects of life with all their realities and how it can be lived in another way and for reminding me often that my happiness and psychological health are more important than anything else.

Umut Can Hekimoğlu, if I am a happier person today if I am more successful, and if I was able to finish this thesis today, your support and love are great. Thank you for giving me courage, being a partner in my struggle, and always listening to my decisions. Thank you for always being there for me, no matter the distance.

I was inspired by the fact that the subject of gender and autonomy concerns the society I live in, my experiences, and the processes my country has gone through while writing this thesis. Even though the subject is robots, the topics of gender, autonomy, and their intersection answered many questions about humans and social issues and encouraged me to ask even more questions. I would like to thank myself for continuing without ignoring and learning after all the adverse social and personal events. I could not have undertaken this journey without you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

1. INTRODUCTION

In my thesis study, I will investigate the relationship between gender identification and the concepts of autonomy in humans and robots. I aim to conduct this study in two parts: In the first part, I will explore these two concepts by explicitly asking people's opinions; in the second part, I will use the implicit method to understand their unconscious association concerning these concepts. In the explicit part, I aim to manipulate gender by naming human and robot agents (using gendered names, neutral-gendered name, no-name, and technical name). This manipulation aims to measure how people’s autonomy attribution is changed across different named agents (human and robot). This measurement is taken after participants watch various actions performed by humans and robots. In the implicit part, I use Implicit Association Task (IAT) to understand people’s implicit biases concerning gender, autonomy, and robots.

I hypothesize that having a state of independence, the capacity to do things, and being competent about doing things may not be enough to be evaluated as autonomous, and “gender” can be a factor modulating it. Depending on who you are in terms of your gender identity, evaluation of your capacity and ability to do something may be subject to change. On one side, I would like to answer the gender-autonomy relation; on the other, I would like to answer whether this relation can also be applied to robots. On the one hand, gender studies show how gender influences the acceptance or rejection of a person's autonomy, and on the other, robotics studies indicate how embodied agents can have the capability to act freely. The different literature shows that gender is a key factor affecting people's autonomy evaluations. Also, gender studies' integration into the human-robot interaction studies shows that gender is a determinative factor in the robot's occupations and emotional intelligence in line with human-human interaction.

The main questions I would like to address in this thesis are: Do we think robots have autonomy? Does the assigned gender of a robot influence this judgment? The scope of these questions, the background of autonomy, and gender will be covered in this chapter.

1.1 Autonomy

I would like to provide a definition and background information to understand the rationale of the autonomy approach in this paper. Firstly, it is essential to define "being autonomous." According to the Merriam-Webster definition of autonomy, it is the ***state of existing or acting separately from others*** [4]. Basically, it means that a person who acts independently from others (influences, controls) has autonomy. A great deal of literature has also discussed this concept, such as gender studies, political science-war studies-, and robotics studies.

Autonomy is a key component of human-robot interaction (HRI). In human-robot interaction, various levels of robot autonomy (LORA) play a significant role [23]**.** That is why HRI should be understood by identifying the factors that influence the perception of robot autonomy. Before channeling the topic into the robot's autonomy, it is important to discuss how I integrate autonomy discussions in various kinds of literature into my research question.

With this definition in mind, I will summarize these different sets of literature discussing autonomy concepts and emphasize the intersection points of these works of literature with the application of my research question.

1.2 Autonomy and Gender Relation

Gender literature discusses autonomy from a feminist perspective concerning participation in the public sphere [37]. It focuses on women's rights to independence, independently governing the self from the man. It means that women in the public sphere must be treated equally to their man counterparts [50]. This treatment relied on the idea that each gender should have equal rights and equity, enabling people to have ownership rights and equal freedom to act, speak and defend themselves. Another example showing how gender is crucial for one's accepting/rejecting autonomy in line with culture is honor killings. In most places in the world, honor killings, the notion that parents have the right to control whom their daughters marry, has wide acceptance due to the caste system, which requires parents to supervise and control their daughters in order to ensure they marry within caste norms [24].

Psychology literature discusses gender and autonomy relations by relying on gender stereotypes which are mental imageries dependent on classification and automatic cognitive processes [28]. These evaluations are about how women and men should be and how they should behave in different spheres of life [5]. Generally, gender stereotype research suggests that women are stereotyped as more communal (e.g., caring and interdependent) than men, while men are stereotyped as more agentic, independent, entrepreneurial, and self-sufficient [5]. Autonomy from the lens of “gender” can provide us a deeper understanding of an ability to act and decide independently(i), have the capacity to (correctly) do the corresponding thing(ii), and can be subject to change in evaluation depending on the gender of the doer. The perspective that autonomy has traditionally been regarded as “long deemed masculine.” [44]. This understanding is in line with the expectation about women who are “supposed to” act interdependently, having less autonomy, whereas men are the ones who have agentic natures and are independent [7]. One research examines how the positive characteristics of women-those promote their communal virtues-conflict with Western societal norms leading to the unequal psychological burden on women compared to men. This study offers a novel explanation for why seemingly positive stereotypes may elicit negative reactions-a diminished sense of autonomy, especially for women [45]. These works of literature show that gender is one of the essential components of accepting/rejecting one’s autonomy, even if it is used as a strategy to deny personal autonomy concerning gender. What needs to be considered is that gender is a key factor in evaluating a person as (in)dependent, so when it is thought of from the social robotics perspective, it raises the question of whether the perceived gender of a robot affects the people’s autonomy evaluations about the robot.

In the next part, I would like to summarize autonomy discussions in robotic and HRI literature, which will enable us to set the relation between gender literature and robot literature autonomy discussions in the further parts of the thesis.

1.3 Autonomy in Robotics Studies

Autonomy is discussed in another field, robotics studies, focusing on robot autonomy rather than humans. This literature applies the automation concept to autonomy by relating the robot’s ability to monitor, produce strategies, achieve goals, decide on an option, and act upon it (goal) [6] by using advanced technologies such as automatic translation, image recognition, and medical diagnosis [49]. It can indicate that a robot demonstrates autonomy due to its ability to perceive and learn the operations and, if required, modify them. The concept of goals included the robot's ability to analyze information, generate an optimum decision, and act accordingly [6]. So, the robot’s automation translated into the robot’s ability to act as an ‘agent.’ An ‘agent’ (autonomous agent) can interact with their environment and work towards their goals independently without immediate assistance/intervention from humans [15].

Another path of robotics literature is human-robot interaction (HRI). In this area, the robot refers to the social robot, an entity with a human-like appearance and abilities (face shapes, limbs, ability to talk, capability to lead a conversation, show emotions, and move like a human). They are called social because they are "relational artifacts" which engage their human partners in intentional social interaction [20]. This engagement changes our perception of social robots because they are not simply tools but agents with “whom” we interact naturally and intuitively [8].

Goodrich and Schultz [23] categorize the human-robot interaction into five categories: level of autonomy to transform the environment into representations and acts(i), creating a communication flow between humans and robots (ii), building a construct in which humans and robots have various roles/tasks(iii), learning and training techniques for artificial intelligence as they pertain to adaptation, learning, and training(iv), change in the way you complete a task when new technology is applied (v). By benefitting from Goodrich and Schultz [23]’s categorization, Rau et al. [47] investigated how human decision-making was affected by a social robot's level of autonomy (high and low) in different group levels (in-group and out-group). In their study, researchers found compelling evidence that highly autonomous robots influence human decisions more than low-autonomous robots. Also, the interaction between autonomy and group orientation was found to be significantly greater when the robot was a member of an ingroup rather than an outgroup member. These show that a robot's level of autonomy can change/ improve its influence on human-robot interaction. This provides an insight into how socially oriented design is required, which is why including social factors such as gender would be helpful in evaluating autonomous agents.

1.4 Intersection points

In the previous sections, we discussed how autonomy concepts could differ and coincide with three pieces of literature. Gender studies show how gender influences acceptance/rejection of a person's autonomy, and robotics studies indicate how embodied agents can have the capability to act freely without human intervention. So, when we think of intersection points of these works of literature, there can be two questions: Does gender also affect the robot's autonomy when designed/called gendered? And Do people reject/accept robots' autonomy as they do humans? To answer these questions, the following section will elaborate on how other factors can contribute to autonomy evaluation concerning gender and how it can be measured.

1.5 Operationalization of Autonomy Concept

In the abovementioned discussions, the autonomy concept has not yet been transformed into a measurable form. That is why it needs to be turned into a question form providing information about people's thoughts and attitudes. Regarding this, four main items provide an opportunity to do under the umbrella of autonomy. These four questions are the explicit questions that are going to be asked conscious report of participants. These explicit questions are about: level of agency (1), level of competency (2), and personification (3). In this thesis, I operationalize the concept of autonomy in three dimensions.

1.5.1 Level of Agency

Firstly, one must have some *capacity* to be autonomous, and this capacity refers to the *agency*. A concept of agency is the ability to act independently and the conditions to do so [2]. It means that I may have autonomy as a state of freedom, but I do not necessarily have the capability to do so.

The concepts of autonomy and agency include "acts, actions," so I used actions as a foundation for measuring both. The construal-level theory (action identification theory) [55] is one theory that aligns with these concepts. Actors are classified as high or low construal depending on whether superordinate goals are associated with their actions. The theory relies on the notion that agents' actions are determined by their high construal constructs, whether they possess autonomy, goal setting, the ability to derive meaning, or intentions [33]. For example, they found that when artificial agents appear to have learning capabilities, they indicate whether they are intelligent, etc. If people program them so that they can learn, their choice of whether to act at a high or low construal level may be affected [33].

In this thesis, I used the Behavioral Identification Form (BIF), which includes actions with these two levels of the agency developed by Vallacher and Wegner [57]. The BIF actions are in the written format that participants read and answer accordingly. However, rather than using them in written format, I animate these actions, and participants watch and then answer according to two levels of agency (high and low).

1.5.2 Level of personification

Assessing humanness can also constitute a path to autonomy and agency. As discussed earlier, individual autonomy refers to the right to control one's life. Freedom of choice allows people to form their identities (Oviatt, 2021). This implies autonomy and agency in acknowledging and reflecting upon one's identity, desires, intentions, beliefs, emotions, and responsibilities [40].

Likewise, according to Harris & Fiske [27], individuals' belief in an inner life, a sense of responsibility, and the capacity to control their situations were closely related to their autonomy and agency. In contrast, dominance, subduing and control lead to the degradation of human qualities, which can result in dehumanization [40]. One study found that participants were more likely to attribute human-like qualities to the robot when it performed a gender-typical task. This tendency to attribute human-like qualities to robots may be due to people's greater familiarity with gender-typical tasks and the expectation that certain tasks are more suited to specific genders [34].

I concluded that what I am referring to here as *personification* is essentially what these "reflections" and "evaluations" (desires, beliefs, emotions) represent. As part of the personification process, inanimate objects are given human characteristics to create the illusion that they are living and feeling entities [35]. Therefore, it is more likely that robot agents who also exhibit human-like traits will likely be attributed autonomy.

Regarding these, the next question I ask is to understand how people personify agents (names, personality traits, etc.) because evaluating an agent by asking this question carries many implications. This question type is open-ended so I can determine the frequency of attributes related to autonomy that are similar to human characteristics.

1.5.3 Competency

In the discussion of autonomy in robotics and psychology literature and then agency, one of the important dimensions is mentioned, which is competence. Competence plays a significant role when discussing autonomy because, as Meyers [42] suggests, for an individual to be autonomous, they need to possess skills and attitudes(competence) that enable them to perform diverse tasks. One study investigates the competency of social robots from a gender perspective. In this study, participants interacted with a robot that performed either a gender-typical task (cooking) or a gender-atypical task (carpentry). The study found that participants rated the robot as more competent when it performed a gender-typical task compared to a gender-atypical task. Additionally, participants were more likely to ask the robot for help with a gender-typical task than a gender-atypical task [34].

This definition of competency relates to the ability to carry out an action (functionally) correctly. Doing the action correctly means the following. For example, when we want to drink water from a glass, if we do not adjust the distance between the glass and our mouth properly, the water will spill over us. This means that the action is not done correctly as it should be. Another example is when we want to start the car; if we press the brake instead of the accelerator to move the car forward, the car will not move. This is an example of a wrongly executed action. In summary, if the action is done correctly, it reaches the purpose of the action, but if it is not done correctly, it does not reach that goal.

Therefore, in this thesis, I explicitly ask people to assess whether the agents' actions were functionally correct.

1.6 Gender Attribution

Lastly, gender attribution is the fourth explicit question in this first part. This question is not directly related to the measurable form of autonomy, but it is a factor that is going to help to evaluate relationships with the agents' autonomy evaluation and gender attribution to these agents. Those mentioned earlier and present studies indicate that gender affects people's evaluations and perceptions of autonomy. It is an effective component because gender is part of identity formation processes. An agent's gender potentially affects the evaluations of the person's capability, state of independence, and ability to do something. As Feltman and Elliot [19] state, competence perception influences achievement outcomes, whether successful (success) or unsuccessful (failure). It has traditionally been thought that gender affects a person's perception of competence (e.g., gender differences in man-dominated fields such as STEM) [58]. Lastly, I asked how people would evaluate the agent to measure gender attribution**,** whether they attribute gender based on name or not.

1.7 Name Manipulation

According to all discussions above, I mentioned how autonomy and some factors could interact. I touched upon the measurable form of autonomy, and the next step of this discussion is to ask what gender manipulation is.

Generally, autonomy is based on the point of origin from within the self [25]. One element that creates a base for the ‘self’ is the name of the agent/person. A person's name can be considered a (semantic/verbal) label that identifies and differentiates them from others [10]. The name we are born with is our first social tagging. Names are used to identify gender in many languages and cultures, including given names, family names, and unofficial names [1]. According to Zwebner et al. [59], names are collections of behaviors, characteristics, and attributes that share meaning and a shared schema.

For example, depending on the semantics of the name, gender is assigned in accordance with gender stereotypes. In Turkish, the man’s name Yılmaz is etymologically derived from a word meaning 'fearless,' thus fitting a stereotype of masculinity. For example, women's names like Gül ('rose') are associated with beauty and thereby tied to the femininity stereotype [1]. Moreover, implicit biases are prevalent and can be easily discerned from a name. This makes it logical to keep us from seeing those names when making important decisions like marking, hiring, and refereeing when we are likely to be influenced by them. Researchers have found that CVs with white or man names tend to be considered better than CVs with nonwhite or women names, and this tendency correlates with implicit bias [31].

The psychology literature indicates that since name-giving includes broad selection processes, some of these are individuals' unusual name preferences, gender-appropriate or non-gender-appropriate names, nicknames, and names that are thought to be representative of the self [32]. Name selection is given value because the name represents the individual (or another object) and is responsible for triggering associations with them [38].

1.8 Social Robots and Name

Social robot literature can correlate name usage where more space is needed. One study demonstrated that a member's choice of a German ('ARMIN’) versus Turkish ('ARMAN') first name had an impact on their group membership [16]. A study shows that within-group members viewed the robot more positively and anthropomorphized it more strongly than those out-groups. Researchers found that robots with man names were assigned to security occupations, whereas robots with woman names were assigned to healthcare occupations [53]. This proves gender stereotypes are related to specific occupations. Another study shows that a name is a helpful cue to activate the gender stereotypes of participants. In this study, robot pictures were shown to the participants in a mock marketing survey where participants were asked to name the robots, rate their potential features and characteristics, and assign tasks to them. Older participants and those with a more feminine self-image showed a more pronounced preference for man names. Participants assigned technical tasks to the robot, rated higher in agency and low in communal attribution but less prone to including this robot in their lives [36]. Based on these pieces of literature's emphasis on name and identity relations, I thought that name could be an effective manipulation for activating people’s gender stereotypes. So, in this thesis work, I used gendered names that strongly fit feminine and masculine and neutral names.

1.9 Implicit Measurement of Autonomy, Gender, and Robot

I have discussed the concepts, their counterparts in different works of literature, and the tools I have transformed into a measurable form. The second part of the study covers a measurement that will bring novelty to the field at least as much as the first. In this part, I will use a method based on the matching speed (reaction time) of words related to the concepts of gender, autonomy, and robot/human, as mentioned above, which is called an Implicit Association Task (IAT).

IAT focuses on whether a concept (e.g., black people) is associated with either an evaluation (e.g., good, bad) or stereotype (e.g., clever). People may have an implicit preference for certain concepts when they categorize words faster-associated evaluation/stereotypes share a response key than other concepts and evaluations group [28].

I am motivated to use this measurement because several studies have shown that combining explicit and implicit measurements will yield more comprehensive results [11, 29,30]. One of the important points of this scope is that these two measurement methods can provide rich information about people’s evaluations. In other words, the person's conscious responses and unconscious responses may differ from each other. So, I designed 3 Implicit Association Tasks to see how these concepts match unconsciously. In these IATs, there are three main categories (independent/dependent, robot/human, and woman/man), and create this measurement with the words associated with these categories.

I aim to determine whether the concepts of robot/human, woman/man, and dependent/independent in people's minds coincide with the findings presented by different kinds of literature when they come together. For example, we have seen in gender studies and political science literature that women and femininity are associated with a lack or little independence; or it was emphasized that the design of robots by humans correlates with their dependence. Finally, this study, which stated that the robot was more masculine [39], raised the question of whether this happens at an unconscious level and how these answers do/not match with the explicit part’s answer.

1.10 Gender in Autonomy and Robotonomy (Robot+Autonomy)

Various aspects of social robots have been analyzed to determine gender attribution (e.g., physical characteristics, tasks, and actions). While those features may be indicative of gender, further research is needed to determine what concept(s) can influence one to attribute gender, capability, and state. For instance, autonomy is an important concept because what the person chooses and does must somehow (partly) be determined by the distinct identity [22]. In this case, the identity is gender, and I closely examine how autonomous acts concerning gender identity potentially affect the evaluations of the agent’s autonomy. In the literature, several studies show how robots’ autonomy (robotonomy coined by me) affects people’s interaction with robots; however, autonomy components have not yet been extensively studied (with gender).

1.11 Importance and Contribution of Studying Autonomy and Gender

It is important to study autonomy, gender, and robots because technological tools have been produced and designed based on people's (gender) stereotypes. These gender stereotypical representations may pose some issues that need to be considered. For example, sex robots designed as pornographic figures aiming to be sex partners for humans have also become prevalent in online channels [12], which have the possibility to emerge serious problems (e.g., consent and sex). Some kinds of robots need a certain level of functional autonomy, learning abilities, physical support, and the ability to adapt to their surroundings (e.g., autonomous vehicles, medical robots, drones, and robotics for human repair and enhancement). Sex robots, which are humanoids, are not included on the list of examples, but they are being developed and attracting much interest. It is presumed that sex robots can be autonomous, learn, be supported physically, and adapt to their surroundings [21]. In the discussion of consent, it is known that an individual unable to give/withhold consent to sex acts (e.g., a child) is prohibited from performing them [41]. In addition, if a person is capable of giving consent but does not perform a sex act on them, this is illegal and an act of rape. Accordingly, the concept of consent might be a viable way to incorporate robots into sexual communities.

The first reason is given that heterosexual men mostly drive the sexbot industry; they are looking for realistically accurate robots that mimic feminine appearance and behavior but interact in a way that promotes harmful pornographic stereotypes of feminine appearance, personality, and sexuality [48]. According to her, sexbots are used and sold to express and reinforce persistent sexist and misogynistic stereotypes about women, resulting in harm to women. However, even if sexbots are not harmful to women, it may be wrong for individuals to interact with sexbots in a way that would harm them if they were women [14]. Sinziana Gutiu [26], a feminist commentator on sex robots, emphasized that humanoid sex robots will erode many people's attitudes toward consent in human-human sexual interactions. According to Gutiu [26], if consent is irrelevant in the robot case, it may lead a significant group of people to believe that consent is not needed from human sex partners.

The literature discussing autonomy and agency concepts will provide a base on how these can be applied in robots. Understanding one's capacity and ability to act independently provides a fundamental understanding of that person's abilities/ capabilities to be autonomous. That is why using BIF provides a tool for creating a stimulus set for our agents (human and robot). Implementing BIF is realized by transforming the written actions in that form into animations performed by the agents. By doing this, we would have the opportunity to measure how participants assign the level of agencies to these agents and ask about the other dimension of autonomy (doer of the tasks and functional appropriateness of the performed action). The last question is asking about the gender of the agent because, in the literature, we were informed that gender is the main component of accepting/rejecting one's autonomy.

Furthermore, the implicit part provides us to understand whether the participants' explicit results about evaluation agents' autonomy in relation to gender manipulation share common grounds with the implicit measurement autonomy/gender/agents' associations. It creates room to compare whether people's general autonomy attribution to gendered name agents will be in line with the dependent/independent association of women/men and robot/human-related words.

Since no study integrates the autonomy concept with gender in human-robot interaction, the robot's design, functionality, and capacity are important autonomy-related parameters of how people will behave toward the robot**.** In the context of this thesis, I mainly focus on the levels of autonomy and gender identity in relation to humans and robots performing different actions.

My research question will try to answer these two questions: Does gender affect a robot's autonomy when they are called gendered? Do people reject/accept robots' autonomy as they do humans? Based on these, my main research question is: Do ***agency attribution of actions***, ***level of personification***, ***level of competency*** and ***gender attribution*** depend on the type of the agents (<u>human</u> and <u>robot)</u> how they are **named?** The second research question is its scope to understand the unconscious associations of related concepts. The question is, how are ***Autonomy***, ***Gender***, and ***Agents*** concepts associated in people's minds, and how do these relate with explicit answers?

1.12 Hypotheses

a) Explicit Part

1. BIF Evaluations:

a. Hypothesis 1: People attribute more high-level agencies to human-named agents than the technical name/No-name agents.

b. Hypothesis 1.1: People attribute more high-level agency to masculine name agents than feminine name agents.

c. Hypothesis 1.2: People attribute higher agency to the human agent than robot agents in human-named and no-name conditions (except technical ones).

2. Level of personification:

d. Hypothesis 2: People prefer personifications like name and personality traits in both agents in more human-named conditions than the technical name/no name.

e. Hypothesis 2.1: People prefer personifications like gender and type (human/robot) in both agents more technical-name/no name than human-named conditions.

f. Hypothesis 2.2: People prefer personifications like name and personality traits to human agents over robot agents in human-named and no-name conditions.

3. Competency:

g. Hypothesis 3: People attribute high-level competence to human-named agents than the technical name/No-name agents.

h. Hypothesis 3.1: People attribute more competency to masculine name agents than feminine name agents.

i. Hypothesis 3.2: People attribute higher competency to the human agent than robot agents in human-named and no-name conditions (except technical ones).

4. Gender Attribution:

j. Hypothesis four. People attribute more gender to human-named agents than the technical name/No-name agents.

k. Hypothesis 4.1. People attribute equal gender to masculine name agents than feminine name agents.

l. Hypothesis 4.2. People attribute less gender to neutral name conditions than gendered name conditions.

m. Hypothesis 4.3. People attribute neutral/less gender when both agents have a technical name.

b) Implicit Part

1) IAT-1(Woman/Dependent-Man/Independent)

n. Hypothesis 1: Man, participants' reaction time is faster in the woman/dependent-man/independent category, whereas woman participants' reaction time is faster in the woman/independent-man/dependent category.

2) IAT-2(Human/Dependent-Robot/Independent

o. Woman and man participants' reaction times are faster in the human/independent robot/dependent category.

3) IAT 3 (Human/Woman, Robot/Man)

p. Woman and man participants' reaction times are faster in the Human/Woman, Robot/Man categories.

CHAPTER II

2. METHODS

2.1 Participants

According to G -power analysis, the number of groups which refer to the between subject factors gender-2 levels and name 5 levels and number of measurements to indicated repeated measures which are agents-2 levels. α err prob, I chose 0.05 and power is %95. The total sample size is given as 110, but I increased the number participants to decrease variability. So, 150 people participated in this study they are mostly the college students. Participants were asked about their age, gender identity, and department/field they study/work as demographic information (Mean Age=20.93 (18-29), SD=2.34; 75 Women, 75 Men, and 1 Other (self-defined)). Unfortunately, I had to exclude the “Other” participant due to an insufficient number (n=1). The Human Research Ethics Committee of Bilkent University approved the study. A consent form was signed by all participants and the flow of the experiment was explained to them before the experiment began.

2.2 Stimuli

There are two types of stimuli in this study. The visual stimuli consisted of 9 short videos, each lasting 4-sec. A gender-neutral agent (human and robot) performed an action in each video. These actions were animated based on the actions placed in the BIF questionnaire. The selection criteria were: 1) one actor should be present in the scenario2) the action is a solid action that can be easily visualized and animated, unlike abstract actions, e.g., thinking.

BIF actions: The videos were taken from BIF. The actions have two levels of agency (high and low). Among the 25 actions, I animated nine (using two-level agencies) using the neutral-gender-looking robot and human. The reason I chose 9 is that other actions include biological processes such as eating and other agents such as a dentist. Also, some of the actions are hard to animate in a short video that would be used as stimuli, such as cleaning the room, which is why the most suitable ones are the 9 actions (watering plants, growing a garden, chopping a tree, painting, reading a book, making a list, ring the bell, locking door, washing clothes). In addition, these actions needed to be revised based on their context and our goals; for example, washing clothes includes washing, but there shouldn't be another technological tool since the robot is also a tool. Instead of it, I animated robot/human agents as if they were rubbing the clothes. Then, the painting action was written as it is executed in a room; however, I value that all actions should take place in the outside environment to eliminate the context effect, so I changed the wording of “room” in the original BIF task to “home.”(See Figure 1) The high level equals 1, and the low level equals 0, as a value assigned by Vallacher & Wegner [57]. Since I selected 9 actions, the maximum value participants can take equals 9, and the minimum is 0.

Gender-neutral robot: The robot agent was taken by the article [3].

Gender-neutral human: The human agents were determined by a pilot study. In this study, 141 participants (99 Women and 42 Men) were asked to rate the gender of human agents on a scale. (-3) and (+3) where [-3 0) means masculine, (0 +3] means feminine, and 0 means neutral. I chose the robot which was rated neutral-close to the zero (Mean= -0.33, Std.:1.5)

2.3 Design

I have a mixed design: the between-subject refers to the naming group, which has 5 levels and the gender of the participants (woman and man), and the within-subject refers to the type of the agent (human and robot).

There are three independent variables in the study. The first is the naming which has 5 levels. In the five names, the gendered human names were selected based on the frequency of names used in Türkiye in data of the Ministry of Interior [52]. The second independent variable is the gender of the participants, man, and woman. The third is the agent: robot and human.

I have 4 dependent variables in the explicit task: Level of Agency, Level of Personification, Level of Competency, and Gender Attribution. These are 4 questions that are asked to measure autonomy. Initial question is the level of agency, I asked the BIF questions, in which each question has an answer corresponding to one of the two levels: high level gets 1 point, and low level gets 0 points. Secondly, it is the level of personification, I asked the "The performer of this action is…." (Not the exact translation) (Original: Bu eylemi gerçekleştiren .... 'dır/tır/tur/tür) question. It is an open-ended question; participants are free to write whatever they think fits. Third question is the competency level question is "To what extent was this action performed properly (correctly)? (Original: Bu eylem ne derecede olması gerektiği gibi (doğru) gerçekleştirilmiştir?). The rating is between 1 and 5. (1: very bad 2: bad 3: Neither bad nor good 4: good 5: very good). Lastly, for the gender attribution question, I asked, “How would you rate Ayşe/Ahmet/Deniz/QX15/Character?”. (Original: Ayşe’yi/Ahmet'i/Deniz'i/QX15'İ/Karakteri nasıl değerlendirirsiniz?)

Figure 1. Each action is represented in single frames to show the context and

the agents.

Besides the explicit task, there is also the implicit part which includes three types of The Implicit Association Test (IAT) IAT-1 includes independent/dependent and woman/man categories and their related words. IAT-2 includes independent / dependent and robot/human categories and their related words; IAT-3 includes woman/man and human/robot categories and their related words.

2.3 Procedure

The study consisted of two parts pre study and main study. Explicit parts were conducted via desktop in the Psychology labs at Bilkent University. In the pre study participants were sent 9 actions in a written format and asked them to rate on a masculine [-3] and feminine [+3] spectrum. The second part of the study, i.e., the explicit part, was conducted through Qualtrics. Participants watched the 9 actions in an animated video format different than in pre study, now they watched 18 videos 9 is done by human other same 9 actions done by robot agent. These videos in this part, including BIF actions. These videos are animated on Unity. After watching each video, participants answered the four questions (See Figure 2).

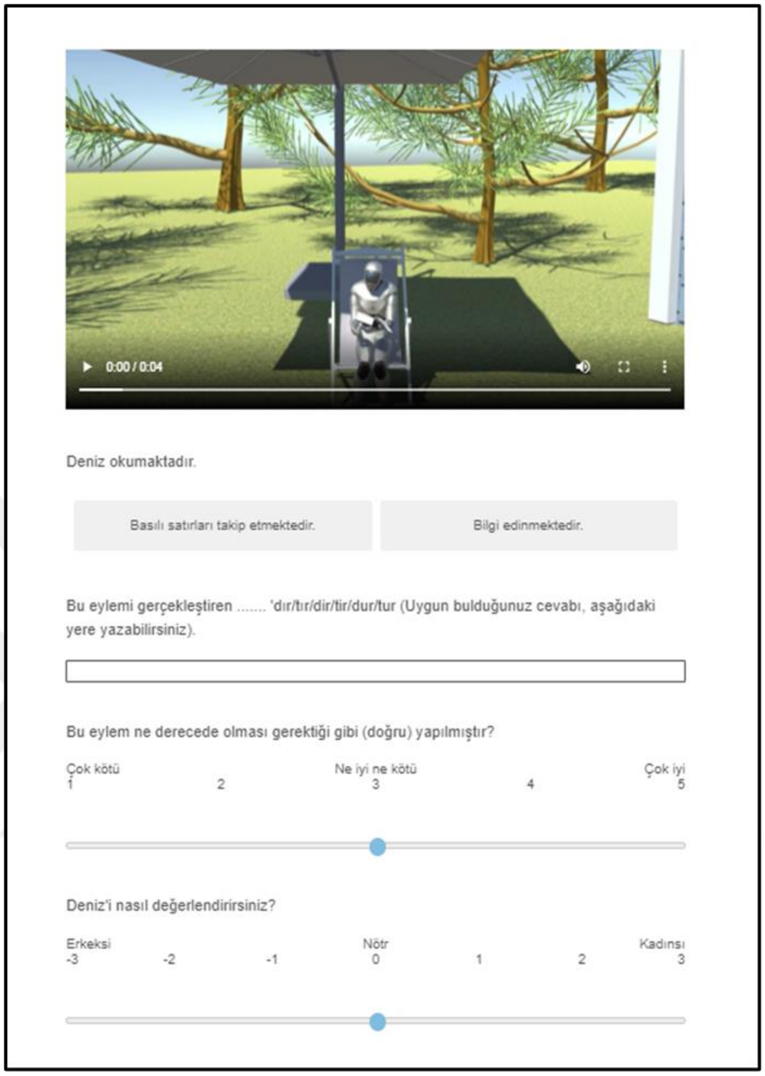

Figure 2. The flow of the explicit part of the experiment

Participants were asked to complete the three IATs in the second part, i.e., the implicit part. IATs were programmed in MATLAB-Psychotoolbox. Before starting the IAT, participants were trained in words and their corresponding categories. After practicing for approximately 2-3 min, I started the IAT. The order of the three IATs was changed for every participant, and I used the Latin square method for this order. It works as follows: For the first participant, the order of IAT is IAT-1 will be shown first, then IAT-2 is second, then IAT-3 is in third place (1,2,3). For the second participant, IAT-2 will be shown; first, IAT-3 is second, IAT-1 is in third place (2,3,1), and IAT order is (3,1,2) for the third participant. For the fourth participant, the order is again 1,2,3.

2.4 Data Analysis

Both quantitative and qualitative tests were run to measure whether people attribute autonomy to gender-neutral robots and human agents based on gendered name manipulation.

2.4.1. Quantitative Analysis

For the explicit part, I conducted 2 (Gender) x 2 (Agent) x 5 (Naming)repeated measures ANOVA on the overall BIF scores, FA, and Gender evaluations. Then, I ran a separate 2 (Agents) x 9 (Actions) repeated measures ANOVA on BIF, FA/competency, and GA/gender attribution evaluations for each naming condition separately (thus, 5 ANOVAs in total).

For implicit measure analysis, I calculate D-score using the online D-score app developed in R [46]. An MIT license is provided with D-scoreApp, making it platform-independent [13].

To understand how each implicit test differs across groups and gender, I conducted two-way 2(Gender) X 5 (Naming) ANOVA. To investigate the correlation between implicit and explicit measurement, I conducted a correlation test for these parameters (IAT-1, IAT-2, IAT-3, and BIF, FA Gender). Lastly, to check the manipulation whether the name affects the gender manipulation or the actions' gendered notion influences participants' gender perception of the agents, I conducted a regression analysis based on the pre-test answers of participants and gender attribution in the explicit part. Through this analysis, I aimed to understand whether participants' ratings on actions' gendered level do (not) predict the gender attribution to the agents.

2.4.2. Qualitative Analysis

The open-ended part is analyzed in two parts; first, I created a theme based on frequency. For example, if the name of the agent is frequently written, I put them in a theme, "Identity," indicating people attribute names in relation to identity to these agents. Another example is that if people assign roles (host) or occupation (painter), I put them in a category called "Roles&Occupation." The total number of inputs for 150 participants is 2700, meaning that 2700 answers were analyzed based on the participants' evaluations. I ended up with 8 categories: Identity, Personality Traits, Gender, Roles & Occupations, Physical State, Type, Actions, and Other. Identity refers to calling the agents by their names; personality trait means attribution of traits as if they have personas; gender means attributing gender (feminine, masculine, or neutral spectrum); Roles& Occupations refers those attributing different roles (e.g., family) occupations (e.g., gardener). Physical state means evaluating the agents as they are experiencing physical states (e.g., tired, healthy). Type refers to whether the agents are robots or humans. Actions mean emphasizing the agents based on the actions they performed (e.g., planting, painting). Lastly, the other category refers to the answers that do not fit the abovementioned categories.

CHAPTER III

3. RESULTS

3.1 Comparison of groups

3.1.1 BIF rating

According to 2 (Gender) X 2(Agent) X 5(Name) omnibus ANOVA analysis, there was a main effect of agents F (1, 140) =25.052 $p<0.001$, η^2 =0.152, meaning that people rated human and robot agents with different agency levels (Figure 3). The post hoc analysis shows that the human agent is attributed a higher level of agency compared to the robot (*MD*= 0.101 SE=0.020, t=5.005, p<0.001). There was no significant main effect of gender F (1, 140) =0.555, p =0.458 and naming F (4, 140) =0.815 p =0.518.

There are no significant interactions among Agents and Naming, Agents and Gender, Naming and Gender, Naming and Gender, and Agents (p>0.05).

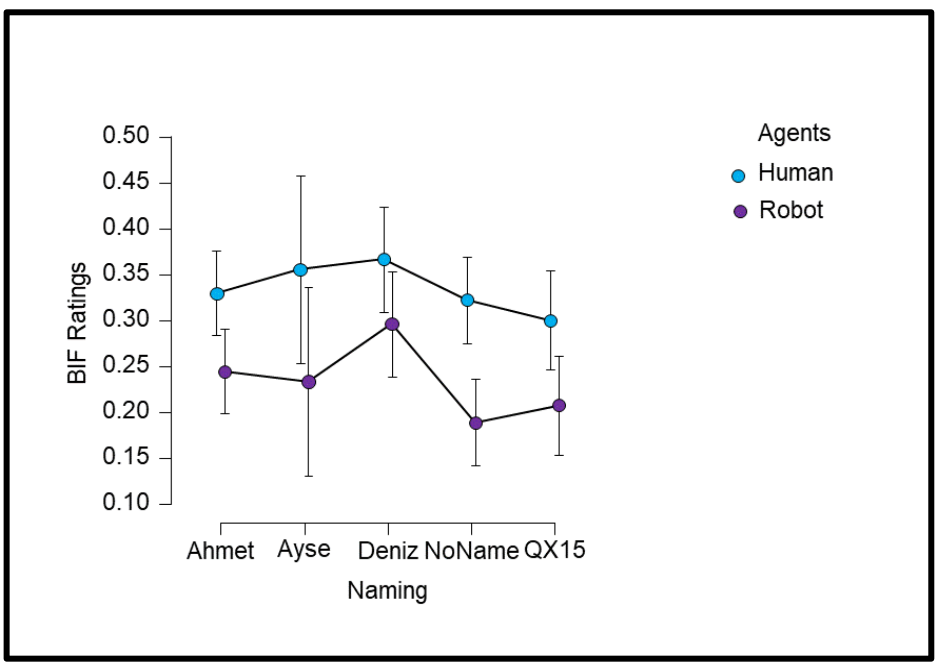

Figure 3. The comparison of BIF scores across all name groups by human and robot agents

3.1.2 FA/Competency rating

According to 2 (Gender) X 2(Agent) X 5(Name) ANOVA analysis, there was no main effect of agents *F* (1, 140) =0.990 *p*=0.321. There was no significant main effect of gender *F* (1, 140) =2.373 p =0.126, and naming *F (*4, 140) =0.508 p =0.730. There were no significant interactions among Agents and Naming, Agents and Gender, Naming and Gender, and Naming and Gender and Agents (p>0.05).

3.1.3 Gender Attribution

According to 2 (Gender) X 2(Agent) X 5(Name) ANOVA analysis, there was no main effect of agents *F* (1, 140) =0.889 *p* =0.347. There was no significant main effect of gender *F* (1, 140) =0.060 p =0.808, and naming *F (*4, 140) =2.363 p =0.056. There were no significant interactions among Agents and Naming, Agents, and Gender, Naming X Gender and, Naming and Gender and Agents, (p>0.05).

Table 1

The summary of results of all groups. The significance level is p<0.05; significant results are shown in bold.

	Main Effect of Agent	**Main Effect of Gender**	**Main Effect of Naming**	**Agent x Gender Interaction**	**Naming x Gender Interaction**	**Agent x Naming Interaction**	**Agent x Naming x Gender Interaction**
BIF	F (1,140) = 25.052 **p<.001 ***** η^2 =0.152	F (1,140) = 0.555 p =0.458	F (4,140) = 0.815 p =0.518	F (1,140) = 1.059 p =0.305	F (4,140) = 1.360 p =0.251	F (4,140) = 0.341 p =0.850	F (4,140) = 0.709 p =0.587
FA	F (1,140) = 0.990 p=0.321	F (1,140) = 2.373 p =0.126	F (4,140) = 0.508 p =0.730	F (1,140) = 0.174 p =0.677	F (4,140) = 0.381 p =0.822	F (4,140) = 0.313 p =0.869	F (4,140) = 2.330 p =0.059
GA	F (1,140) = 0.889 p =0.347	F (1,140) = 0.060 p =0.808	F (4,140) = 2.363 **p=0.056 (Marginal)**	F (1,140) = 0.176 p =0.268	F (4,140) = 0.139 p =0.968	F (4,140) = 1.313 p =0.268	F (4,140) = 0.725 p =0.576

Table 2

The BIF summary of each group. The significance level is p<0.05; significant results are shown in bold.

BIF	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
Feminine	F (1, 29) =13.716 **p <.001***** η^2 =0.196	F (1, 29) =2.295 p =0.095	F (8, 232) =1.738 p =0.093
Masculine	F (1, 29) =22.032 **p <.001***** η^2 =0.331	F (1, 29) =7.165 **p =0.012** η^2 =0.010	F (1, 29) =4.145 **p <.001***** η^2 <0.025
Neutral	F (1, 29) =22.099 **p <.001***** η^2 =0.327	F (1, 29) =3.145 p =0.087	F (8, 232) =0.728 p =0.667
Technical	F (1, 29) =20.465 **p <.001***** η^2 =0.294	F (1, 29) =6.154 **p =0.019** η^2 =0.013	F (8, 232) =0.942 p =0.482
No Name	F (1, 29) =12.992 **p <.001***** η^2 =0.220	F (1, 28) =16.749 **p =<.001***** η^2 =0.027.	F (8, 232) =1.407 p =0.194

Table 3

The FA summary of each group. The significance level is p<0.05, and significant results are shown in bold.

FA/Competency	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**

Feminine	F (1, 29) =10.533 **p <.001***** η^2 =0.211	F (1, 29) =0.011 p =0.919	F (8, 232) =1.163 p =0.322
Masculine	F (1, 29) =12.058 **p =.001***** η^2 =0.235	F (1, 29) =0.025 p =0.877	F (8, 232) =0.283 p =0.971
Neutral	F (1, 29) =17.044 **p =.001***** η^2 =0.297	F (1, 29) =0.023 p =0.880	F (8, 232) =1.781 p =0.082
Technical	F (1, 29) =16.192 **p =.001***** η^2 =0.248	F (1, 29) =1.074 p =0.309	F (8, 232) =0.706 p =0.686
No Name	F (1, 29) =13.760 **p =.001***** η^2 =0.247	F (1, 29) =0.296 p =0.590	F (8, 232) =1.096 p =0.367

Table 4

The Gender Attribution summary of each group. The significance level is p<0.05; significant results are shown in bold.

Gender Attribution	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
Feminine	F (1, 29) =16.997 **p =.001***** η^2 =0.245	F (1, 29) =2.615 p =0.117	F (8, 232) =1.262 p =0.264
Masculine	F (1, 29) =19.669 **p =.001***** η^2 =0.289	F (1, 29) =0.020 p =0.888	F (1, 29) =2.384 **p =0.017** η^2 <0.014
Neutral	F (1, 29) =13.063 **p =.001***** η^2 =0.209	F (1, 29) =0.098 p =0.756	F (1, 29) =2.468 **p =0.014** η^2 <0.016

Technical	F (1, 29) =14.003 **p =.001***** η^2 =0.208	F (1, 29) =1.466 p =0.236	F (8, 232) =1.223 p =0.286
No Name	F (1, 29) =15.433 **p =.001***** η^2 =0.249	F (1, 28) =0.810 p =0.376	F (8, 232) =0.880 p =0.534

3.1.4. Qualitative Analysis

The 150 participants' responses to 18 videos (9 human and 9 robot) which makes up 2700 responses in total, were analyzed based on 9 themes. The most frequent themes are identity which refers to defining the agent with names (in naming conditions), and type refers to defining agents, either robot or human (Figure 4).

i.Identity

Pearson's chi-square test of association revealed a significant relationship between the gender of participants and agents in personality trait attribution (1, N = 885) = 1.012, p =0.314.

ii. Personality Traits

Pearson's chi-square test of association revealed a significant relationship between the gender of participants and agents in personality trait attribution (1, N = 285) = 4.246, p =0.039. Women attributed to personality traits both agents are estimated to be -0.525 times less likely to attribute identity compared to men. A 95% confidence interval indicated high precision (-1.027 to -0.024) in estimating the population odds ratio.

iii. Gender Attribution

Pearson's chi-square test of association revealed a significant relationship between the gender of participants and agents in personality gender attribution (1, N = 207) = 6.326, p =0.012. Women attributed to gender both agents are estimated to be 0.970 times more likely to attribute gender compared to men. A 95% confidence interval indicated high precision (0.196 to 1.744) in estimating the population odds ratio.

iv. Roles and Occupations

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality roles and occupation attribution (1, N = 219) = 0.068 p =0.794.

v.Physical States

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 41) = 2.932, p =0.087.

vi. Type

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 817) = 0.080, p =0.778.

vii. Actions

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality action attribution (1, N = 190) = 0.166, p =0.684.

vii. Other

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in a personality other (non-fit abovementioned categories) (1, N = 56) = 1.972, p =0.160.

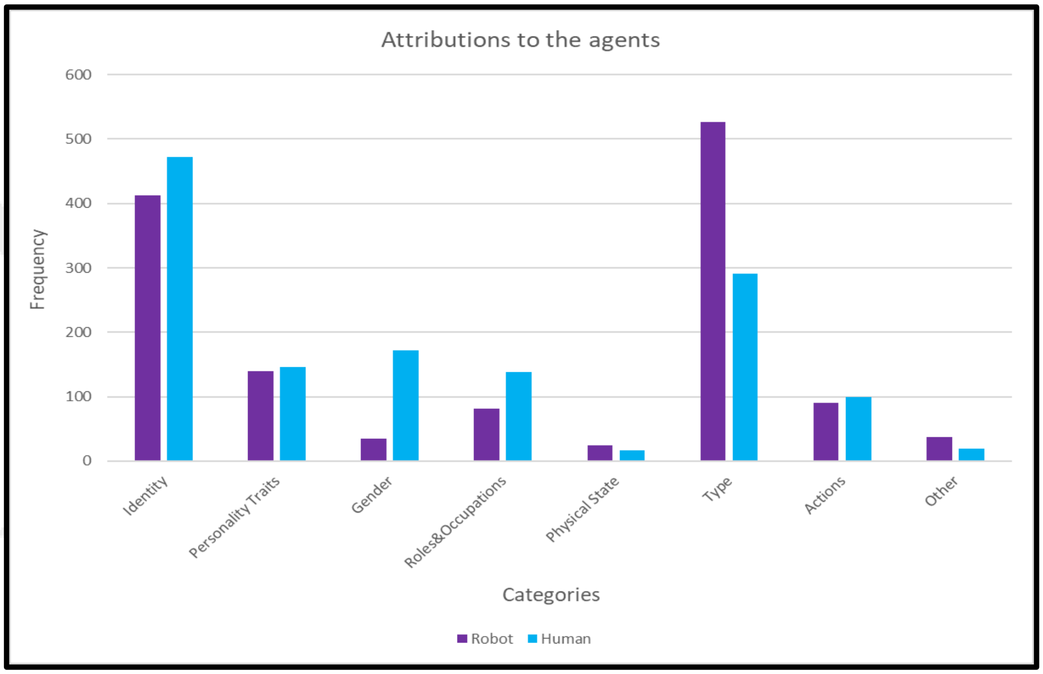

Figure 4.The all-conditions personification attributions' bar graph across 9 categories for the robot and the human agent.

3.2 Individual Naming Group Analysis

3.2.1 Feminine Name Condition

3.2.1.1 BIF rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =13.716 p<.001, η^2 =0.196, meaning that participants rated agent's agency level to actions differently (Figure 5). The post hoc analysis shows that the chop action BIF level is rated significantly different from ringing the bell action (*MD*= -0.483 *SE*=0.084, *t*=-5.741, *p*<0.001) and reading action (*MD*= -0.483 *SE*=0.084, *t*=-5.741, *p*<0.001). The making a list action is rated significantly different from the ringing the bell action (*MD*= -0.383 *SE*=0.084, *t*= -4.553, *p*<0.001) and reading action (*MD*= 0.133 *SE*=0.084, *t*=-4.553, *p*<0.001). The watering the plant action is rated significantly different from ringing the bell action (*MD*= -0.533 *SE*=0.084, *t*=-6.335, *p*<0.001) and reading action (*MD*= -0.533 *SE*=0.084, *t*=-6.335, *p*<0.001). The painting the wall action is rated significantly from ringing the bell action (*MD*= -0.450 *SE*=0.084, *t*=-5.345, *p*<0.001) and reading action (*MD*= -0.450 *SE*=0.084, *t*=-5.345, *p*<0.001). The ringing the bell action is rated significantly different from plant action (*MD*= -0.567 *SE*=0.084, *t*=6.731, *p*<0.001) and laundry action (*MD*= 0.517 *SE*=0.084, *t*=6.137, *p*<0.001). The locking the door action significantly differs from the planting action (*MD*= 0.300 *SE*=0.084, *t*=-3.563, *p*=016). The planting action significantly differs from the reading action (*MD*= -0.567 *SE*=0.084, *t*=-6.731, p<0.001). The reading action significantly differs from the laundry action (*MD*= 0.517 *SE*=0.084, *t*=6.137, p<0.001). There was no significant main effect of agents F (1, 29) =2.295 *p* =0.095. There are no significant interactions between Agent and Action (p>0.05)

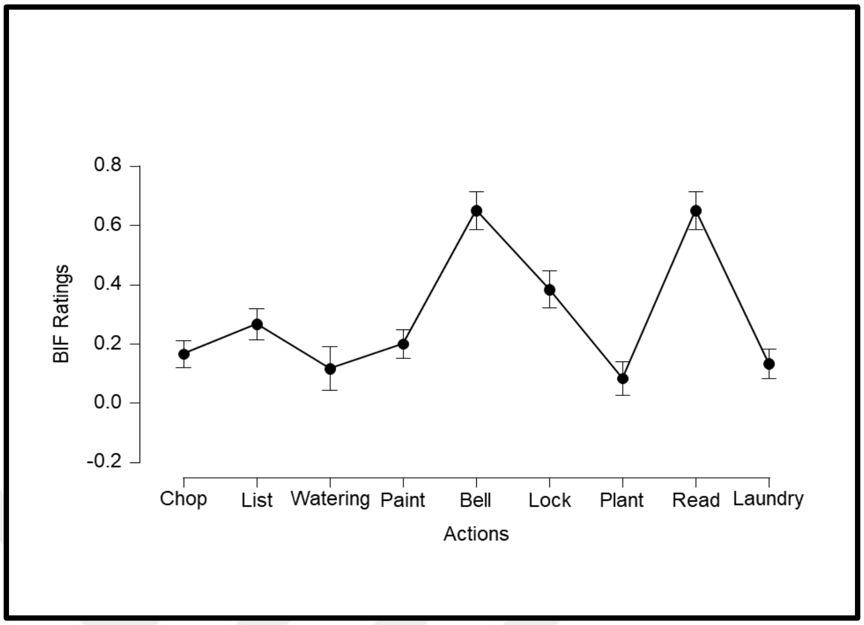

Figure 5. The feminine condition BIF ratings across all 9 actions by participants

3.2.1.2 F A/Competency rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =10.533 p<.001, η^2 =0.211, meaning that participants rated competency of agents across actions differently (Figure 6). The post hoc analysis shows that the chop action competency level is rated significantly different from making a list action (*MD*= -1.333 *SE*=0.250, *t*=-4.533, *p*<0.001), watering action (*MD*= -1.017 *SE*=0.250, *t*=-4.066, *p*=0.002), painting action (*MD*= -1.100 *SE*=0.250, *t*=-4.400, *p*<0.001); ringing the bell action (*MD*= -1.833 *SE*=0.250, *t*=-7.333, *p*<0.001); planting action (*MD*= -1.500 *SE*=0.250, *t*=-5.999, *p*<0.001); reading action (*MD*= -1.267 *SE*=0.250, *t*=-5.066, *p*<0.001) and laundry action (*MD*= -0.950 *SE*=0.250, *t*=-3.800, *p*=0.007). The making a list action is rated significantly different from the locking the door action (*MD*= 0.900 *SE*=0.084, *t*= 3.600, *p*<0.014) and reading action (*MD*= 0.133 *SE*=0.250, *t*=-4.553, *p*<0.001).

The watering the plant action is rated significantly different from the ringing the bell action (*MD*= -0.817 *SE*=0.250, *t*=-3.266, *p*<0.045). The painting of the wall action is rated significantly from locking the door action (*MD*= 0.867 *SE*=0.250, *t*=-3.466, *p*=0.023). The ringing the bell action is rated significantly different from the locking the door action (*MD*= -0.567 *SE*=0.084, *t*=6.731, *p*<0.001) and laundry action (*MD*= 1.600 *SE*=0.250, *t*=6.399, *p*<0.001). The locking the door action significantly differs from the planting action (*MD*= -1.267 *SE*=0.250, *t*=-5.066, *p*=0.001).

There was no significant main effect of agents F (1, 29) =0.011 *p* =0.919. There is no significant interaction between Actions and Agents (p>0.05).

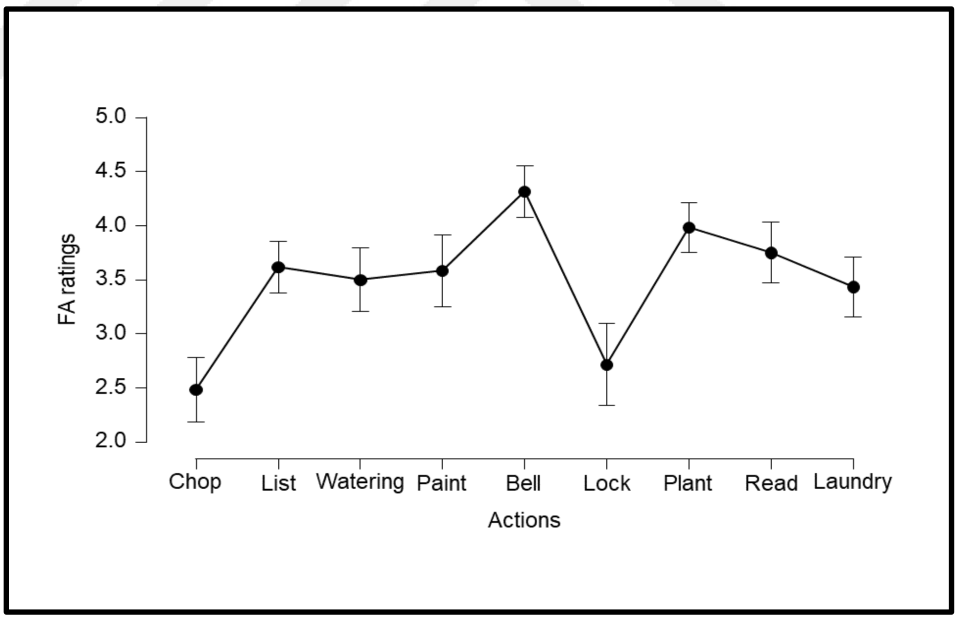

Figure 6. The feminine condition FA ratings across all 9 actions by participants

3.2.1.3 Gender Attribution

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =16.997 $p<.001$, η^2 =0.245, meaning that people attributed different gender levels to agents across actions (Figure 7).

The post hoc analysis shows that the chop action gender attribution level is rated significantly different from making a list action (*MD*= -1.383 *SE*=0.213, *t*=-6.495, *p*<0.001), watering action(*MD*= -1.783 *SE*=0.213, *t*=-8.371, *p*<0.001); ringing the bell action (*MD*= -1.400 *SE*=0.213, *t*=-6.572, *p*<0.001); locking the door action (*MD*= -1.017 *SE*=0.213, *t*=-4.772, *p*<0.001) planting action (*MD*= -1.450 *SE*=0.213, *t*=-6.806, *p*<0.001); reading action (*MD*= -1.417 *SE*=0.213, *t*=-6.650, *p*<0.001) and laundry action (*MD*= -2.033 *SE*=0.213, *t*=-9.545, *p*<0.001). The making a list action is rated significantly different from the locking the painting action (*MD*= 0.817 *SE*=0.213, *t*= 3.834, *p*<0.006). The watering of the plant action is rated significantly different from the painting action (*MD*= 1.217 *SE*=0.213, *t*= 5.711, *p*<0.001). The painting of the wall action is rated significantly from ringing the bell action (*MD*= -0.833 *SE*=0.213, *t*=-3.912, *p*<0.004); planting action (*MD*= -0.883 *SE*=0.213, *t*=-4.146, *p*=0.002); reading action (*MD*= -0.850 *SE*=0.213, *t*=-3.990, *p*=0.003) and laundry action (*MD*= -1.467 *SE*=0.213, *t*=-6.885, *p*<0.001) The locking the door action is significantly different from the laundry action (*MD*= -1.017 *SE*=0.213, *t*=-4.772, *p*=0.001).

There was no significant main effect of agents F (1, 29) =2.615 p =0.117. There is no significant interaction between the Agent and the Action (p>0.05)

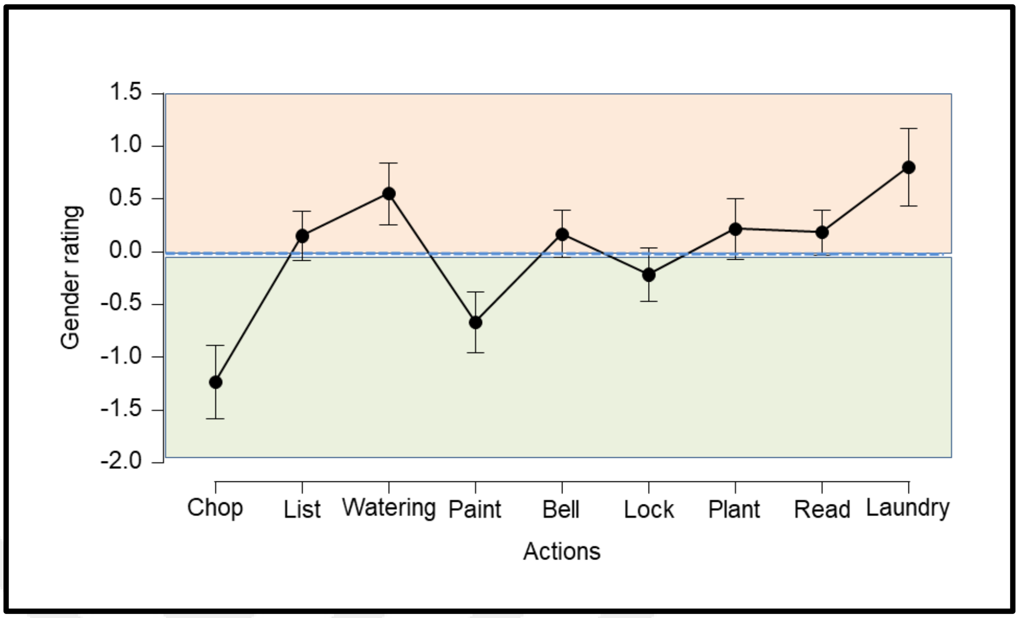

Figure 7. These are the feminine condition gender attribution ratings showing the extent actions are pronounced responses (more femininity or masculinity) across all 9 actions by participants.

Table 5

The BIF, FA, and Gender attribution summary of the feminine condition. The significance level is $p<0.05$; significant results are shown in bold.

Feminine name Condition	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
BIF	$F(1, 29) = 13.716$ **p <.001***** $\eta^2 = 0.196$	$F(1, 29) = 2.295$ $p = 0.095$	$F(8, 232) = 1.738$ $p = 0.093$
FA/Competency	$F(1, 29) = 10.533$ **p <.001***** $\eta^2 = 0.211$	$F(1, 29) = 0.011$ $p = 0.919$	$F(8, 232) = 1.163$ $p = 0.322$
Gender Attribution	$F(1, 29) = 16.997$ **p =.001***** $\eta^2 = 0.245$	$F(1, 29) = 2.615$ $p = 0.117$	$F(8, 232) = 1.262$ $p = 0.264$

3.2.1.4 Personification

The 30 participants' responses to 18 videos (9 human and 9 robot) which makes up 540 responses in total, were analyzed based on 9 themes. The most frequent themes are identity, which refers to defining the agent as Ayşe, and type, which defines agents as either robot or human (Figure 8).

i. Identity

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in identity attribution (1, $N = 245$) = 0.128, $p = 0.721$. It means there is no significant identity attribution to human and robot agents between women and men participants. They attribute identity to the robot and human agents by calling them their names on a similar level.

ii. Personality Traits

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality trait attribution (1, $N = 33$) = 0.509, $p = 0.475$.

iii. Gender Attribution

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality gender attribution (1, $N = 56$) = 2.777, $p = 0.096$.

iv. Roles and Occupations

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality roles and occupation attribution (1, N = 16) = 0.085, p =0.771.

v. Physical States

Due to the small sample size(n=4), the chi-square value was not calculated (p=1.000).

vi. Type

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 133) = 0.505, p =0.477.

vii. Actions

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality action attribution (1, N = 44) = 0.010, p =0.920.

vii. Other

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in a personality other (non-fit abovementioned categories) (1, N = 12) = 0.037, p =0.848.

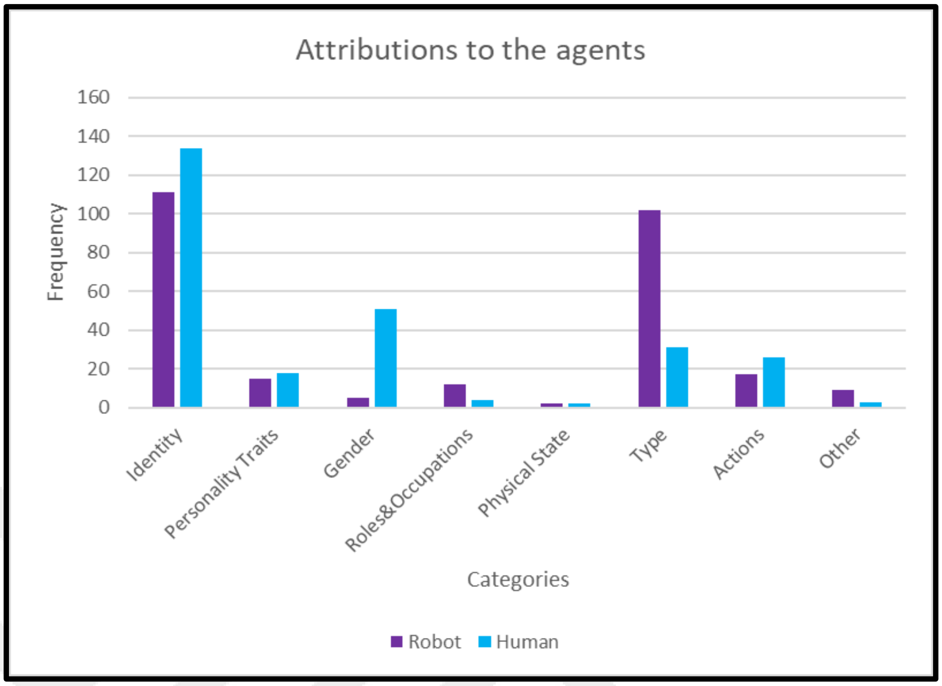

Figure 8. The feminine condition personification attributions bar graph across 9 categories for the robot and the human agent.

3.2.2 Masculine Name Condition

3.2.2.1 BIF rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =22.032 p<.001, η^2 =0.331, meaning that participants rated agent's agency level to actions differently.

The post hoc analysis shows that the chop action BIF level is rated significantly different from ringing the bell action (*MD*= -0.750 *SE*=0.078, *t*=-9.621, *p*<0.001) and reading action (*MD*= -0.533 *SE*=0.078, *t*=-6.842, *p*<0.001). The making a list action is rated significantly different from the ringing the bell action (*MD*= -0.583 *SE*=0.078, *t*= -7.583, *p*<0.001) and reading action (*MD*= -0.367 *SE*=0.078, *t*=-4.704, *p*<0.001). The watering the plant action is rated significantly different from ringing the bell action (*MD*= -0.667 *SE*=0.078, *t*=-8.552, *p*<0.001) and reading action (*MD*= -0.450 *SE*=0.078, *t*=-5.773, *p*<0.001). The painting the wall action is rated significantly from ringing the bell action (*MD*= -0.617 *SE*=0.078, *t*=-7.910, *p*<0.001) and reading action (*MD*= -0.400 *SE*=0.078, *t*=-5.131, *p*<0.001). The ringing the bell action is rated significantly different from locking the door action (*MD*= 0.567 *SE*=0.078, *t*=7.269, *p*<0.001), planting action (*MD*= 0.750 *SE*=0.078, *t*=9.621, *p*<0.001 and laundry action (*MD*= 0.767 *SE*=0.078, *t*=9.835, *p*<0.001). The locking the door action is significantly different from the reading action (*MD*= -0.350 *SE*=0.078, *t*=-4.490, *p*=001). The planting action is significantly different from the reading action (*MD*= -0.533 *SE*=0.078, *t*=-6.842, p<0.001). The reading action is significantly different from the laundry action (*MD*= 0.550 *SE*=0.078, *t*=7.055, p<0.001). There was a significant main effect of agents F (1, 29) =7.165 *p* =0.012, η^2 =0.010. The post hoc analysis shows that the robot agent agency level is rated significantly different from the human agent (*MD*= -0.085 *SE*=0.032, *t*=-2.2677, *p*=0.012).

There was a significant main interaction of ActionsXAgents *F* (8, 232) =4.145 p <0.001, η^2 <0.025(Figure 9). The post hoc analysis shows that the robot-chopping a tree action rated significantly different agency level from robot-ringing the bell action (*MD*= -0.700 *SE*=0.093, *t*=-7.522, *p*<0.001), human-ringing the bell action (*MD*= -0.833 *SE*=0.095, *t*=-8.736, *p*<0.001) and human-reading action (*MD*= -0.767 *SE*=0.095, *t*=-8.037, *p*<0.001). The robot-making a list action is rated significantly different from robot-ringing the bell action (*MD*= -0.500 *SE*=0.093, *t*=-5.373, *p*<0.001), human-ringing the bell action (*MD*= -0.633 *SE*=0.095, *t*=-

6.639, *p*<0.001) and human-reading action (*MD*= -0.567 *SE*=0.095, *t*=-5.940, *p*<0.001). The robot-watering plant action is rated significantly different from robot-ringing the bell action (*MD*= -0.667 *SE*=0.093, *t*=-7.163, *p*<0.001), human-ringing the bell action (*MD*= -0.800 *SE*=0.095, *t*=-8.386, *p*<0.001) and reading action (*MD*= -0.733 *SE*=0.095, *t*=-7.688, *p*<0.001). The robot-painting wall action is rated significantly different from robot-ringing the bell action (*MD*= -0.533 *SE*=0.093, *t*=-5.731, *p*<0.001), human-ringing the bell action (*MD*= -0.667 *SE*=0.095, *t*=-6.989, *p*<0.001) and human-reading action (*MD*= -0.600 *SE*=0.095, *t*=-6.290, *p*<0.001). The robot-ringing-the-bell action is significantly different from robot-locking the door action (*MD*= 0.500 *SE*=0.093, *t*=5.373, *p*<0.001), robot-planting action (*MD*= 0.733 *SE*=0.093, *t*=7.880, *p*<0.001; robot-reading action (*MD*= 0.367 *SE*=0.093, *t*=3.940, =0.015); robot-laundry action (*MD*= 0.700 *SE*=0.093, *t*=7.522, *p*<0.001); human-chopping tree action (*MD*= 0.667 *SE*=0.095, *t*=6.989, *p*<0.001); human-making a list action (*MD*= 0.533 *SE*=0.095, *t*=5.591, *p*<0.001);human-watering plant action (*MD*= 0.533 *SE*=0.095, *t*=5.591, *p*<0.001);human-locking door action (*MD*= 0.500 *SE*=0.095, *t*=6.639, *p*<0.001);human planting action (*MD*= 0.633 *SE*=0.095, *t*=-7.639, *p*<0.001) and human-laundry action (*MD*= 0.700 *SE*=0.095, *t*=7.338, *p*<0.001). The robot-locking door action is significantly rated different from human-ringing the bell action (*MD*= -6.333 *SE*=0.095, *t*=-6.639, *p*<0.001) and human-reading action (*MD*= -0.567 *SE*=0.095, *t*=-5.940, *p*<0.001). The robot-planting action is rated significantly different from the robot-reading action (*MD*= -0.367 *SE*=0.093, *t*=-3.940, p=0.015); human-ringing the bell action (*MD*= -0.867 *SE*=0.095, *t*=-9.085, *p*<0.001) and human-reading action (*MD*= -0.800 *SE*=0.095, *t*=-8.386, *p*<0.001). The robot-reading action is rated significantly different from the human-ringing bell action (*MD*= -0.500 *SE*=0.095, *t*=-5.242, *p*<0.001) and human-reading action (*MD*= -0.433 *SE*=0.075, *t*=-5.787, *p*<0.001). The robot-laundry action is rated significantly different from the human-ringing bell action (*MD*= -0.833 *SE*=0.095, *t*=-8.736, *p*<0.001) and human-reading action (*MD*= -0.767 *SE*=0.095, *t*=-8.037, *p*<0.001). The human-chopping tree action is rated significantly different from the human-ringing bell action (*MD*= -0.800 *SE*=0.093, *t*=-8.596, *p*<0.001) and human-reading action (*MD*= -0.733 *SE*=0.093, *t*=-7.880, *p*<0.001). The human-making a list action is rated significantly different from the human-ringing bell action (*MD*= -0.667 *SE*=0.093, *t*=-7.163, *p*<0.001) and human-reading action (*MD*= -0.600 *SE*=0.093, *t*=-6.447, *p*<0.001). The human-watering plant action is rated significantly different from the human-ringing bell action (*MD*= -0.667 *SE*=0.093, *t*=-7.163, *p*<0.001) and human-reading action (*MD*= -0.600 *SE*=0.093, *t*=-6.447, *p*<0.001). The human-painting action is rated significantly different from the human-ringing bell action (*MD*= -0.700 *SE*=0.093, *t*=-7.7533, *p*<0.001) and human-reading action (*MD*= -0.633 *SE*=0.093, *t*=-6.805, *p*<0.001). The human-ringing the bell action is rated significantly different from the human-locking door (*MD*= 0.633 *SE*=0.093, *t*= 6.805, *p*<0.001), human-planting action (*MD*= 0.767 *SE*=0.093, *t*= 8.238, *p*<0.001). and human-laundry action (*MD*= 0.833 *SE*=0.093, *t*=-8.954, *p*<0.001). The human-locking door action is rated significantly different from the human-reading action (*MD*= -0.567 *SE*=0.093, *t*= -6.089, *p*<0.001). The human-planting action is rated significantly different from the human-reading action (*MD*= -0.700 *SE*=0.093, *t*= -7.522, *p*<0.001). The human-reading action is rated significantly different from the human-laundry action (*MD*= 0.767 *SE*=0.093, *t*= 8.238, *p*<0.001).

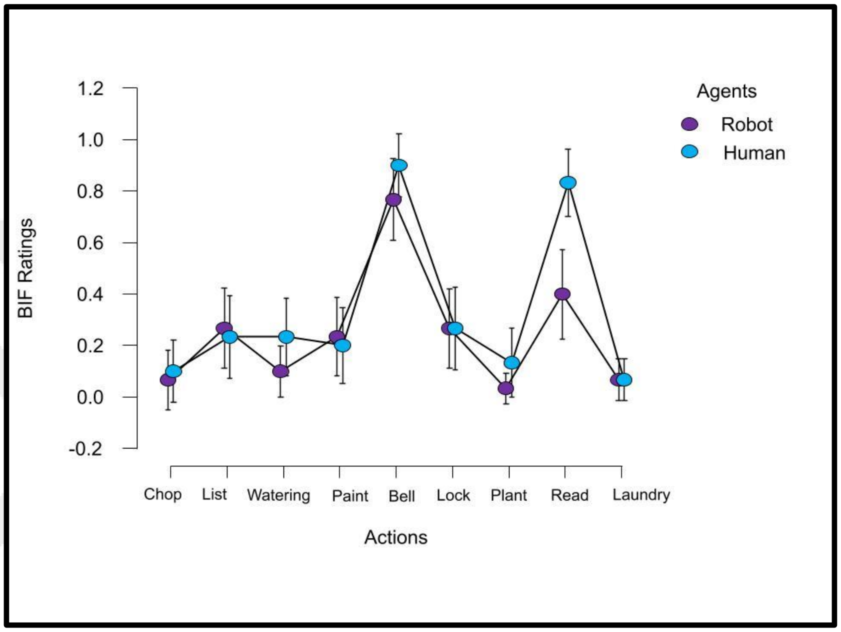

Figure 9. The masculine condition BIF ratings across all 9 actions by agents

3.2.2.2 FA/Competency rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =12.058 p<.001, η^2 =0.235, meaning that participants rated competency of agents across actions differently (Figure 10).

The post hoc analysis shows that the chop action competency level is rated significantly different from making a list action (*MD*= -1.050 *SE*=0.238, *t*=-4.417, *p*<0.001), watering action (*MD*= -1.050 *SE*=0.238, *t*=-4.417, *p*=0.002), painting action (*MD*= -0.900 *SE*=0.238, *t*=-3.786, *p*=0.007); ringing the bell action (*MD*= -1.567 *SE*=0.238, *t*=-6.591, *p*<0.001); planting action (*MD*= -1.500

SE=0.238, *t*=-6.310, *p*<0.001) and reading action (*MD*= -1.217 *SE*=0.238, *t*=-4.698, *p*<0.001). The making a list action is rated significantly different from the locking the door action (*MD*= 1.150 *SE*=0.238, *t*= 4.838, *p*<0.001). The watering the plant action is rated significantly different from the locking the door action (*MD*= 1.150 *SE*=0.238, *t*=4.838, *p*=0.045). The painting of the wall action is rated significantly from locking the door action (*MD*= 1.00 *SE*=0.238, *t*=4.207, *p*=0.023). The ringing the bell action is rated significantly different from the locking the door action (*MD*= 1.667 *SE*=0.238, *t*=7.012, *p*<0.001) and laundry action (*MD*= 1.600 *SE*=0.250, *t*=6.399, *p*<0.001). The locking the door action is significantly different from the planting action (*MD*= -1.600 *SE*=0.238, *t*=-6.731, *p*=0.001), reading action (*MD*= -1.217 *SE*=0.238, *t*=-5.118, *p*=0.001) and laundry action (*MD*= -0.833 *SE*=0.238, *t*=-3.506, *p*=0.020).

There was no significant main effect of agents F (1, 29) =0.025, p =0.877. There are no significant interactions between Action and Agents (p>0.05).

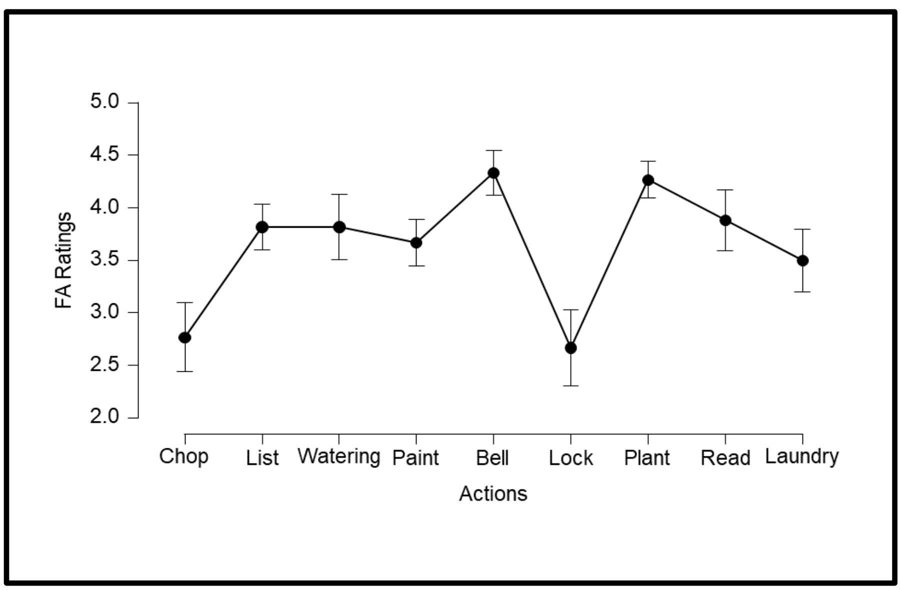

Figure 10. The masculine condition FA ratings across all 9 actions by participants

3.2.2.3 Gender Attribution

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =19.669 p<.001, η^2 =0.289, meaning that participants attributed different gender levels to agents across actions (Figure 11). The post hoc analysis shows that the chop action gender attribution level is rated significantly different from making a list action (MD= -1.367 SE=0.207, t=-6.602, p<0.001), watering action (MD= -2.083 SE=0.207, t=-10.064, p<0.001); painting action (MD= -0.717 SE=0.207, t=-3.462, p=0.023); ringing the bell action (MD= -1.517 SE=0.207, t=-7.326, p<0.001); locking the door action (MD= -1.233 SE=0.207, t=-7.326, p<0.001) planting action (MD= -1.433 SE=0.207, t=-6.924, p<0.001); reading action (MD= -1.433 SE=0.207, t=-6.924, p<0.001) and laundry action (MD= -2.100 SE=0.207, t=-10.144, p<0.001). The making a list action is rated significantly different from the watering plant action (MD= -0.717 SE=0.207, t= -3.462, p=0.023) and laundry action (MD= -0.733 SE=0.207, t=-3.542, p=0.017). The watering of the plant action is rated significantly different from the painting action (MD= 1.367 SE=0.207, t= 6.602, p<0.001) and locking door action (MD= 0.850 SE=0.207, t=4.106, p=0.002). The painting of the wall action is rated significantly from ringing the bell action (MD= -0.800 SE=0.207, t=-3.865, p=0.005); planting action (MD= -0.717 SE=0.207, t=-3.462, p=0.023); reading action (MD= -0.717 SE=0.207, t=-3.462, p=0.023) and laundry action (MD= -1.383 SE=0.207, t=-6.682, p<0.001). The locking of the door action is significantly different from the laundry action (*MD*= -0.867 *SE*=0.207, *t*=-4.187, *p*=0.001). There was no significant main effect of agents F (1, 29) =0.020 *p* =0.888.

There was a significant main effect of Actions X Agents F (1, 29) =2.384 *p* =0.017, η^2 <0.014. The post hoc analysis shows that the robot-chopping a tree action rated significantly different agency level from the robot-making a list action (*MD*= -1.267 *SE*=0.246, *t*=-5.147, *p*<0.001), robot-watering action (*MD*= -1.967 *SE*=0.246, *t*=-7.991, *p*<0.001) human-ringing the bell action (*MD*= -1.633 *SE*=0.246, *t*=-6.637, *p*<0.001) robot-locking the door action (*MD*= -1.000 *SE*=0.246, *t*=-4.063, *p*=0.009); robot-planting action (*MD*=-1.533 *SE*=0.246, *t*=-6.230, *p*<0.001); robot-laundry action (*MD*=-2.300 *SE*=0.246, *t*=-9.345, *p*<0.001); human-listing action (*MD*=-1.533 *SE*=0.271, *t*=-5.656, *p*<0.001); human-ringing the bell action (*MD*= -1.567 *SE*=0.271, *t*=-5.410, *p*<0.001); human-locking the door action (*MD*= -1.533 *SE*=0.271, *t*=-5.656, *p*<0.001);human-planting action (*MD*= -1.400 *SE*=0.271, *t*=-5.164, *p*<0.001); human-reading action (*MD*= -1.400 *SE*=0.271, *t*=-7.254, *p*<0.001) and human-laundry action (*MD*= -1.967 *SE*=0.271, *t*=-7.254, *p*<0.001). The robot-making a list action is rated significantly different from the robot-laundry action (*MD*= -1.033 *SE*=0.246, *t*=-4.199, *p*=0.005), human-chopping tree action (*MD*= 1.200 *SE*=0.271, *t*=4.426, *p*=0.002) and human-watering action (*MD*= -1.000 *SE*=0.271, *t*=-3.680, *p*=0.041). The robot-watering plant action is rated significantly different from the robot-painting action (*MD*= *1*.100

$SE=0.246$, $t=4.470$, $p=0.002$), human-chopping tree action ($MD=1.900$ $SE=0.271$, $t=7.009$, $p<0.001$) and human-painting action ($MD=1.333$ $SE=0.271$, $t=4.918$, $p<0.001$). The robot-painting wall action is rated significantly different from the robot-laundry action ($MD=-1.433$ $SE=0.246$, $t=-5.824$, $p<0.001$), human-watering action ($MD=1.400$ $SE=0.271$, $t=-5.164$, $p<0.001$) and human-laundry action ($MD=-1.100$ $SE=0.271$, $t=-4.058$, $p<0.001$). The robot-ringing-the-bell action is significantly different from the human-chopping tree action ($MD=1.567$ $SE=0.271$, $t=5.779$, $p<0.001$), human-painting action ($MD=1.000$ $SE=0.271$, $t=3.689$, $p=0.041$; robot-reading action ($MD=0.367$ $SE=0.093$, $t=3.940$, $p=0.015$); The robot-locking door action is significantly rated different from robot-laundry action ($MD=-1.300$ $SE=0.246$, $t=-5.282$, $p<0.001$) and human-watering action ($MD=-1.267$ $SE=0.271$, $t=-1.967$, $p<0.001$). The robot-planting action is rated significantly different from the human-chopping tree action ($MD=1.467$ $SE=0.271$, $t=5.410$, $p<0.001$). The robot-reading action is rated significantly different from the human-chopping tree action ($MD=1.467$ $SE=0.271$, $t=5.410$, $p<0.001$). The robot-laundry action is rated significantly different from the human-chopping tree action ($MD=2.233$ $SE=0.271$, $t=8.238$, $p<0.001$) and human-painting action ($MD=1.667$ $SE=0.271$, $t=6.148$, $p<0.001$). The human-chopping tree action is rated significantly different from the human-making a list action ($MD=-1.467$ $SE=0.246$, $t=-5.959$, $p<0.001$) and human-watering action ($MD=-2.200$ $SE=0.246$, $t=-8.939$, $p<0.001$);human-ringing the bell action (*MD*= -*1*.400 *SE*=0.246, *t*=-5.688, *p*<0.001); locking door action (*MD*= -1.467 *SE*=0.246, *t*=-5.959, *p*<0.001); human-planting action (*MD*= -1.133 *SE*=0.246, *t*=-5.418, *p*<0.001); human-reading action (*MD*= -1.133 *SE*=0.246, *t*=-5.418, *p*<0.001) and human-laundry action (*MD*= -1.900 *SE*=0.246, *t*=-7.720, *p*<0.001) The human-making a list action is rated significantly different from the human-painting action (*MD*= 0.900 *SE*=0.246, *t*=3.657, *p*=0.044). The human-watering plant action is rated significantly different from the human-painting action (*MD*= 1.633 *SE*=0.246, *t*=6.637, *p*<0.001). The human-painting action is rated significantly different from the human-locking the door action (*MD*= -0.900 *SE*=0.246, *t*=-3.657, *p*=0.044) and human-laundry action (*MD*= -1.333 *SE*=0.246, *t*=-5.418, *p*<0.001).

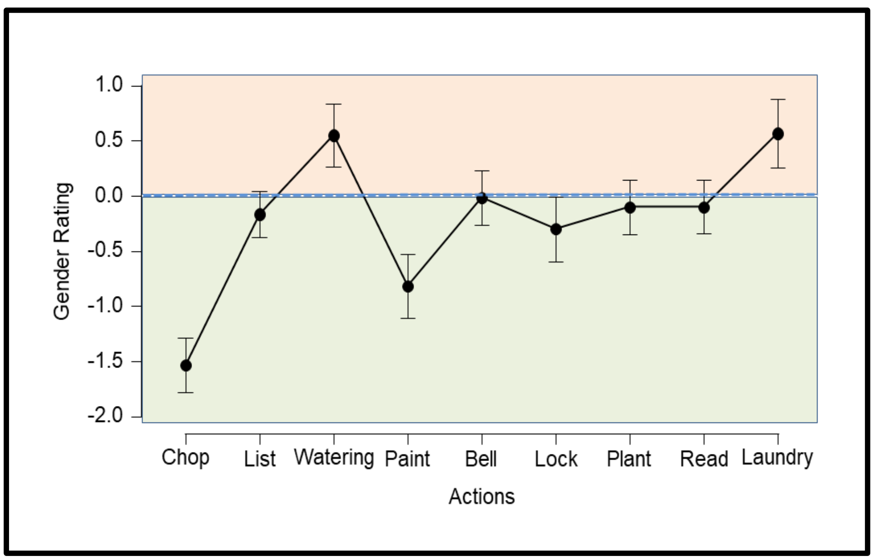

Figure 11. These are the masculine condition gender attribution ratings showing the extent actions are pronounced responses (more femininity or masculinity) across all 9 actions by participants.

Table 6

The BIF, FA, and Gender attribution summary of the masculine condition. The significance level is p<0.05; significant results are shown in bold.

Masculine name Condition	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
BIF	$F(1, 29) = 22.032$ **$p < .001$***** $\eta^2 = 0.331$	$F(1, 29) = 7.165$ **$p = 0.012$** $\eta^2 = 0.010$	$F(8, 232) = 4.145$ **$p < .001$***** $\eta^2 < 0.025$
FA/Competency	$F(1, 29) = 12.058$ **$p = .001$***** $\eta^2 = 0.235$	$F(1, 29) = 0.025$ $p = 0.877$	$F(8, 232) = 0.283$ $p = 0.971$
Gender Attribution	$F(1, 29) = 19.669$ **$p = .001$***** $\eta^2 = 0.289$	$F(1, 29) = 0.020$ $p = 0.888$	$F(8, 232) = 2.384$ **$p = 0.017$** $\eta^2 < 0.014$

3.2.2.4 Personification

The 30 participants' responses to 18 videos (9 human and 9 robot) which makes up 540 responses in total, were analyzed based on 9 themes. The most frequent themes are identity, which refers to defining the agent as Ahmet, and type, which defines agents as either robot or human (Figure 12).

i. Identity

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in identity attribution (1, N = 231) = 0.069, p =0.792. It means there is no significant identity attribution to human and robot agents between women and men participants. They equally attribute identity to the robot and human agents by calling them their names.

ii.Personality Traits

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality trait attribution (1, N = 79) = 0.5277, p =0.599.

iii.Gender Attribution

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality gender attribution (1, N = 11) = 0.244, p =0.621.

iv.Roles and Occupations

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality roles and occupation attribution (1, N = 39) = 0.542, p =0.462.

v.Physical States

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 17) = 1.893, p =0.169.

vi.Type

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 99) = 3.423, p =0.064.

vii.Actions

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality action attribution (1, N = 39) = 1.739, p =0.187.

viii. Other

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in a personality

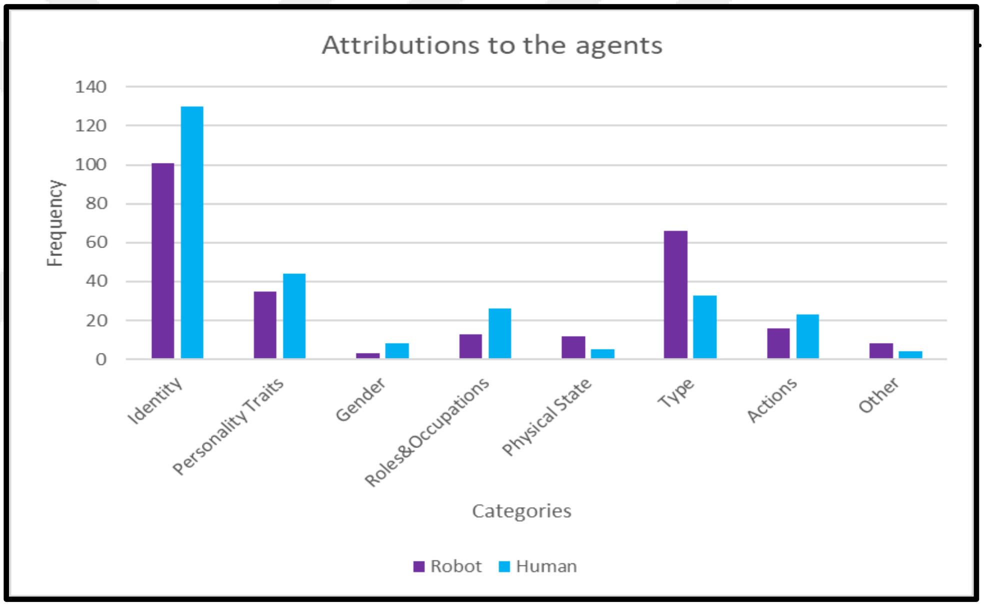

Figure 12. The masculine condition personification attributions' bar graph across 9 categories for the robot and the human agent

3.2.3 Neutral Name Condition

3.2.3.1 BIF rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =22.099 p<.001, η^2 =0.327, meaning that participants rated agent's agency level to actions differently (Figure 13).

The post hoc analysis shows that the chop action BIF level is rated significantly different from ringing the bell action ($MD= -0.600$ $SE=0.082$, $t=-7.343$, $p<0.001$) and reading action ($MD= -0.467$ $SE=0.082$, $t= -5.711$, $p<0.001$). The making a list action is rated significantly different from the watering action ($MD= 0.267$ $SE=0.082$, $t= 3.263$, $p=0.046$), ringing the bell action ($MD= -0.517$ $SE=0.082$, $t= -6.323$, $p<0.001$) reading action ($MD= -0.383$ $SE=0.082$, $t=-4.691$, $p<0.001$) and laundry action ($MD= 0.300$ $SE=0.082$, $t= 3.671$, $p<0.011$). Watering the plant action is rated significantly different from ringing the bell action ($MD= -0.783$ $SE=0.082$, $t=-9.586$, $p<0.001$), locking the door action ($MD= -0.267$ $SE=0.082$, $t= -3.263$, $p=0.046$) and reading action ($MD= -0.650$ $SE=0.082$, $t=-7.954$, $p<0.001$). Painting the wall action is rated significantly from ringing the bell action ($MD= -0.533$ $SE=0.082$, $t=-6.527$, $p<0.001$), reading action ($MD= -0.400$ $SE=0.082$, $t=-4.895$, $p<0.001$) and laundry action ($MD= 0.283$ $SE=0.082$, $t= 3.467$, $p=0.023$). Ringing the bell action is rated significantly different from the locking action ($MD= 0.517$ $SE=0.082$, $t= 6.323$, $p<0.001$), planting action ($MD= 0.617$ $SE=0.082$, $t=7.547$, $p<0.001$) and laundry action ($MD= 0.817$ $SE=0.082$, $t=9.994$, $p<0.001$). The locking of the door action is significantly different from the reading action ($MD= -0.383$ $SE=0.082$, $t=-4.691$, $p=001$) and laundry action ($MD= 0.300$ $SE=0.082$, $t= 3.671$, $p=0.011$). The planting action is significantly different from the reading action ($MD= -0.483$ $SE=0.082$, $t=-5.915$, $p<0.001$). The reading action is significantly different from the laundry action ($MD= 0.683$ $SE=0.082$, $t=8.362$, $p<0.001$). There was no significant main effect of agents F (1, 29) =3.145 p =0.087. There are no significant interactions between Actions and Agents (p>0.05).

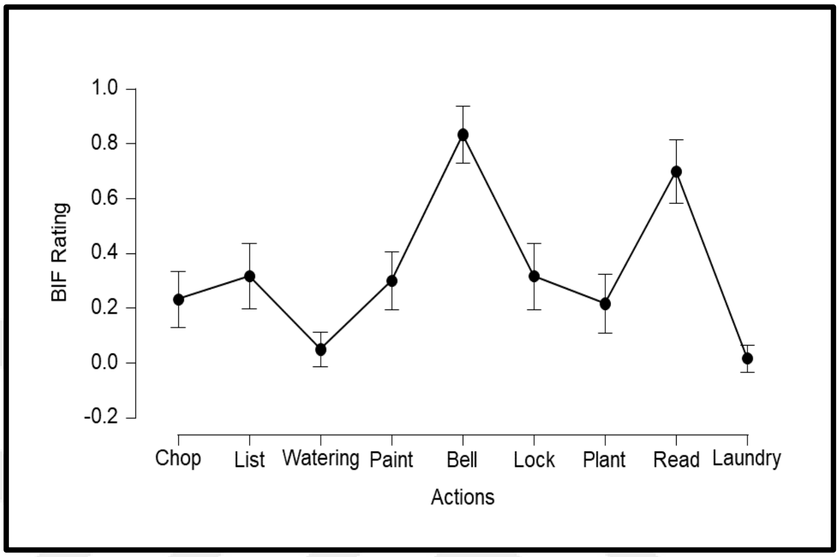

Figure 13. The neutral-name condition BIF ratings across all 9 actions by participants

3.2.3.2 FA/Competency rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =17.044 $p<.001$, η^2 =0.297, meaning that participants rated competency of agents across actions differently (Figure 14). The post hoc analysis shows that the chop action competency level is rated significantly different from making a list action (*MD*= -1.483 *SE*=0.238, *t*=-6.244, *p*<0.001), watering action (*MD*= -1.033 *SE*=0.238, *t*=-4.350, *p*<0.001), painting action (*MD*= -1.133 *SE*=0.238, *t*=-4.771, *p*<0.001); ringing the bell action (*MD*= -2.000 *SE*=0.238, *t*=-8.419, *p*<0.001); planting action (*MD*= -1.617 *SE*=0.238, *t*=-6.805, *p*<0.001), reading action (*MD*= -1.233 *SE*=0.238, *t*=-5.192, *p*<0.001) and laundry action (*MD*= -1.150 *SE*=0.238, *t*=-4.841, *p*<0.001). The making a list action is rated significantly different from the locking the door action (*MD*=

1.567, *SE*=0.238, *t*= 6.595, *p*<0.001). The watering the plant action is rated significantly different from the ringing the bell action (*MD*= -0.967 *SE*=0.238, *t*=-4.069, *p*=0.002) and locking the door action (*MD*= 1.117 *SE*=0.238, *t*=4.700, *p*<0.001). The painting of the wall action is rated significantly from ringing the bell action (*MD*= -0.867 *SE*=0.238, *t*=-3.648, *p*=0.012) and locking the door action (*MD*= 1.217 *SE*=0.238, *t*=5.121, *p*<0.001). The ringing the bell action is rated significantly different from the locking the door action (*MD*= 2.083 *SE*=0.238, *t*=8.769, *p*<0.001) and laundry action (*MD*= 0.850 *SE*=0.238, *t*=3.578, *p*=0.015). The locking the door action is significantly different from the planting action (*MD*= -1.700 *SE*=0.238, *t*=-7.156, *p*=0.001), reading action (*MD*= -1.317 *SE*=0.238, *t*=-5.542, *p*=0.001) and laundry action (*MD*= -1.233 *SE*=0.238, *t*=-5.192, *p*=0.001). There is no main effect of agents *F* (1, 29) =0.023 *p* =0.880. There are no significant interactions between Actions and Agents (p>0.05).

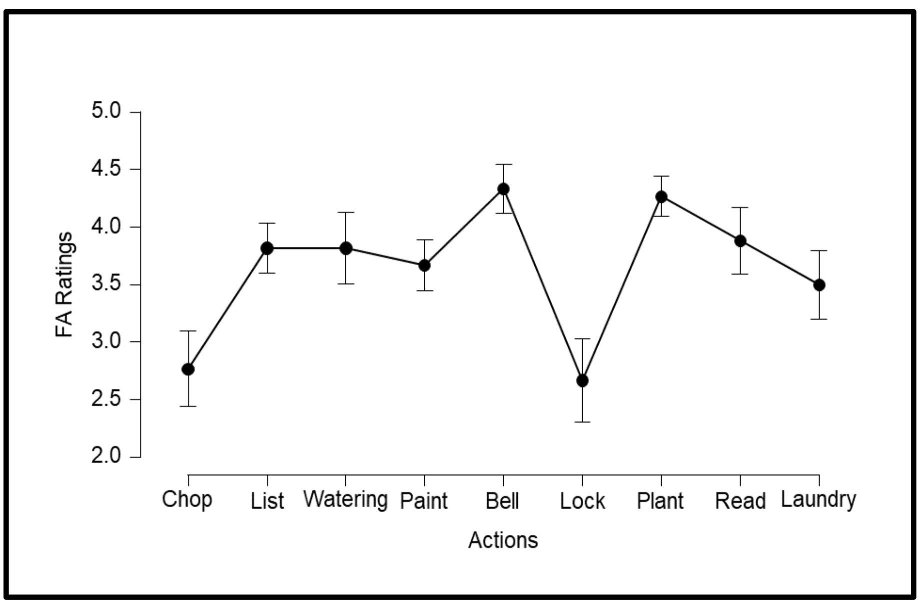

Figure 14. The neutral-name condition FA ratings across all 9 actions by participants

3.2.3.3 Gender Attribution

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =13.063 $p<.001$, η^2 =0.209, meaning that participants attributed different gender levels to the agents across actions (Figure 15). The post hoc analysis shows that the chop action gender attribution level is rated significantly different from making a list action (*MD*= -0.690 *SE*=0.207, *t*=-3.285, *p*=0.043), watering action (*MD*= -1.259 *SE*=0.210, *t*=-5.994, *p*<0.001); ringing the bell action (*MD*= -0.948 *SE*=0.210, *t*=-4.516, *p*<0.001); planting action (*MD*= -0.897 *SE*=0.210, *t*=-4.270, *p*<0.001); reading action (*MD*= -0.897 *SE*=0.210, *t*=-6.405, *p*<0.001) and laundry action (*MD*= -2.100 *SE*=0.207, *t*=-10.144, *p*<0.001). The making a list action is rated significantly different from the laundry action (*MD*= -0.897 *SE*=0.210, *t*=-4.270, *p*<0.001). The watering of the plant action is rated significantly different from the painting action (*MD*= 1.121 *SE*=0.210, *t*= 5.337, *p*<0.001) and locking door action (*MD*= 0.707 *SE*=0.210, *t*=3.367, *p*=0.032). The painting of the wall action is rated significantly from ringing the bell action (*MD*= -0.810 *SE*=0.210, *t*=-3.859, *p*=0.005); planting action (*MD*= -0.759 *SE*=0.210, *t*=-3.613, *p*=0.013); reading action (*MD*= -1.207 *SE*=0.210, *t*=-5.748, *p*<0.001) and laundry action (*MD*= -1.448 *SE*=0.210, *t*=-6.898, *p*<0.001). The locking the door action is significantly different from the reading action (*MD*= -0.793 *SE*=0.210, *t*=-3.77, *p*=0.007) and laundry action (*MD*= -1.034 *SE*=0.210, *t*=-4.927, *p*=0.001). The planting action is significantly different from the laundry action (*MD*= -0.690 *SE*=0.210, *t*=-3.85, *p*=0.043).

There was no significant main effect of agents F (1, 29) =0.098 *p* =0.756. There was a significant interaction of Actions and Agents *F* (8,232) =2.468 *p* =0.014, η^2 <0.016. The post hoc analysis shows that the robot-chopping a tree action rated significantly different agency level from the robot-making a list action (*MD*= -1.267 *SE*=0.249, *t*=-3.738, *p*=0.033), robot-watering action (*MD*= -1.138 *SE*=0.249, *t*=-4.569, *p*<0.001) robot-planting action (*MD*= -1.172 *SE*=0.249, *t*=-4.807, *p*<0.001) robot-reading the door action (*MD*= -1.517 *SE*=0.249, *t*=-6.092, *p*<0.001); robot-laundry action (*MD*= -1.517 *SE*=0.249, *t*=-6.092, *p*<0.001); human-watering action (*MD*= -1.483 *SE*=0.289, *t*=-5.137, *p*<0.001); human-ringing the bell action (*MD*= -1.103 *SE*=0.289, *t*=-3.823, *p*=0.026); human-reading action (*MD*= -1.276 *SE*=0.289, *t*=-4.430, *p*=0.002);human-laundry action (*MD*= -1.759 *SE*=0.289, *t*=-6.093, *p*<0.001). The robot-watering plant action is rated significantly different from the robot-painting action (*MD= 0*.966 *SE*=0.249, *t*=3.876, *p*=0.019). The robot-painting wall action is rated significantly different from the robot-planting action (*MD*= -1.00 *SE*=0.249, *t*=-4.015, *p*=0.011); robot-reading action (*MD*= -1.345 *SE*=0.249, *t*=-5.399, *p*<0.001); robot-laundry action (*MD*= -1.345 *SE*=0.249, *t*=-5.399, *p*<0.001); human-watering action (*MD*= -1.310 *SE*=0.289, *t*=-4.540, *p*<0.001); human-reading action (*MD*= -1.103 *SE*=0.289, *t*=-3.823, *p*=0.026) and human-laundry action (*MD*= -1.586 *SE*=0.289, *t*=-5.496, *p*<0.001). The robot-planting action is rated significantly different from the human-chopping tree action (*MD= 1*.069 *SE*=0.289, *t*=3.704, *p*=0.041). The robot-reading action is rated significantly different from the human-chopping tree action (*MD= 1*.414

$SE=0.289$, $t=4.898$, $p<0.001$), human-painting action ($MD= 1.310$ $SE=0.289$, $t=4.540$, $p<0.001$) and human-locking the door action ($MD= 1.060$ $SE=0.289$, $t=3.704$, $p=0.041$). The robot-laundry action is rated significantly different from the human-chopping tree action ($MMD= 1.414$ $SE=0.289$, $t=4.898$, $p<0.001$) and human-painting action ($MD= 1.310$ $SE=0.289$, $t=4.540$, $p<0.001$) and human-locking the door action ($MD= 1.060$ $SE=0.289$, $t=3.704$, $p=0.041$). The human-chopping tree action is rated significantly different from the watering action ($MD= -1.379$ $SE=0.249$, $t=-5.538$, $p<0.001$) and human-ringing the bell action ($MD= -1.000$ $SE=0.249$, $t=-4.015$, $p=0.011$); human-reading action ($MD= -1.172$ $SE=0.249$, $t=-4.707$, $p<0.001$) and human-laundry action ($MD= -1.655$ $SE=0.249$, $t=-6.645$, $p<0.001$). The human-making a list action is rated significantly different from the human-watering action ($MD= -0.931$ $SE=0.249$, $t=-3.738$, $p=0.033$) and laundry action ($MD= -1.207$ $SE=0.249$, $t=-4.846$, $p<0.001$). The human-watering plant action is rated significantly different from the human-painting action ($MD= 1.276$ $SE=0.249$, $t=5.122$, $p<0.001$) and human-locking door action ($MD= 1.034$ $SE=0.249$, $t=4.153$, $p=0.006$). The human-painting action is rated significantly different from the human-reading action ($MD= 1.069$ $SE=0.249$, $t=-4.292$, $p=0.003$) and human-laundry action ($MD= -1.552$ $SE=0.249$, $t=-6.230$, $p<0.001$). The human-locking door action is rated significantly different from the human-laundry action ($MD= -1.310$ $SE=0.249$, $t=-5.261$, $p<0.001$). The human-planting action is rated significantly different from the human-laundry action ($MD= -1.034$ $SE=0.249$, $t=-4.153$, $p=0.006$).

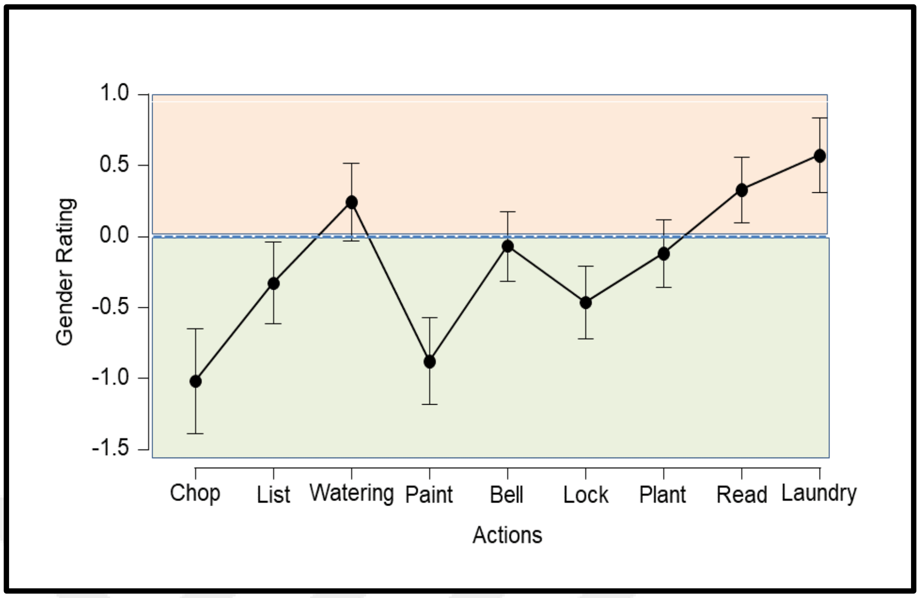

Figure 15. The neutral-name condition gender attribution ratings to what extent actions are pronounced responses (more femininity or masculinity) across all 9 actions by participants.

Table 7

The BIF, FA, and Gender attribution summary of the neutral condition. The significance level is $p<0.05$, and significant results are shown in bold.

Neutral name Condition	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
BIF	$F(1, 29) = 22.099$ **p <.001***** $\eta^2 = 0.327$	$F(1, 29) = 3.145$ $p = 0.087$	$F(8, 232) = 0.728$ $p = 0.667$
FA/Competency	$F(1, 29) = 17.044$ **p =.001***** $\eta^2 = 0.297$	$F(1, 29) = 0.023$ $p = 0.880$	$F(8, 232) = 1.781$ $p = 0.082$
Gender Attribution	$F(1, 29) = 13.063$ **p =.001***** $\eta^2 = 0.209$	$F(1, 29) = 0.098$ $p = 0.756$	$F(8, 232) = 2.468$ **p =0.014** $\eta^2 < 0.016$

3.2.3.4 Personification

The 30 participants' responses to 18 videos (9 human and 9 robot) which makes up 540 responses in total, were analyzed based on 9 themes. The most frequent theme is the identity which refers to defining the agent as Deniz (Figure 16).

i. Identity

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in identity attribution (1, N = 306) = 0.559, p =0.455. It means there is no significant identity attribution to human and robot agents between women and men participants. They equally attribute identity to the robot and human agents by calling them their names.

ii. Personality Traits

Pearson's chi-square test of association revealed a significant relationship between the gender of participants and agents in personality trait attribution (1, N = 71) = 6.724, p =0.010. It means there is a significant identity attribution to human and robot agents between women and men participants. Women attributed to personality traits both agents are estimated to be -1.399 times less likely to attribute identity compared to men. A 95% confidence interval indicated high precision (-2.489 to -0.309) in estimating the population odds ratio.

iii. Gender Attribution

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality gender attribution (1, N = 13) = 0.503, p =0.478.

iv. Roles and Occupations

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality roles and occupation attribution (1, N = 23) = 0.289, p =0.591.

v.Physical States

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 8) = 1.324, p =0.250.

vi. Type

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 95) = 1.690, p =0.194.

vii. Actions

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality action attribution (1, N = 35) = 0.430, p =0.512.

viii. Other

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in a personality other (non-fit abovementioned categories) (1, N = 2) = 1.052, p =0.305.

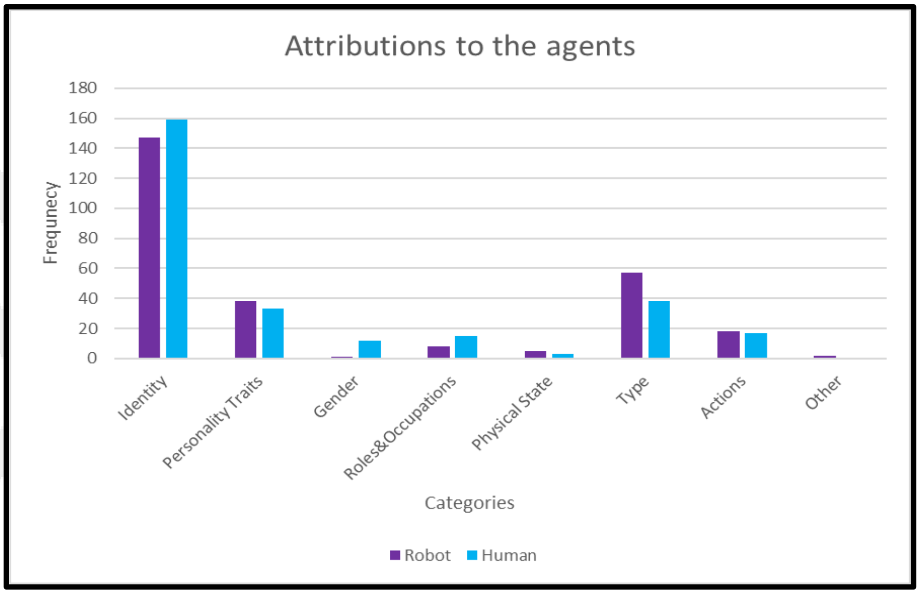

Figure 16. The neutral-name condition personification attributions' bar graph across 9 categories for the robot and the human agent.

3.2.4. Technical Name Condition

3.2.4 .1 BIF rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =20.465 $p<0.001$, η^2 =0.294, meaning that participants rated agent's agency level to actions differently (Figure 17).

The post hoc analysis shows that the chop action BIF level is rated significantly different from ringing the bell action (MD= -0.650 SE=0.074, t=-8.784, p<0.001) and reading action (MD= -0.383 SE=0.074, t=-5.181, *p*<0.001). The making a list action is rated significantly different from the ringing the bell action (*MD*= -0.433 *SE*=0.074, *t*= -5.856, *p*<0.001) and laundry action (*MD*= 0.283 *SE*=0.074, *t*= 3.829, *p*=0.006). The watering the plant action is rated significantly different from ringing the bell action (*MD*= -0.650 *SE*=0.074, *t*=-8.784, *p*<0.001) and reading action (*MD*= -0.383 *SE*=0.074, *t*=-5.181, *p*<0.001). The painting the wall action is rated significantly from ringing the bell action (*MD*= -0.650 *SE*=0.074, *t*=-8.784, *p*<0.001) and reading action (*MD*= -0.383 *SE*=0.074, *t*=-5.181, *p*<0.001). The ringing the bell action is rated significantly different from the locking the door action (*MD*= 0.450 *SE*=0.074, *t*=6.082, *p*<0.001), planting action (*MD*= 0.650 *SE*=0.074, *t*=8.784, *p*<0.001), reading action (*MD*= 0.267 *SE*=0.074, *t*=3.604, *p*=0.014) and laundry action (*MD*= 0.717 *SE*=0.074, *t*=9.685, *p*<0.001). The locking the door action is significantly different from the laundry action (*MD*= -0.267 *SE*=0.074, *t*=3.604, *p*=0.014). The planting action is significantly different from the reading action (*MD*= -0.383 *SE*=0.074, *t*=-5.181, p<0.001). The reading action is significantly different from the laundry action (*MD*= 0.450 *SE*=0.074, *t*=6.082, p<0.001). There was a significant main effect of agents F (1, 29) =6.154 *p* =0.019, η^2 =0.013 (Figure 18).

The post hoc analysis shows that the human agent agency level is rated significantly different from the robot agent (*MD*= 0.093 *SE*=0.037, *t*=-2.481, *p*=0.019). There are no significant interactions between Actions and Agents (p>0.05).

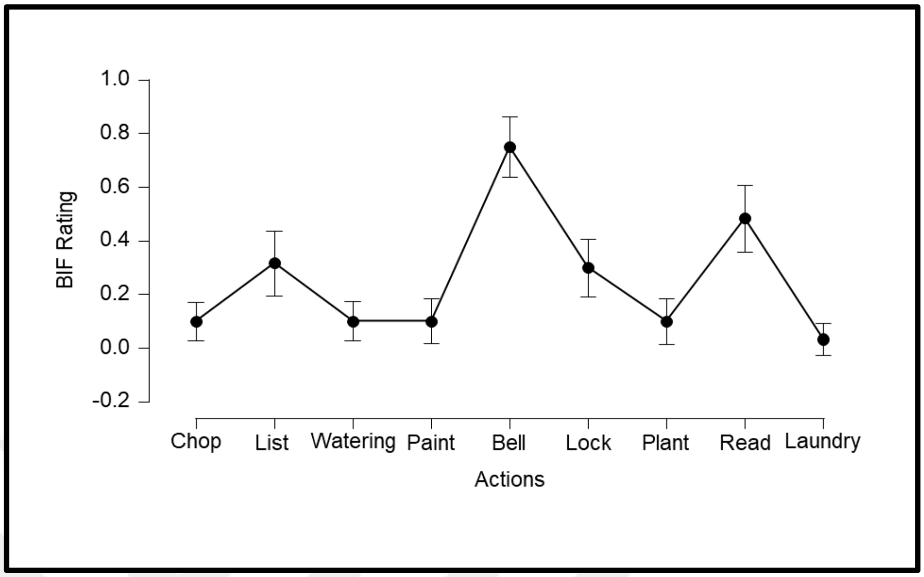

Figure 17. The technical-name condition BIF ratings across all 9 actions by participants

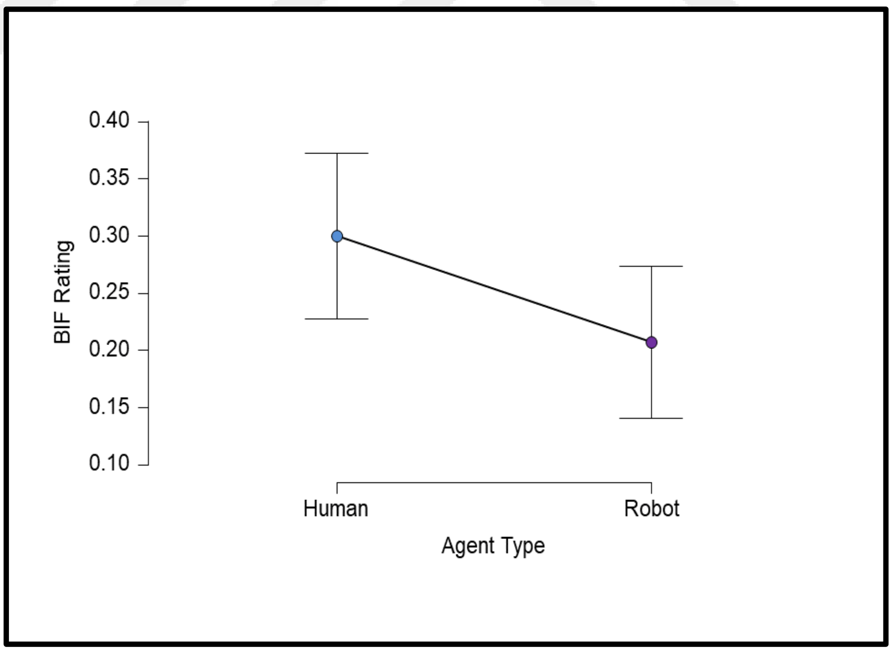

Figure 18. The technical-name condition BIF rating difference between the human and robot agents.

3.2.4 2 FA/Competency rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =16.192 $p<0.001$, η^2 =0.248, meaning that participants rated competency of agents across actions differently (Figure 19).

The post hoc analysis shows that the chop action competency level is rated significantly different from making a list action (*MD*= -1.033 *SE*=0.216, *t*=-4.774, *p*<0.001), watering action (*MD*= -0.867 *SE*=0.216, *t*=-4.004, *p*=0.003), painting action (*MD*= -0.900 *SE*=0.216, *t*=-4.158, *p*=0.002); ringing the bell action (*MD*= -1.483 *SE*=0.216, *t*=-6.852, *p*<0.001); planting action (*MD*= -1.567 *SE*=0.216, *t*=-7.237, *p*<0.001), reading action (*MD*= -1.167 *SE*=0.216, *t*=-5.390, *p*<0.001) and laundry action (*MD*= -0.717, *SE*=0.216, *t*=-3.311, *p*=0.039). The making a list action is rated significantly different from the locking the door action (*MD*= 1.300, *SE*=0.216, *t*= 6.006, *p*<0.001). The watering the plant action is rated significantly different from the locking the door action (*MD*= 1.133 *SE*=0.216, *t*=-5.236, *p*<0.001). The painting of the wall action is rated significantly from locking the door action (*MD*= 1.167 *SE*=0.216, *t*=5.390, *p*<0.001). The ringing the bell action is rated significantly different from the locking the door action (*MD*= 1.750 *SE*=0.216, *t*=8.084, *p*<0.001) and laundry action (*MD*= 0.767, *SE*=0.216, *t*=3.542, *p*=0.017). The locking of the door action is significantly different from the planting action (*MD*= -1.833 *SE*=0.216, *t*=-8.469, *p*=0.001), reading action (*MD*= -1.433 *SE*=0.216, *t*=-6.622, *p*=0.001) and laundry action (*MD*= -0.983 *SE*=0.216, *t*=-4.543, *p*=0.001). The planting action is significantly different from the laundry action (*MD*= 0.850 *SE*=0.216, *t*= 3.927, *p*=0.004). There was no significant main effect of agents F (1, 29) =1.074 *p* =0.309. There are no significant interactions between Actions and Agents (p>0.05).

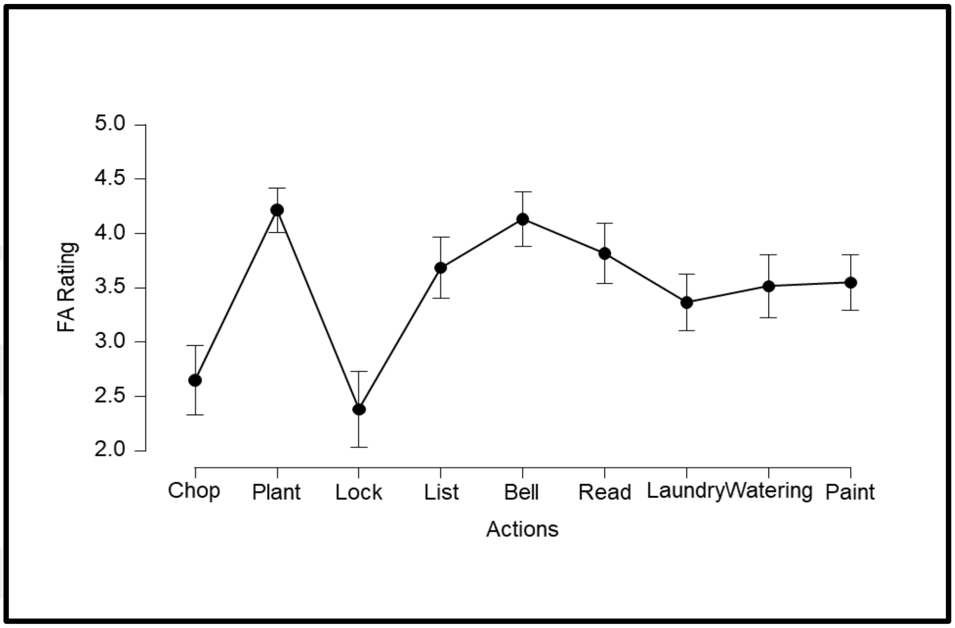

Figure 19. The technical-name condition FA ratings across all 9 actions by participants

3.2.4.3 Gender Attribution

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions $F(1, 29)=14.003$ $p<.001$, $\eta^2=0.208$, meaning that participants attributed different gender levels to the agents across actions(Figure 20).The post hoc analysis shows that the chop action gender attribution level is rated significantly different from making a list action (*MD*= -1.150 *SE*=0.211, *t*=-5.445, *p*<0.001), watering action (*MD*= -1.733 *SE*=0.211, *t*=-8.207, *p*<0.001); ringing the bell action (*MD*= -1.117 *SE*=0.211, *t*=-5.287, *p*<0.001); locking door action (*MD*= -1.017 *SE*=0.211, *t*=-4.814, *p*<0.001), planting action (*MD*= -0.950 *SE*=0.211, *t*=-4.498, *p*<0.001); reading action (*MD*= -1.483 *SE*=0.211, *t*=-7.024, *p*<0.001) and laundry action (*MD*= -1.450 *SE*=0.211, *t*=-6.866, *p*<0.001). The making a list action is rated significantly different from the painting action (*MD*= 0.850 *SE*=0.211, *t*= 4.025, *p*=0.003). The watering of the plant action is rated significantly different from the painting action (*MD*= 1.433, *SE*=0.211, *t*= 6.787, *p*<0.001), locking door action (*MD*= 0.717 *SE*=0.211, *t*=3.393, *p*=0.029) and planting action (*MD*= 0.783 *SE*=0.211, *t*= 3.709, *p*=0.009). The painting of the wall action is rated significantly from ringing the bell action (*MD*= -0.817 *SE*=0.211, *t*=-3.867, *p*=0.005); locking door action (*MD*= -0.717 *SE*=0.211, *t*=-3.393, *p*=0.029); reading action (*MD*= -1.183 *SE*=0.211, *t*=-5.603, *p*<0.001) and laundry action (*MD*= -1.150 *SE*=0.211, *t*=-5.445, *p*<0.001). There was no significant main effect of agents F (1, 29) =1.466 *p* =0.236. There are no significant interactions between Actions and Agents (p>0.05).

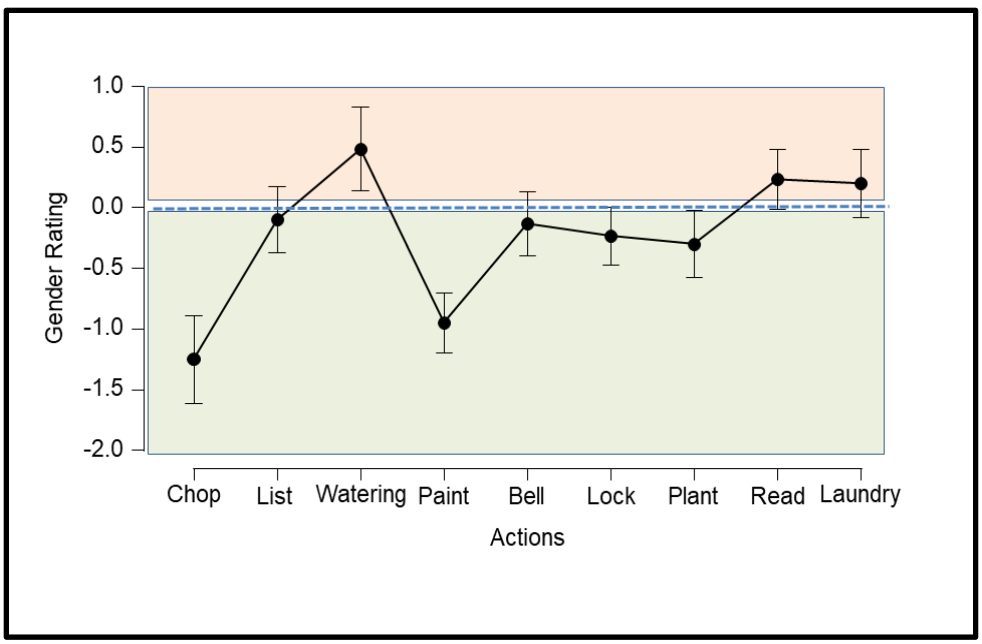

Figure 20. The technical-name condition gender attribution ratings to what extent actions are pronounced responses (more femininity or masculinity) across all 9 actions by participants.

Table 8

The BIF, FA, and Gender attribution summary of the technical condition. The significance level is $p<0.05$; significant results are shown in bold.

Technical name Condition	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
BIF	$F(1, 29) = 20.465$ **$p < .001$***** $\eta^2 = 0.294$	$F(1, 29) = 6.154$ **$p = 0.019$** $\eta^2 = 0.013$	$F(8, 232) = 0.942$ $p = 0.482$
FA/Competency	$F(1, 29) = 16.192$ **$p = .001$***** $\eta^2 = 0.248$	$F(1, 29) = 1.074$ $p = 0.309$	$F(8, 232) = 0.706$ $p = 0.686$
Gender Attribution	$F(1, 29) = 14.003$ **$p = .001$***** $\eta^2 = 0.208$	$F(1, 29) = 1.466$ $p = 0.236$	$F(8, 232) = 1.223$ $p = 0.286$

3.2.4 .4 Personification

The 30 participants’ responses to 18 videos (9 human and 9 robot) which makes up 540 responses in total, were analyzed based on 9 themes. The most frequent theme is type refers to defining agents, either robot or human (Figure 21).

i.Identity

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in identity attribution (1, N = 103) = 0.572, p =0.450. It means there is no significant identity attribution to human and robot agents between women and men participants. They equally attribute identity to the robot and human agents by calling them their names.

ii.Personality Traits

Pearson’s chi-square test of association revealed a significant relationship between the gender of participants and agents in personality trait attribution (1, N = 34) <0.001, p =0.982. There is a significant identity attribution to human and robot agents between women and men participants. Women attributed to personality traits both agents are estimated to be -0.018 times less likely to attribute identity compared to men. A 95% confidence interval indicated high precision (-1.552 to 1.516) in estimating the population odds ratio.

iii.Gender Attribution

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality gender attribution (1, N = 61) = 3.421, p =0.064.

iv.Roles and Occupations

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality roles and occupation attribution (1, N = 44) = 1.490, p =0.222.

v.Physical States

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 2) = 0.105, p =0.746.

vi. Type

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 198) = 0.275, p =0.600.

vii.Actions

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality action attribution (1, N = 59) = 0.001, p =0.971.

viii.Other

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in a personality other (non-fit abovementioned categories) (1, N = 28) = 1.972, p =0.160.

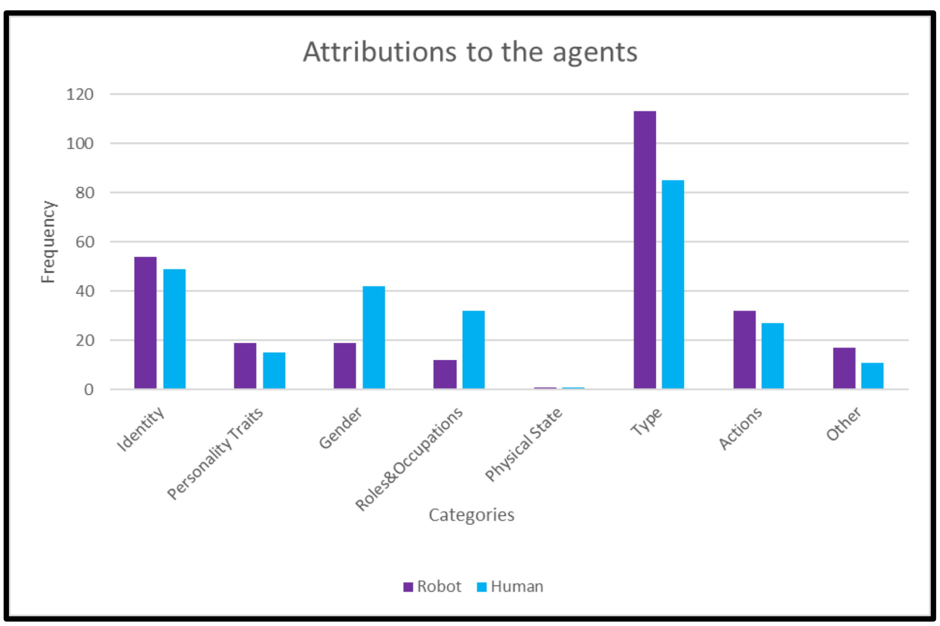

Figure 21. The technical-name condition personification attributions' bar graph across 9 categories for the robot and the human agent.

3.2.5 No-Name Condition

3.2.5 .1 BIF rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =12.992 $p<0.001$, η^2 =0.220, meaning that participants rated agent's agency level to actions differently (Figure 22).

The post hoc analysis shows that the chop action BIF level is rated significantly different from ringing the bell action (MD= -0.567 SE=0.080, t=-7.117, p<0.001) and reading action (MD= -0.400 SE=0.080, t= -5.024, p<0.001). The making a list action is rated significantly different from the ringing the bell action (MD= -0.467 SE=0.080, t= -5.861, p<0.001) and reading action (MD= -0.300 SE=0.080, t=-3.768, p<0.001). The watering the plant action is rated significantly different from ringing the bell action (MD= -0.600 SE=0.080, t=-7.536, p<0.001) and reading action (MD= -0.433 SE=0.080, t=-5.442, p<0.001). The painting the wall action is rated significantly from ringing the bell action (MD= -0.450 SE=0.080, t=-5.652, p<0.001) and reading action (MD= -0.283 SE=0.080, t=-3.559, p=0.016). The ringing the bell action is rated significantly different from the locking door action (MD= 0.367 SE=0.080, t= 4.605, p<0.001), planting action (MD= 0.533 SE=0.080, t=6.698, p<0.001) and laundry action (MD= 0.550 SE=0.080, t=6.908, p<0.001). The planting action is significantly different from the reading action (MD= -0.367 SE=0.080, t=-4.605,

$p<0.001$). The reading action is significantly different from the laundry action (*MD*= 0.383 *SE*=0.080, *t*=4.815, $p<0.001$). There was a significant main effect of agents $F(1, 28) = 16.749$ $p<0.001$, $\eta^2 = 0.027$ (Figure 23). The post hoc analysis shows that the robot agent's agency level is significantly different from the human agent (*MD*= 0.133 *SE*=0.033, *t*=-4.093, $p<0.001$). There are no significant interactions between Actions and Agents ($p>0.05$).

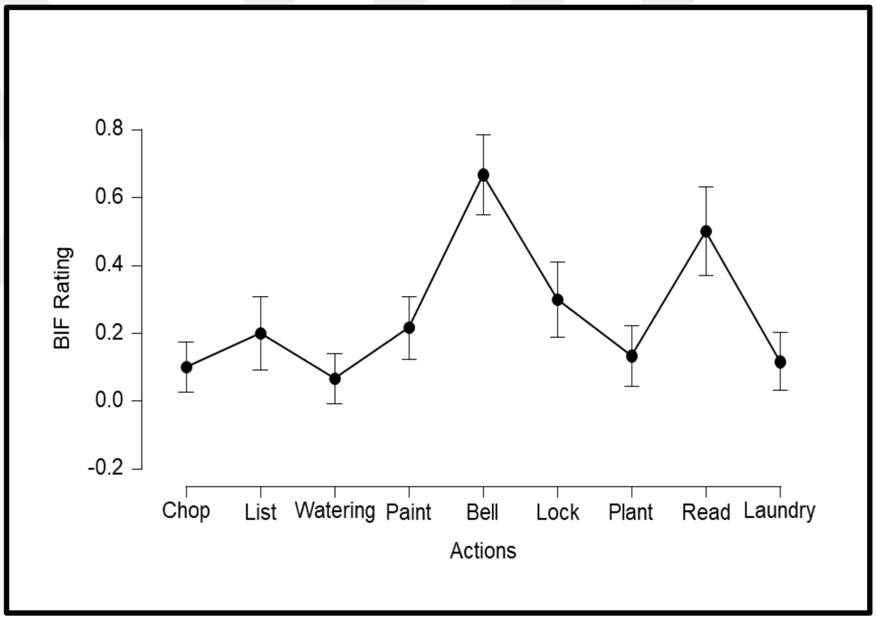

Figure 22. The no-name condition BIF ratings across all 9 actions by participants

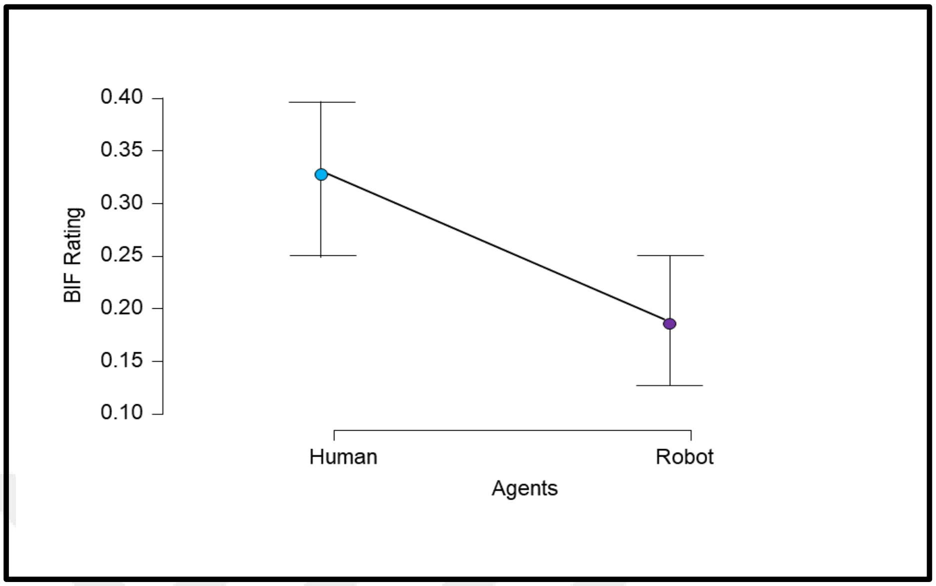

Figure 23. The no-name condition BIF rating difference between human and robot agents

3.2.5 .2 FA/Competency rating

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =13.760 $p<0.001$, η^2 =0.247, meaning that people rated all agents' competency levels differently (Figure 24).

The post hoc analysis shows that the chop action competency level is rated significantly different from making a list action (*MD*= -1.117 *SE*=0.217, *t*=-5.153, *p*<0.001), watering action (*MD*= -0.783 *SE*=0.217, *t*=-3.615, *p*=0.013), painting action (*MD*= -0.900 *SE*=0.217, *t*=-4.153, *p*=0.002); ringing the bell action (*MD*= -1.250 *SE*=0.217, *t*=-5.768, *p*<0.001); planting action (*MD*= -1.350 *SE*=0.217, *t*=-6.230, *p*<0.001), reading action (*MD*= -1.150 *SE*=0.217, *t*=-5.307, *p*<0.001) and laundry action (*MD*= -1.067, *SE*=0.217, *t*=-4.922, *p*<0.001). The making a list action is rated significantly different from the locking the door action (*MD*= 1.383, *SE*=0.217, *t*= 6.384, *p*<0.001). The watering the plant action is rated significantly different from the locking the door action (*MD*= 1.050 *SE*=0.217, *t*=4.845, *p*<0.001). The painting of the wall action is rated significantly from locking the door action (*MD*= 1.167 *SE*=0.217, *t*=5.384, *p*<0.001). The ringing the bell action is rated significantly different from the locking the door action (*MD*= 1.517 *SE*=0.217, *t*=6.99, *p*<0.001). The locking the door action is significantly different from the planting action (*MD*= -1.617 *SE*=0.217, *t*=-7.461, *p*=0.001), reading action (*MD*= -1.417 *SE*=0.217, *t*=-6.538, *p*=0.001) and laundry action (*MD*= -1.333 *SE*=0.217, *t*=-6.153, *p*=0.001). There was no significant main effect of agents F (1, 29) =0.296 p =0.590. There are no significant interactions between Actions and Agents (p>0.05).

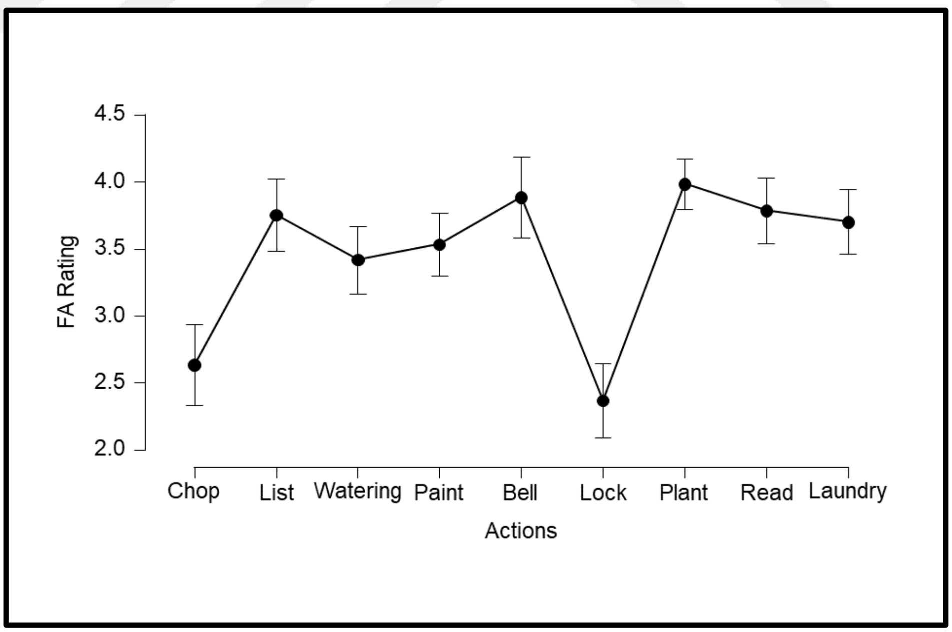

Figure 24. The no-name condition FA ratings across all 9 actions by participants

3.2.5.3 Gender Attribution

According to 2(Agent) x 9 (Action) ANOVA analysis, there was a main effect of actions F (1, 29) =15.433 $p<0.001$, η^2 =0.249, meaning that participants attributed different gender levels to the agents across actions (Figure 25). The post hoc analysis shows that the chop action gender attribution level is rated significantly different from making a list action (*MD*= -1.167 *SE*=0.232, *t*=-5.023, *p*<0.001), watering action (*MD*= -1.717 *SE*=0.232, *t*=-7.391, *p*<0.001); ringing the bell action (*MD*= -1.100 *SE*=0.232, *t*=-4.736, *p*<0.001); planting action (*MD*= -1.200 *SE*=0.232, *t*=-5.166, *p*<0.001); reading action (*MD*= -1.767 *SE*=0.232, *t*=-7.606, *p*<0.001) and laundry action (*MD*= -1.767 *SE*=0.232, *t*=-7.606, *p*<0.001). The making a list action is rated significantly different from the painting action (*MD*= 0.900 *SE*=0.232, *t*= 3.875, *p*=0.005). The watering of the plant action is rated significantly different from the painting action (*MD*= 1.450, *SE*=0.232, *t*= 6.243, *p*<0.001), locking door action (*MD*= 1.00 *SE*=0.232, *t*=4.305, *p*<0.001). The painting of the wall action is rated significantly from ringing the bell action (*MD*= -0.833 *SE*=0.232, *t*=-3.588, *p*=0.015), planting action (*MD*= -0.933 *SE*=0.232, *t*=-4.018, *p*=0.003); reading action (*MD*= -1.500 *SE*=0.232, *t*=-6.458, *p*<0.001) and laundry action (*MD*= -1.500 *SE*=0.232, *t*=-6.458, *p*<0.001). There was no significant main effect of agents F (1, 29) =0.810 *p* =0.376, There are no significant interactions between Actions and (p>0.05).

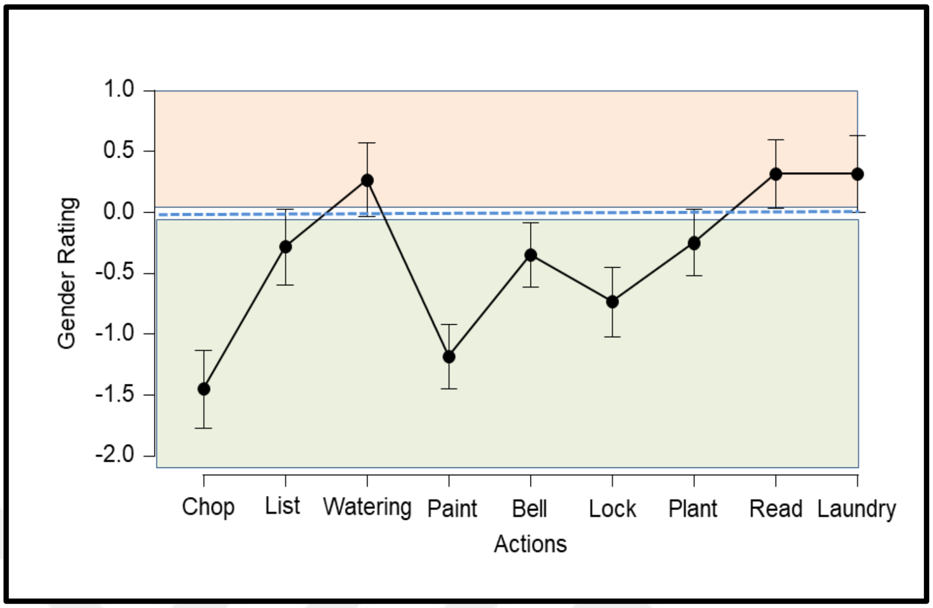

Figure 25. The no-name condition gender attribution *ratings to* what extent actions are pronounced responses (more femininity or masculinity) across all 9 actions by participants.

Table 9

The BIF, FA, and Gender attribution summary of the no-name condition. The significance level is p<0.05; significant results are shown in bold.

No name Condition	**Main Effect of Action**	**Main Effect of Agent**	**Agent x Action Interaction**
BIF	F (1, 29) =12.992 **p <.001***** η^2 =0.220	F (1, 28) =16.749 **p =<.001***** η^2 =0.027.	F (8, 232) =1.407 p =0.194
FA/Competency	F (1, 29) =13.760 **p =.001***** η^2 =0.247	F (1, 29) =0.296 p =0.590	F (8, 232) =1.096 p =0.367

Gender Attribution	F (1, 29) =15.433 **p =.001***** η^2 =0.249	F (1, 28) =0.810 p =0.376	F (8, 232) =0.880 p =0.534

3.2.5.4 Personification

The 30 participants' responses to 18 videos (9 human and 9 robot) which makes up 540 responses in total, were analyzed based on 9 themes. The most frequent theme is type refers to defining agents, either robot or human Figure 26).

i.Identity

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in identity attribution (1, N = 103) = 0.572, p =0.450. It means there is no significant identity attribution to human and robot agents between women and men participants. They equally attribute identity to the robot and human agents by calling them their names.

ii.Personality Traits

Pearson's chi-square test of association revealed a significant relationship between the gender of participants and agents in personality trait attribution (1, N = 34) <0.001, p =0.982. There is a significant identity attribution to human and robot agents between women and men participants. Women attributed to personality traits both agents are estimated to be -0.483 times less likely to attribute identity compared to men. A 95% confidence interval indicated high precision (-1.612 to 0.646) in estimating the population odds ratio.

iii.Gender Attribution

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality gender attribution (1, N = 61) = 3.421, p =0.064.

iv. Roles and Occupations

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality roles and occupation attribution (1, N = 44) = 1.490, p =0.222.

v. Physical States

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 2) = 0.105, p =0.746.

vi. Type

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality type(human/robot) attribution (1, N = 198) = 0.275, p =0.600.

vii.Actions

Pearson’s chi-square test of association revealed a non-significant relationship between the gender of participants and agents in personality action attribution (1, N = 59) = 0.001, p =0.971.

vii.Other

Pearson's chi-square test of association revealed a non-significant relationship between the gender of participants and agents in a personality other (non-fit abovementioned categories) (1, N = 28) = 1.972, p =0.160.

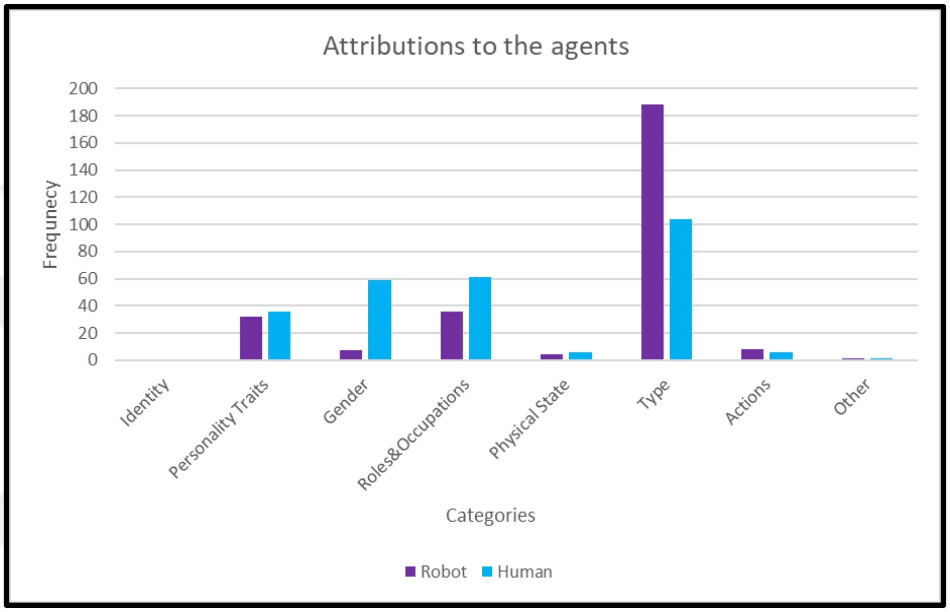

Figure 26. The no-name condition personification attributions' bar graph across 9 categories for the robot and the human agent.

3.3 Pre-Study Results

According to pre-study results where the participants rated the 9 actions on the feminine and masculine spectrum previously to understand whether participants have pre-conceived evaluations about the actions' gendered notions. The result shows that chopping tree (*M*: -2.03 *SD*:0.21) and painting the wall (*M*: -0.83 *SD*:0.14) were rated on masculine spectrum. Watering plant (*M*:0.98 *SD*:0.20), washing clothes (*M*:1.05 *SD*: 0.21) were rated on a feminine spectrum. Locking door (*M*:0.30 *SD*: 0.21), making a list (*M*:0.50 *SD*: 0.13) and reading book (*M*:0.20, *SD*:0.11) were rated as feminine side but closer to the neutral. Growing a garden (*M*:0.05, *SD*:0.08) and ringing the bell (*M*:0.05 *SD*: 0.07) were rated closer to zero-neutral.

3.4 Regression Analysis

A linear regression analysis helps to understand to what extent participants' gender attributions to the actions in the pre-study predict their gender attribution to the agents in the explicit part. Our results showed that it does not significantly predict human agent gender attribution F *(*1,148) =0.611, *p*=0.436; β=0.142. Also, it did not predict gender attribution for robot agents *F (*1,148) =0.922, *p*=0.338; β=-0.112.

3.5 Implicit Part Results

3.5.1 IAT-1 Results

The first implicit test is Woman/Man -Dependent/Independent; I conducted 2 (Gender) X 5(Naming) ANOVA. The results show a significant effect of gender F (1, 140) =6.601 1 *p* =0.011, η^2 =0.045, meaning that women and men participants reacted differently. The post hoc analysis shows that women participants' implicit responses significantly differed from men's (*MD*=0.136, *SE*= 0.053, *t*=2.569, *p*=0.011). There was no significant effect of naming (group effects) F (4, 140) =0.600 1 *p* =0.663. There was no significant interaction between Gender and Naming F (4, 140) =0.031 1 *p* =0.998 (Figure 27).

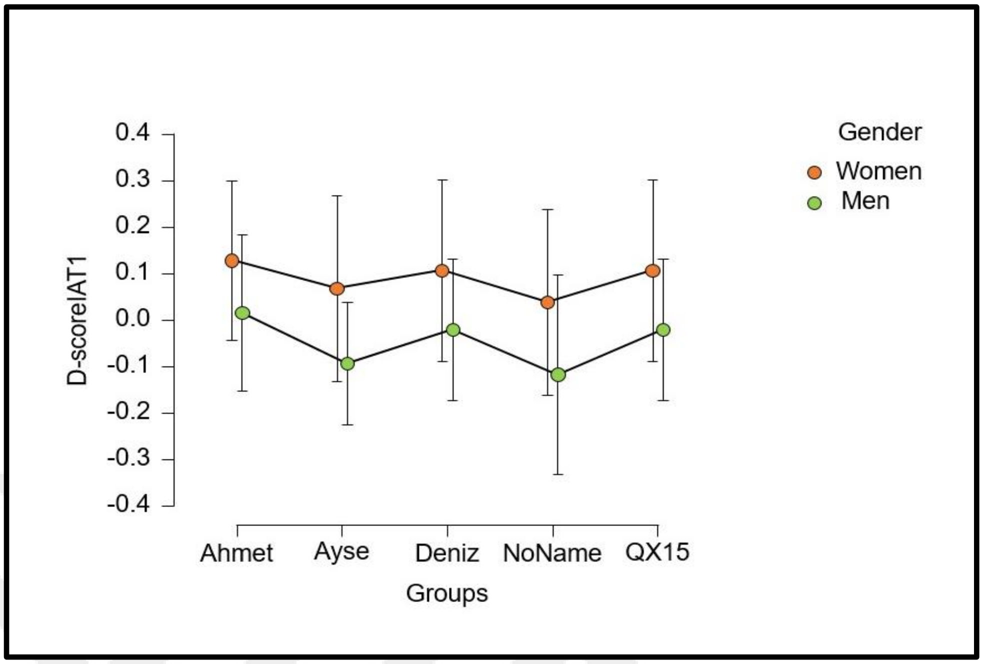

Figure 27. The IAT-1 d score comparison across 5 name categories for the woman and man participants

3.5.2 IAT-2 Results

The first implicit test is Robot/Human -Dependent/Independent; I conducted 2 (Gender) X 5(Naming) ANOVA. The results show no significant effect of gender F (1, 140) =0.310 1 *p* =0.578, meaning that women and men participants do not react differently. There was no significant effect of naming (group effects) F (4, 140) =1.654 1 *p* =0.164. There was no significant interaction between Gender and Naming *F* (4, 140) =0.962 1 *p* =0.431(Figure 28).

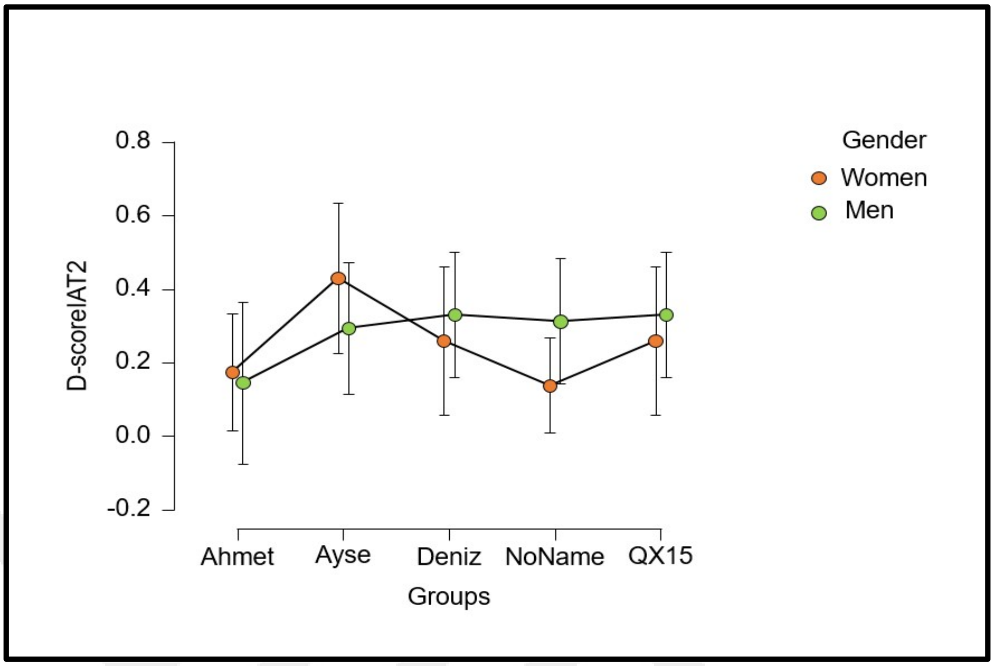

Figure 28. The IAT-2 d score comparison across 5 name categories for the woman and man participants.

3.5.3 IAT-3 Results

The first implicit test is Robot/Human -Woman/Man; I conducted 2 (Gender) X 5 (Naming) ANOVA. The results show a significant effect of gender F (1, 140) =37.131 1 p <0.001, η^2 =0.210, meaning that women and men participants reacted differently. The post hoc analysis shows that women participants' implicit responses significantly differed from men's (*MD*=-0.288, *SE*= 0.047, t=-6.093, p=0.001). There was a significant effect of naming (group effects) F (4, 140) =11.523 1 p <0.001, η^2 =0.248. The post hoc analysis shows that the Ahmet (Masculine) condition's implicit results is significantly different from Ayse (Feminine) condition (*MD*= 0.400, *SE*= 0.075, t=5.351, p=0.001), Deniz

(Neutral-Name) condition (*MD*=0.388, *SE*= 0.075, *t*=5.194, *p*=0.001); No-Name condition (*MD*= 0.423, *SE*= 0.075, *t*=5.662, *p*=0.001) and QX15 (Technical) condition (*MD*=0.388, *SE*= 0.075, *t*=5.194, *p*=0.001). There was a significant interaction between Gender and Naming *F* (4, 140) =4.361 1 *p* =0.002, η^2 =0.111. The women participants in masculine conditions' reactions are significantly different from women participants in feminine condition (*MD*=0.609, *SE*= 0.106, *t*=5.762, *p*=0.001), women participants in neutral name condition (*MD*=0.637, *SE*= 0.106, *t*=6.025, *p*=0.001); women participants in no-name condition (*MD*=0.513, *SE*= 0.106, *t*=4.848, *p*=0.001); women participants in technical name condition (*MD*=0.637, *SE*= 0.106, *t*=6.025, *p*=0.001) and man participants in no-name condition (*MD*=0.364, *SE*= 0.106, *t*=3.445, *p*=0.034). The men participants in masculine conditions reactions are significantly different from women participants in the feminine condition (*MD*=0.579, *SE*= 0.106, *t*=5.476, *p*=0.001), women participants in neutral name condition (*MD*=0.607, *SE*= 0.106, *t*=5.739, *p*=0.001); women participants in no-name condition (*MD*=0.482, *SE*= 0.106, *t*=4.562, *p*=0.001); women participants in technical name condition (*MD*=0.607, *SE*= 0.106, *t*=5.739, *p*=0.001) and man participants in no-name condition (*MD*=0.364, *SE*= 0.106, *t*=3.445, *p*=0.034). The women participants in feminine conditions reactions are significantly different from men participants in feminine condition (*MD*=-0.388, *SE*= 0.106, *t*=-3.670, *p*=0.015), men participants in neutral name condition (*MD*=-0.440, *SE*= 0.106, *t*=-4.156, *p*=0.003); men participants in technical name condition (*MD*=-0.440, *SE*= 0.106, *t*==4.156, *p*=0.001). The men participants in the feminine condition's reactions are significantly different from women participants in the neutral name condition (*MD*=0.416, *SE*= 0.106, *t*= *3*.933, *p*=0.006); women participants in the technical name condition (*MD*=0.416, *SE*= 0.106, *3*.933, *p*=0.006). The women participants in neutral name conditions reactions are significantly different from men participants in neutral name conditions (*MD*=-0.467, *SE*= 0.106, *t*= -*4*.419, *p*=0.001); women participants in technical name condition (*MD*=0.416, *SE*= 0.106, *3*.933, *p*=0.006). The men participants in the neutral name conditions reactions are significantly different from men participants in the technical name condition (*MD*=-0.467, *SE*= 0.106, t=-*4*.419, *p*=0.001). The women participants in technical name conditions reactions are significantly different from men participants in technical name conditions (*MD*=-0.467, *SE*= 0.106, t=-*4*.419, *p*=0.001) (Figure 29).

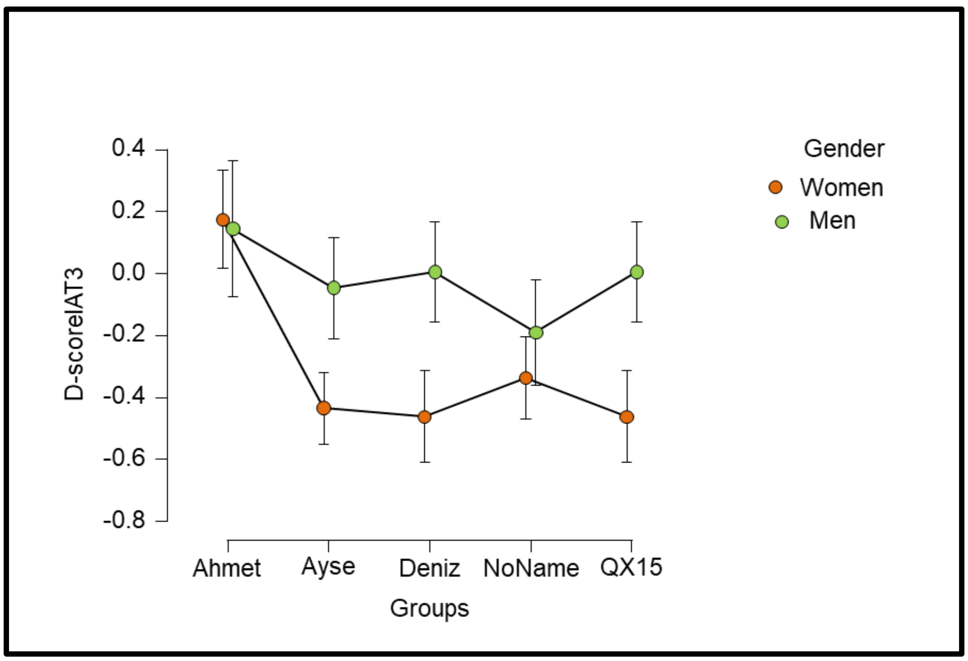

Figure 29. The IAT-3 d score comparison across 5 name categories for the woman and man participants.

Table 10

The IAT-1, IAT-2, and IAT-3 d score comparison.

Implicit Tests	**Gender**	**Naming**	**Gender X Naming Interaction**
IAT-1 (Woman/Man - Dependent/Independent)	$F(1,140) = 6.601$ **$p = 0.011$** $\eta^2 = 0.045$	$F(1, 140) = 0.600$ $p = 0.663$	$p > 0.05$
IAT-2 (Robot/Human - Dependent/Independent)	$F(1, 140) = 0.310$ $p = 0.578$	$F(1, 140) = 1.654$ $p = 0.164$	$p > 0.05$
IAT-3 (Robot/Human - Woman/Man)	$F(1, 140) = 37.131$ **$p < 0.001$***** $\eta^2 = 0.210$	$F(1,140) = 11.523$ **$p < 0.001$***** $\eta^2 = 0.248$	$F(1, 140) = 4.361$ **$p = 0.002$** $\eta^2 = 0.111$.

3.6 Comparison of Explicit and Implicit Parts

3.6.1 Feminine Condition

A correlation analysis was conducted to examine to what extent explicit answers could relate to participants' implicit association test reactions. Our results showed that no explicit part variables were predicted by implicit score ($p>0.05$).

3.6.2 Masculine Condition

In this condition, results showed that none of the explicit part variables predicted by implicit score ($p>0.05$).

3.6.3 Neutral Condition

In this condition, results show that the BIF scores for robot agent are negatively correlated by IAT2 (Robot/Human -Dependent/Independent) r=-0.376, p=0.041 and positively predicted by IAT 3(Robot/Human -Woman/Man) r=0.504, p=0.005.

3.6.4 Technical Condition

In this condition, results show that the BIF scores for robot agents are negatively correlated by IAT1 (Woman/Man Dependent/Independent) r=-0.462, p=0.010.

3.6.5 No Name Condition

In this condition, results show that the BIF scores for human agents are positively correlated by IAT2 (Robot/Human -Dependent/Independent) r=0.382, p=0.037.

CHAPTER IV

4. DISCUSSION

Researchers in HRI are increasingly studying gendering robots [43, 51, 53, 54, 56, 17, 18]. This study explored whether gender could potentially affect people's autonomy attributions to human and robot agents. As we discussed autonomy in three main works of literature, it can be observable that one person's or agent's perceived autonomy level affects the attitudes and treatments towards them. Then autonomy is taken as a state of being independent and acting separately. Moreover, understanding one's capacity and ability to act independently relies on the agency concept. It provides a fundamental understanding of that person's abilities/ capabilities to be autonomous. About agency, including measurable agency concepts concerning autonomy, becomes important. That is why using BIF provides a tool for creating a stimulus set for our agents (human and robot). Implementing BIF is realized by transforming the written actions in that form into animations performed by the agents. By doing this, we would have the opportunity to measure how participants assign the level of agencies to these agents and ask about the other dimension of autonomy (doer of the tasks and functional appropriateness of the performed action). The last question is asking about the gender of the agent because, in the literature, we were informed that gender is the main component of accepting/rejecting one's autonomy. To provide more elaborative measurement, I also included implicit measurement (IAT) to see which words related to the concepts (gender, robot, autonomy) are prone to associate faster (showing strong association).

The results show that my hypothesis for BIF scores failed to reject human and robot agents' agency scores change depending on the naming group difference. Rather, in general for agency-level attribution, the agent type has the main effect on BIF evaluations meaning that the human agent is attributed to a higher-level agency relative to the robot agent. In my hypothesis for FA scores is failed to reject human and robot agents' agency scores change depending on the naming group difference. People equally attribute the level of functional appropriateness of doing the action for both agents in all name conditions. In my hypothesis for GA scores, it failed to reject human and robot agents' gender attribution scores change depending on the naming group difference. In gender attribution, people attribute similar femininity and masculinity to robot and human agents. For the human condition, except for the feminine-name condition, the human agent was mostly attributed to masculinity. For the robot agent, all name conditions of the robot were attributed to masculinity. When carefully examined, the significance level of the naming effect is within limits, which may indicate that applying other strong name manipulation to agents in future studies may potentially reveal the naming effect. To be sure about whether people attributed gender based on the agents or the potential effect of the actions' gendered level, it was found that people did not attribute gender based on actions' perceived gender level.

In the level of personification analysis, the most attributed categories among 9 categories are the name and type of the agent. People attribute names assigned and type mostly to robot and human agents. As line with hypothesis, I reject the null hypothesis that for feminine names, masculine names, and neutral name conditions (human-names), name is attributed more than the type of the agent. For no name and technical name condition, type is more attributed to the agents than name. As line with my hypothesis, I reject the null hypothesis that the implicit part results show that women participants responded to women-independent/ men-dependent association faster than men-independent/ women-dependent. In contrast, men participants responded to men-independent/ women-dependent faster than women-independent/ men-dependent. Both men and women participants responded to human-independent/ robot-dependent association faster than robot-independent/ human-dependent. Both men and women participants responded to the human-women/ robot-men association faster than the human-men/robot-women association.

4.1. Contributions

The study makes contribution to HRI literature with three aspects. First, this study builds on previous work on gendered robots by integrating with autonomy. So far, we have seen that the concept of autonomy in robotics literature is defined as the state of industrial robots performing certain operations without human intervention. In the social robot literature, it is defined as the state of interacting with humans without human intervention. Since these definitions do not give this concept of autonomy in a measurable form, making it measurable in 3 basic elements is one of my significant contributions to this field. I thought that autonomy is more than just a 'state'; it is closely related to the fact that autonomy requires a capacity to perform an action (1), how well agent does the action (2), and to what extent the human-like qualities are attributed to the agents (3). Measuring this way will provide a comprehensive measurement method and bring innovation and contribution to literature.

Secondly, my study demonstrates the importance of integrating animations technique to HRI literature. Humans should share a similar socio-physical space with robots to study our real-world interactions with them. Sometimes these interactions have some limitations, and today's technology may not provide comprehensive depictions of robots, which is why it may not reflect real-life interaction reality. Or sometimes participants feel uncomfortable while interacting with real-life robots [9]. On the other side, autonomy is the concept that requires similar real-life evaluations of this capacity through the actions taken(agency), which makes the more vivid interaction required. BIF actions, which are the agency measurement, are in written format that participants read and chose accordingly. To create a more interactive context, I animated these actions. Therefore, animation techniques allow for the creation of interactive content, the manipulation of context, and the manipulation of characters. We can observe their responses as robots interact with humans in specific social and cultural contexts.

Thirdly, the implicit measurement of autonomy, robot, and gender concepts/categories are novel measurement that provides comprehensive literature measurement. The purpose of creating such a type of measurement is to determine whether people's perceptions of robots and humans, women and men, and dependents and independents align with the findings presented by various kinds of literature when taken together. According to literature in gender studies and political science, women and femininity are associated with a lack of independence, or humans design robots with a sense of dependency.

4.2. Implications

4.2.1 Explicit part

Behavioral Identification Form/BIF

As I highlighted in the introduction, I expect differences in agency-level attribution of actions in human and robot agents, as found in the study [33]. It may indicate that animated versions of BIF actions lead participants to evaluate robots as humans in how they act. People are more inclined to choose low-level agency agents, which may be because the actions are thought to require more physically done jobs and are less likely to be mental effort-requiring ones. It provides insight that participants are more likely to evaluate the robots' and humans' agency level in physical-focused jobs at the same level regardless of their gender level.

Personification

In this part, participants are free to write anything after they watch the agent do the action. Participants equally and mostly attribute identity to both agents, meaning that when they do the same action and have the same gendered names, they attribute identity to both agents. The second most attributed thing is the type of agents. It was seen mostly in technical and no-name conditions. It implies that participants prefer to define agents in types when the agents have no-name or technical names. They primarily define the robot agent with its type more than the human. It implies that when the robot and human agents are assigned human-gendered names, people evaluate this robot agent as having an identity like humans.

Functional Appropriateness/Competency

In line with the BIF scores, participants evaluate the correctness of doing an action for human and robot agents at the same level. It may imply that doing an action appropriately or not does differ according to the agent's type. Participants do not evaluate the human agent as more correctly capable of doing the action by default than the robot agent. Instead, if the robot has similar body capability as humans, it is also evaluated in the limit of its way of executing the action. It may imply that when robots are placed in jobs that humans are capable of, their competency is not immediately disregarded; instead, it provides more room for robots to be accepted.

Gender Attribution

Participants attribute different gender levels to the agents varying according to actions. However, participants did not attribute gender differently. It means that for every action, the human and robot agent attribute similar levels of gender.

To be confident about whether people attribute gender based on the actions' perceived gender level (such as chopping a tree is evaluated as more masculine, that is why the agent attributed more masculine), participants completed a pre-study. The perceived gender level of these actions does not predict the participants' gender attribution to the agents. It shows that participants make gender attribution according to their perception of the agent during the experiment. Although I received feedback from participants, such as "since chopping a tree was seen as a more masculine action, I rated this agent masculine," there is no significant interaction in the regression analysis.

4.2.2 Implicit Part

One of the measurements that formed the second part of the study is the implicit measurement. The motivation for using this measure is to compare the explicit measure results with the implicit measure's responses. First, I would like to understand to what extent people unconsciously match masculine and feminine to dependent and independent states in the implicit measure. Likewise, I would like to understand to what extent robots related words to the dependent and independent related categories at an unconscious level.

This information will allow me to understand the answers given according to different noun categories in explicit measurement more clearly.

a) <u>IAT 1</u>

IAT-1 includes independent/dependent and woman/man categories and their related words. I hypothesized that men participants reacted more quickly to the dependent/women and independent/men association than independent/women and dependent/men, indicating a strong linkage in all conditions. Women participants react fast to independent/women and dependent/men faster than dependent/women and independent/men in all conditions. It may be due to the idea that in the socio-cultural context of

Türkiye, women are portrayed, treated, and experienced domination by men frequently, so men participants' inclination relies on there. For women participants who are not submitting these portrayals, they challenge and object by highlighting women's independence to strengthen their association.

b) IAT 2

IAT-2 includes independent/dependent and robot/human categories and related words. Both women and men participants react fast to independent/human and dependent/robot association than dependent/human and independent/robot. It may imply that robots are evaluated as "machines" controlled by humans and cannot act by themselves. This has quite different results from what was found in the explicit part. Both autonomy components for robots are similar to human agents; based on these, we can infer that people attribute autonomy to robots. However, humans' unconscious minds still perceive them as "dependent agents."

c) IAT 3

IAT-3 includes woman/man and human/robot categories and their related words. In the other naming conditions, women's response was more strongly associated with robot/man and human/woman categories than human/man and robot/woman compared to men. Women have been inclined to show stronger biases toward this association compared to men.

4.3 Limitations and Future Directions

Studying has several limitations. In the explicit part, a pilot study was conducted in order to select a gender-neutral human agent. Although there are many alternatives in finding characters, gender-neutral-looking human agents are not found in many databases, so this selection was not made among many options. Rather I adjusted a one-appropriate character to look more gender-neutral, meaning that a person cannot immediately decide whether this agent is masculine or feminine characteristics. Our results could be generalized by using a variety of gender-neutral human characters in future studies.

This study can be followed up by determining whether people's evaluations of a gendered human and robot agent can change (instead of a gender-neutral robot/human). It is also possible to evaluate autonomy perception by integrating the Aşkın et al. (2023) gendered actions study that examines how neutral agents perform gendered behaviors. It is possible to investigate whether a masculine-looking robot is considered less competent if it does makeup or takes care of a child. This would show how effective actions are in autonomy and gender relations.

A second limitation is the way of animating the actions. Some actions in the study, -such as chopping a tree, and locking a door-, need to be appropriately animated. For example, while the key must rotate with the hand when the agent inserts the key into the hole, this action is animated simply as inserting the key into the hole. Or, in the act of cutting a tree, since no part of the tree (wood) fell while the ax was passing through the tree (this did not make the participants think that the tree was cut down), it is because the action was not adequately animated. However, this situation has given us an essential insight that the situation where an action cannot be performed properly is evaluated similarly in both human and robot agendas. In other words, the fact that the action was not performed well enough evaluated both agents as not doing the job correctly. This means that we do not automatically rate robots as less capable than humans or as humans as more capable than robots.

In the explicit part of the videos, the robot stimuli are dynamic-moving, another limitation of our study. However, in the implicit part, where the stimuli are words. This raises the question of whether the modality differences can affect the participant's evaluations. Future studies can test this possibility by integrating the same modalities in these two parts.

The third limitation is naming manipulation effectiveness. The gendered names were selected based on the prevalence of the names in Türkiye in 2020 in the Ministry of the Interior's database. Since the Ministry of Interior published the previous year's results, such as 2020 data published in 2021, I could not use the most updated/prevalent names. Also, the sample we collected data from has different demographic characteristics compared to the rest of the population. That is why a pilot study needs to be conducted to understand what kinds of names are frequently used and then use these names to provide a more accurate representation of the current demographic context.

The fourth limitation is there is not another coder for the qualitative analysis while determining the categories. Since more than one coder would provide higher inter-rater reliability, there should be at least one more rater to determine the categories.

As a future direction for this study, participants' familiarity with technology can be included to measure autonomy perception. Since people's usage of these tools:

experiencing vivid interactions via games, virtual reality devices and ability to communicate with artificial intelligence and bots can change their autonomy perception. I received feedback from the participants who are familiar with video games, some of which liked the way actions are animated but some of them criticized animations' real-life depictions. That is why including technological familiarity questionnaire provides more insights about the interaction with the robot.

Moreover, the actions I animated are heavily based on physical capabilities, except reading the book action. There is no action category corresponding to mental capacities in the stimulus set used in this study. It may influence the autonomy evaluations about the agents, which is why including mental and physical actions together with gender manipulations may be investigated in future studies.

4.4 Conclusion

This thesis aims to understand the effect of gender perception on the autonomy evaluations of an entity. The agency and competency, which are the autonomy components, do not differ between humanoid robots and human agents, showing that humanoid robots could also be evaluated as human capacity and state. Even if the actions are perceived as not adequately performed by these agents, indicating incompetence, human, and robot agents were found similarly. Although the level of humanoid features attributed to humans and robots appear in similar patterns in human-named conditions, we see that the human agent is more defined by name, meaning that participants call these agents in human-named conditions either

Ayşe, Ahmet or Deniz than the robot agent, basically they attribute identity to the human agent more than the robot. While we see that robots are defined more based on their types meaning that people are more likely to call robot agents "robots." This may indicate that although the robot is given human characteristics and capacities, it is possible for people to see the robot as its type.

The results in the implicit part showed us that people's stereotypes are strong at the unconscious level, unlike the explicit ones. Looking specifically at the gender and (in)dependency IAT, the stronger link in men participants between femininity and dependency concepts and the stronger connection between women participants in terms of femininity and independence. These are closely related to the gender inequality atmosphere in Turkey because women are seen as dependent. Men are more likely to match this femininity and dependency in the cultural context unconsciously; on the other side, women have a stronger bond to independence as they have to show reaction and resistance to protect their independence leading them to consolidate it more.

In the scope of this thesis, gender stereotypes may also affect the autonomy perception in robots, so there should be more room to investigate gender and autonomy in different contexts. Thus, the robot's design and placement in and for different areas can be changed according to implicit and explicit evaluations of gender and autonomy.

As a final note, throughout my thesis, I gave importance to using gender-inclusive language, I tried to handle the gender category in a non-binary way, and I always tried to make the measurements over a continuum. Since the developments on the subject of gender will always continue, I would like to say that the language, measurements, and comments in my work are not written from a sexist lens, and I hope that what is written here will not harm the subject of gender in case of future changes. If what is written here contradicts future changes, I would like to remind you that it is thought that these articles were written according to the gender perception valid for that period and that no adverse purpose was carried out.

BIBLIOGRAPHY

[1] Ackermann, T., & Zimmer, C. (2021). The sound of gender – correlations of name phonology and gender across languages. *Linguistics*, *59*(4), 1143–1177. https://doi.org/10.1515/ling-2020-0027

[2] agency. (n.d.). In *The Merriam-Webster.com Dictionary*. Retrieved October 17, 2022, from https://www.merriam-webster.com/dictionary/agency.

[3] Aşkın, G., Saltık, İ., Boz, T.E. *et al.* Gendered Actions with a Genderless Robot: Gender Attribution to Humanoid Robots in Action. *Int J of Soc Robotics* (2023). https://doi.org/10.1007/s12369-022-00964-0

[4] autonomy. (n.d.). In *The Merriam-Webster.com Dictionary*. Retrieved October 16, 2022, from https://www.merriam-webster.com/dictionary/autonomy

[5] Basow, S. A. (1980). *Sex-role stereotypes: Traditions and alternatives*. Thomson Brooks/Cole.

[6] Beer, J. M., Fisk, A. D., & Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of Human-Robot Interaction*, *3*(2), 74. https://doi.org/10.5898/jhri.3.2.beer

[7] Bekker, M. H., & van Assen, M. A. (2008). Autonomy-connectedness and gender. *Sex Roles*, *59*(7-8), 532–544. https://doi.org/10.1007/s11199-008-9447-x

[8] Breazeal, C., Gray, J., Hoffman, G., & Berlin, M. (2004). Social robots: Beyond tools to partners. ROMAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759), pp. 551–556.

[9] Castelo, N., & Sarvary, M. (2022). Cross-cultural differences in comfort with humanlike robots. *International Journal of Social Robotics*, *14*(8), 1865–1873. https://doi.org/10.1007/s12369-022-00920-y

[10] Dion, K. L. (1983). Names, identity, and self. *Names*, *31*(4), 245–257. https://doi.org/10.1179/nam.1983.31.4.245

[11] Echebarria Echabe, A. (2013). Relationship between implicit and explicit measures of attitudes: The impact of application conditions. *Europe's Journal of Psychology*, *9*(2), 231–245. https://doi.org/10.5964/ejop.v9i2.544

[12] Edwards, C. (2019, October 28). Bot Bonkers Sex robot reviews and 'unboxing' videos are latest bizarre YouTube trend. Retrieved from https://www.thesun.co.uk/tech/10442670/sex-robot-reviews-unboxing-videos-youtube/.

[13] Epifania, O. M., Anselmi, P., & Robusto, E. (2020). DscoreApp: A shiny web application for the computation of the Implicit Association Test D-Score. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02938

[14] Eskens, R. (2017). Is Sex with Robots Rape? *Journal of Practical Ethics*, 62–76. Retrieved from https://philpapers.org/rec/ESKISW.

[15] Esterle, L. (2022). Deep learning in multiagent systems. *Deep Learning for Robot Perception and Cognition*, 435–460. https://doi.org/10.1016/b978-0-32-385787-1.00022-1

[16] Eyssel, F., & Kuchenbrandt, D. (2011). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724–731. https://doi.org/10.1111/j.2044-8309.2011.02082.x

[17] Eyssel, F., & Hegel, F. (2012, July 27). (S)he's Got the Look: Gender Stereotyping of Robots1. Retrieved December 13, 2020, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1559-1816.2012.00937.x

[18] Eyssel, F., Kuchenbrandt, D., Bobinger, S., de Ruiter, L., & Hegel, F. (2012). 'If you sound like me, you must be more human'. *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '12*. https://doi.org/10.1145/2157689.2157717

[19] Feltman, R., and Elliot, A. (2014). "Competence," in Encyclopedia of Quality of Life and Well-Being Research, ed. A. C. Michalos (Dordrecht: Springer), doi: 10.1007/978-94-007-0753-5

[20] Formosa, P. (2021). Robot autonomy vs. human autonomy: Social robots, Artificial Intelligence (AI), and the nature of autonomy. *Minds and Machines*, *31*(4), 595–616. https://doi.org/10.1007/s11023-021-09579-2

[21] Frank, L., & Nyholm, S. (2017). Robot sex and consent: Is consent to sex between a robot and a human conceivable, possible, and desirable? *Artificial Intelligence and Law*, *25*(3), 305–323. https://doi.org/10.1007/s10506-017-9212-y

[22] Friedman, M. (2003). A conception of autonomy. *Autonomy, Gender, Politics*, 3–29. https://doi.org/10.1093/0195138503.003.0001

[23] Goodrich, M. A., & Schultz, A. C. (2007). Human-robot interaction: A survey. *Foundations and Trends® in Human-Computer Interaction*, *1*(3), 203–275. https://doi.org/10.1561/1100000005

[24] Gorar, M. (2021). *Female sexual autonomy, virginity, and honour-based violence with special focus on the UK*. Virtual Commons - Bridgewater State University. Retrieved May 12, 2022, from https://vc.bridgew.edu/jiws/vol22/iss5/5

[25] Griffiths, M. (1988). The Philosopher's Child. In *Feminist perspectives in philosophy* (pp. 89–90). essay, Macmillan.

[26] Gutiu, S. M. (2016). Chapter 8: The roboticization of consent. In *Robot Law* (pp. 186–212). essay. Retrieved from https://www.elgaronline.com/view/edcoll/9781783476725/9781783476725.00016.xml.

[27] Harris, L. T., & Fiske, S. T. (2011). Dehumanized perception: A Psychological Means to Facilitate Atrocities, Torture, and Genocide? *Zeitschrift Für Psychologie*, *219*(3), 175–181. https://doi.org/10.1027/2151-2604/a000065

[28] Harvard Implicit Association Test. (2011). *Project implicit*. About the IAT. Retrieved January 10, 2023, from https://implicit.harvard.edu/implicit/iatdetails.html.

[29] Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*(3), 390–421. https://doi.org/10.1037/a0018916

[30] Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Personality and Social Psychology Bulletin*, *31*(10), 1369–1385. https://doi.org/10.1177/0146167205275613

[31] Holroyd, J., & Saul, J. (2018). Implicit Bias and Reform Efforts in Philosophy: A Defence. *Philosophical Topics*, *46*(2), 71–102. https://www.jstor.org/stable/26927951

[32] Joubert, C. E. (1993). Personal names as a psychological variable. *Psychological Reports*, *73*(3_suppl), 1123–1145. https://doi.org/10.2466/pr0.1993.73.3f.1123

[33] Kim, T. W., & Duhachek, A. (2020). Artificial Intelligence and persuasion: A construal-level account. *Psychological Science*, *31*(4), 363–380. https://doi.org/10.1177/0956797620904985

[34] Kuchenbrandt, D., Häring, M., Eichberg, J. et al. Keep an Eye on the Task! How Gender Typicality of Tasks Influence Human–Robot Interactions. Int J of Soc Robotics 6, 417–427 (2014). https://doi.org/10.1007/s12369-014-0244-0

[35] Kshirsagar, S. (2002). A multilayer personality models. *Proceedings of the 2nd International Symposium on Smart Graphics - SMARTGRAPH '02*. https://doi.org/10.1145/569005.569021

[36] Ladwig, R. C., & Ferstl, E. C. (2018). What's in a name? *Proceedings of the 4th Conference on Gender & IT - GenderIT '18*. https://doi.org/10.1145/3196839.3196851

[37] Landes, J. B. (1984). WOMEN AND THE PUBLIC SPHERE: A Modern Perspective. *Social Analysis: The International Journal of Social and Cultural Practice*, *15*, 20–31. http://www.jstor.org/stable/23169275

[38] Lawson, E.D. (1987). Personal names and naming: An annotated bibliography. New York: Greenwood.

[39] Lie M (1995) Technology and masculinity: the case of the computer. Eur J Women’s Stud 2(3):379–394

[40] Martínez, C. A., & Leongómez, J. D. (2019). Dehumanization affects agency attributions, retributive justice judgments, and resocialization attitudes towards former perpetrators in post-conflict. https://doi.org/10.31234/osf.io/w5tka

[41] Megan. (2022, April 5). *Children and consent: What you need to know*. The Innocent Lives Foundation. Retrieved April 26, 2023, from https://www.innocentlivesfoundation.org/children-and-consent-what-you-need-to-know/.

[42] Meyers, D. (1987). Personal autonomy and the paradox of feminine socialization. Journal of Philosophy,

[43] Nomura, T. (2017). Robots and Gender. Retrieved December 13, 2020, from https://journals.sagepub.com/doi/full/10.1089/gg.2016.29002.nom

[44] Piper, M., & Veltman, A. (2014). *Autonomy, oppression, and gender*. Oxford University Press.

[45] Proudfoot, D., & Kay, A. C. (2023). Communal expectations conflict with autonomy motives: The western drive for autonomy shapes women’s negative responses to positive gender stereotypes. Journal of Personality and Social Psychology, 124(1), 1–21. https://doi.org/10.1037/pspa0000311.supp (Supplemental)

[46] R Core Team (2018). R: A Language and Environment for Statistical Computing. Vienna: R Core Team.

[47] Rau, P.-L. P., Li, Y., & Liu, J. (2013). Effects of a social robot's autonomy and group orientation on human decision-making. *Advances in Human-Computer Interaction, 2013*, 1–13. https://doi.org/10.1155/2013/263721

[48] Richardson, K. (2016). The asymmetrical 'relationship'. *ACM SIGCAS Computers and Society*, *45*(3), 290–293. https://doi.org/10.1145/2874239.2874281

[49] Sakai, T., & Nagai, T. (2022). Explainable autonomous robots: A survey and perspective. *Advanced Robotics*, *36*(5-6), 219–238. https://doi.org/10.1080/01691864.2022.2029720

[50] Singh, R. (1998). *Gender autonomy in Western Europe*. *Springer Link*. Palgrave Macmillan, 27-28 UK. Retrieved from https://link.springer.com/book/10.1057/9780230379022.

[51] Strait M, Briggs P, Scheutz M (2015) Gender, more so than age, modulates positive perceptions of language-based human-robot interactions. In: Salem

[52] T.C. İçişleri Bakanlığı Bilgi Teknolojileri Genel Müdürlüğü. (2020). *Türkiye'de geçen Yıl 1 Milyon 180 bin 840 Bebek Dünyaya Geldi*. T.C. İçişleri Bakanlığı. https://www.icisleri.gov.tr/turkiyede-gecen-yil-1-milyon-180-bin-840-bebek-dunyaya-geldi

[53] Tay, B., Jung, Y., and Park, T. (2014). When stereotypes meet robots: the double-edge sword of robot gender and personality in human-robot interaction. *Comput. Human Behav.* 38, 75–84. doi: 10.1016/j.chb.2014.05.014

[54] Tay, B., Park, T., Jung, Y., Tan, Y., & Wong, A. (2013, July 21). When Stereotypes Meet Robots: The Effect of Gender Stereotypes on People's Acceptance of a Security Robot.Retrieved December 15, 2020, from https://link.springer.com/chapter/10.1007/978-3-642-39360-0_29

[55] Trope, Y., and Liberman, N. (2010). Construal level theory of psychological distance. Psychological Review 117: 440. doi:10.1037/a0018963

[56] Trovato, G., Lucho, C., & Paredes, R. (2018). She's Electric—The Influence of Body Proportions on Perceived Gender of Robots across Cultures. *Robotics*, *7*(3). doi:10.3390/robotics7030050

[57] Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, *57*(4), 660–671. https://doi.org/10.1037/0022-3514.57.4.660

[58] Zeldin, A. L., Britner, S. L., and Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. J. Res. Sci. Teach. 45, 1036–1058. doi: 10.1002/tea.20195

[59] Zwebner, Y., Sellier, A.-L., Rosenfeld, N., Goldenberg, J., & Mayo, R. (2017). We look like our names: The manifestation of name stereotypes in facial appearance. *Journal of Personality and Social Psychology*, *112*(4), 527–554. https://doi.org/10.1037/pspa0000076