

Kandidatnummer: 10091

Can your uncle be a wedding planner? Proscriptive stereotypes and their effect on attitudes about people's compatability with occupational roles

PSY2900 - Bacheloroppgave i Psykologi
NTNU Trondheim

Bachelor's thesis in Psychology

Supervisor: Kim, Jonathan

May 2022

Kandidatnummer: 10091

**Can your uncle be a wedding planner?
Proscriptive stereotypes and their
effect on attitudes about people's
compatibility with occupational roles**

PSY2900 - Bacheloroppgave i Psykologi
NTNU Trondheim

Bachelor's thesis in Psychology
Supervisor: Kim, Jonathan
May 2022

Norwegian University of Science and Technology

This thesis marks the end of my three years as a bachelor student at NTNU. The years have sadly been affected by unfortunate circumstances, but it has still been a good experience through and through. I am very grateful for the opportunity to be part of this project, which was about one of the topics I found to be the most confusing, but also one of the most interesting, of all the topics I've been introduced to during these three years.

The experimental elements used in this thesis was chosen by the project leader, Jonathan Kim, as well as the whole part of setting up the experiment. This includes choosing the tests we used, (yes-no forced choice task, ASI, IAT) and coding the different experiments on PsyToolkit.

The responsibility of the student group was to recruit participants to the study, and the rest of the preparation was done by the project leader.

Some of the sources used in this thesis was from the reading list we received at the start of the semester. Throughout the whole project we as a group have been receiving a lot of guidance on the general topic, how to prepare the data, how to perform the analyses, how to write the thesis, and other issues we could possibly have on the project. I would personally like to thank the project leader for all this help which would be very hard to do without. I would also like to thank my friend for reading through my text and finding what is hopefully the most of my many clumsy phrasings and spelling errors. This is the extent of help that I have used in the creation of this thesis, and I declare that this is a work of my own.

Abstract

In this study we examined the participant's attitudes on people's compatibility with gender stereotyped occupations and how this could be affected by different sources of stereotype

information. This was done through a forced-choice task. The results implied that participants responded significantly slower when presented with family nouns compared to names, but there was no significant difference between the female gendered family items or the male gendered family items. A previously reported male bias on neutral items was no longer prevalent which can imply that the language amendments have successfully phased out the effect of grammatical gendering on generic nouns. An interesting finding was that the participants responded consistently more positively and in shorter response time when presented with a neutral stereotyped role regardless of the item gender they compared it with. This can imply that the participants used proscriptive stereotypes more than prescriptive when deciding whether the item, whether family member or name, could fit the occupational role presented.

Gender roles and gender stereotypes are a social phenomenon which can vary in form and strength from culture to culture. It is also something that evolves and changes within a culture as well. While stereotypes are a cultural product it is also something that is constructed and present in a person itself. It is in many ways an internalization of attitudes present in the social world. By taking information from the social world the person creates their own understanding and beliefs about groups of people and how they ought to be. This is learned by communication and interactions with others and through the mediums of communications. It is a convenient way to make a quick assessment of people by just assuming they adhere to the stereotypes you already have about the group in question. Understandably defaulting to this is very reductive and will often lead to very mistaken assumptions and expectations of groups. Not feeling like people belong in a setting can naturally lead to them not wanting to be part of that setting. Results from a qualitative study on why a lot of women, especially women with minority backgrounds, leave STEM studies implied that a sense of belonging was one of the main factors leading to them leaving the field (Rainey et al, 2018). Since stereotypes are closely related to the perceived typicality of a group seeing few people of a social group in an occupation might make people perceive that this occupation is something this group isn't fit for. This might lead to people of this group not wanting to pursue these options. The prescriptive nature of stereotypes can to some degree be a self-fulfilling prophecy by pushing out what is not considered to be incongruent with the characteristics of the occupation.

Challenging these notions can be a very difficult, but equally as important, issue that needs to be addressed through evidence-based policy.

What is considered to be stereotypical of a group is what is thought to be typical of the group in question (Prentice & Carranza, 2002). Stereotypes are prescriptive and proscriptive in nature and deviating from these rules can have consequences. The consequences can range from being sanctioned to being rewarded depending on what kind of stereotypes and expectations are being broken. Prentice and Carranza separated gender stereotypes expectations into four different categories where the different categories were based on two separate dimensions. The first dimension was prescriptive stereotypes which were traits that were considered positive. The expectations of a person to have this trait could vary depending on which gender they were. A gender could have an intensified prescription to this trait which meant it was an expectation of a person of this gender to possess this trait and failing to do so would be looked down upon. A trait that was an intensified prescription for one gender was often a relaxed prescription for the other gender. A relaxed prescription is a trait which is considered positive, but it is considered less important for people of a gender to have. It can be considered relaxed not just because it is considered more important for the other gender, but because it was considered less important when the gender was specified than when people were asked to judge how important the trait was for a person to have regardless of gender. The genderless person often had different expectations than both the gendered ones which creates a baseline to be used to determine if a trait is intensified or relaxed for a gender.

Language is one of many factors which have been linked to understanding and creating meaning of the world. Not only is language one of the main ways we gather and exchange information, but it is also an important tool for the coding process of information (Loftus & Palmer, 1974). Naturally this would also link language to the generation and maintenance of social stereotypes which in itself is just a set of “meanings” about a specific group of people. Language might be the primary tool we use to understand and make meaning of the world and is therefore a very important factor on any issue regarding social stereotypes. Language is also an expression of culture itself, and differences in culture can be expressed in the linguistic differences. Because of the connection between our language and our own generation of meaning, language might be one of the most important areas of study regarding decision making and social cognition. This importance is not unnoticed and many policies considering

the use of language and its effect on both generating and maintaining stereotypes have been tried out in multiple countries in multiple different ways (Gygax et al, 2008).

In countries where the primary language is gendered, such as France and Spain, it has been suggested to combat a male bias to use strategies of feminization of the language (Gabriel et al, 2018). Since the generic plural is grammatically male gendered this can result in highlighting the women and being more inclusive, but in non-gender marked language this can have an opposite effect by making women seem as something different and separate to the norm. In these languages it is suggested to use strategies that are keeping with the neutral gendering of the language. Norway has historically had a gendered language, but has since the 1970s had a neutralization strategy to address a male bias (Gabriel et al, 2018). Today that leaves Norwegian as a language that in some sense is between the gendered and the non-gendered languages. A study conducted at NTNU in 2008 measured how participants used language cues and stereotype cues to decide if two succeeding sentences made sense with each other (Gygax et al, 2008). The results from this study showed that stereotypes had a proportionally higher impact on the participants decision making, but there was still some male bias on neutral stereotyped roles which was explained as an “aftertaste of gendering of the language”. These results were later supported in a latter experiment which challenges this hypothesis (Gabriel et al, 2017).

The primary research question this thesis set out to examine was; does activation of family roles affect perceptions of whether a person can be part of a gender stereotyped occupation, and if so, is this effect related to stereotypical gender roles in the family. The first two hypothesis is directly connected to this research question while the third and last hypothesis is connected to the secondary research question; has language policy had a measurable effect on diminishing a male bias from semantic information. Other topics of interest that will be brought up in the discussion part is how perceived gender stereotypes can affect hiring and what people perceive as viable career options, and how visible representation can affect this perception.

Hypothesis 1: There will be a significant difference in average response time between the trials where the participants are presented with a name item and the trials where participants are presented with a family role item.

This hypothesis is based on assumptions made from schematic activation and implicit biases (Cacciari & Padovani, 2007). Schematic activation is premised on that when we hear or read a word, attributes connected to this word are activated. These attributes are connected to the stereotypical beliefs regarding this word. This relates to the hypothesis since the assumption is that the family nouns will activate the setting and context the family words are associated with and inhibit the other possible interpretations that are “less likely”. The activation-selection model is also very relevant. This model assumes that we decide the meaning of ambiguous words by using primers or contextual clues. When a word is used a lot of attributes connected to the word are activated and if these attributes relate to a succeeding word it is furtherly weighted (Gorfein et al, 2007). The interpretation that is given the most weight is the one that is chosen as being most likely correct. If the family role is activated and perceived as entirely separated, and often in opposition of the occupation, then this could mean that the activation of this role will force the participant to use higher effort to consider these two roles as compatible because the meaning opposed has been activated. In our experiment that would result in a significant difference in the average response time between the trials using the different item types; names or family roles.

Hypothesis 2: the difference will be bigger for feminine family items due to a stronger family role stereotype regarding women.

This builds further on the previous hypothesis and assumes that stereotypically there is a bigger emphasis on women’s role in the family, and an assumption that women should value family over career. While Norway is considered one of the most gender egalitarian countries there is still a prevalent stereotype about parenthood that emphasizes the mother’s role as the main caregiver. The father's role in the family has historically been the “breadwinner” while the mother has been the one with the main responsibilities regarding raising the children and taking care of the home. While policies in Scandinavian countries have had some effect on closing this “childcare gap” it is still not fully egalitarian (Ellingsæter & Gulbrandsen, 2007). While working dual-earning parents are very common in Norway in large part due to a supportive system through welfare care, the demand for these services often exceeds the number of spaces. As Ellingsæter and Gulbrandsen (2007) points out, a change in the number

of mothers in the workforce has been a driving factor for attitude and policy changes. Still the emphasis on the mother as the main provider of childcare is prevalent.

Previous studies on hiring practices have used this assumption as an indicator to why women are less considered for positions which demand higher dedication and time in the job (Shaw & Hoerber, 2003). Research on discourses in sports management has implied similar issues where the agenda setting of organizations are less accessible for employees with family role responsibilities. This leads to them getting systematically excluded from these important settings for social networking necessary to be considered for higher positions in organizations (Shaw & Frisby, 2006). The expectations for the family responsibilities are heavily stereotyped as feminine. Sarah Hampson even referred to this attitude as seeing motherhood as an “antithesis of an ideal worker” (Hampson, 2018).

Men might have a more relaxed prescription to these expectations which allows them to be perceived as more fitting and capable for these positions. While not all the occupations in this experimentation are management position or other occupation that is perceived to expect much effort and time, and the feminine stereotyped occupations are by nature supposed to be occupations that are considered more compatible for women, the effect should at least be visible on the masculine gender stereotyped roles.

Hypothesis 3: There won't be any significant difference in choices or response time between female and male items on neutral roles.

This hypothesis is drawn directly from the study from 2008 previously mentioned (Gygax et al, 2008). As mentioned, this hypothesis was directly challenged by a latter experiment (Gabriel et al, 2017), and it was speculated that the fixed time might make participants rely more on semantic information than in a self-paced task. While our experiment is time limited and supposed to be performed as quickly as possible, it does not give participants as much cognitive load as the two experiments previously mentioned. Our experiment only uses two nouns to decide if the first noun could be part of the group that is the second noun. Their experiment on the other was to decide if two succeeding sentences made sense in relation to each other.

The results from both experiments showed that gender stereotypically neutral roles were responded faster to on average when the primer was male. The effect was not large, but it was

significant and considerable. This male bias was explained as a result of the neutral plural form of the occupation historically being masculine. The paper mentioned as an implication for future studies that this effect could possibly diminish over time and expected that within 20 years the small male bias explained by the grammatical form of the words would not appear in similar studies. Our study is conducted nearly 15 years after, and this might be enough time to see the effect that was hypothesized.

Method

Ethical

Due to the data management elements of the experiment the project needed approval from NSD before conducting the experiments. The project was approved by NSD before it started preparing for the collection of data. The NSD reference number is 523093. This study was keeping in line with the NESH ethical guidelines on research in social sciences and the humanities (NESH, 2021)

Participants

There were in total, before the deselection process, 44 participants that completed the experiment. 33 of the participants were women (75 %), while the remaining 11 were men (25 %) The mean age of the participants was 24 years with a range from 18 to 68. Most of the participants were students, $N = 36$ (82 %).

The data collection was done through an online survey made in and hosted on PsyToolkit. It was decided early on to conduct the experiment this way instead of an in-lab experiment on campus. This decision was heavily influenced by the uncertainty of future complications due to coronavirus restrictions. The decision was also made on the grounds that previous studies on this specific paradigm has tested the quality of conducting studies through PsyToolkit compared to results in a lab and showed no significant loss of quality (Kim et al, 2019). The recruitment of participants was our main responsibility as the student group in the project. Different methods such as proximity recruitment of fellow students and other convenient potential participants was encouraged and recommended.

Materials and Research Design

The participants responded to four different experimental elements in this study. The study consisted of a questionnaire, a two-alternatives forced choice task, an IAT, and an ASI. The questionnaire was a set of demographic questions regarding the participant such as their age, if they were students or not, if they were native Norwegian speakers, and their preferred handedness.

Yes-No Choice Task

The main body of the study was the two-alternatives forced choice task. This experiment examined participants' stereotypical beliefs about gender roles. In particular it examined gender stereotypical beliefs about occupational roles. There were two different versions of this task that was used in this experiment. The difference between these two versions was the order in which the two blocks of the experiment were presented. The two blocks of the experiment were the "Name Block" and the "Family Block" which had different items used. All participants completed both blocks, but the order of the blocks was different between the two versions. The participants were presented with either a name (e.g *David*) or a family role (e.g *Mødre*) and an occupational role in plural form (e.g *Fyrvoktere*). The task required the participant to respond as quickly as possible whether the name or family member could also be part of the group presented. The role group (occupational roles) consisted of three categories of items; masculine stereotyped, feminine stereotyped, and occupations without a specific gender stereotype attached to it. While this would always be the case since the occupations might be heavily gender stereotype associated, but would still be possible for any person to be, the participant couldn't respond "yes" every trial since filler items with correct yes-no answers were added in the data and would randomly appear throughout the test. These filler items were roles such as "*Konger*" or "*Dronninger*" which would force the participant to think about the possibility for the participant to be part of this group and not just default to answering yes.

The experiment has 36 different roles (occupations), 12 masculine stereotyped, 12 neutral, and 12 feminine stereotyped. These occupations were pulled from a previous study on

perceived stereotypicality in occupational roles in different languages, one of the languages being Norwegian (Misersky et al 2014). A similar process was done to get the names which would be considered very clearly masculine and feminine and taken from an unpublished paper. The full list of names, occupations and family roles can be found in the appendix.

By measuring if there was any systematic difference in participants choices (yes or no) on occupation trials we could measure an explicit bias on gender roles. The response time was also measured which would give another dimension of stereotypical beliefs to examine. Systematic differences in response time could be considered a measure of implicit attitudes on gender stereotypes.

Implicit Association Test

The implicit association test is very similar in functioning as the yes-no choice task. In this test the participant is presented with two different sets of stimuli: one adjective group and one noun group. The adjective group can be any trait you want to test as an association. The most basic version of IAT tests uses positive adjectives (nice, great, wonderful) versus negative adjectives (mean, horrible, bad). In this IAT test the adjective group is agentic words vs dependent words, and the items were taken from a previous study on attitudes towards female authority (Rudman and Kilianski, 2000). While the noun group is men vs women. The task involves sorting the different word groups as quickly as possible by indicating if it belongs on the left side or the right side. The men and the women will be sorted opposite of each other, as will the agentic and the dependent words. After a while the noun group changes places, while the adjective group stays. The test measures if there is a systematic response time difference between these two blocks and uses that as an indicator for an implicit association between the adjectives and the nouns. If the women group would be faster and more accurately sorted when it is paired with dependent words then this could imply that there is an implicit association connecting these groups of words (Greenwald et al, 1998).

Ambivalent Sexism Inventory

Ambivalent Sexism Inventory is a questionnaire where participants respond using a five-point Likert scale (Glick & Fiske, 1996). We decided on using the short form ASI in this

experiment which consists of 12 questions. The longer form reverses the measurements, but we decided the short form is adequately accurate for an experiment of this sort. The 12 questions can be divided in two; 6 questions on hostile sexism and 6 questions on benevolent sexism. The questions were translated to Norwegian in a previous study, and we used these translations in our experiment (Bendixen et al, 2014)

Procedure

Participants were sent access to the test through a link online. The test was hosted on PsyToolkit. Everything presented in the study was in Norwegian. The first thing presented to the participants was information on how the study was going to be conducted and their rights as participants. The participants were asked to wear glasses if they needed it, and how much time it would take to complete the experiment. They were informed on their rights to withdraw at any stage of the experiment, their rights to ask any question about the study that they may wonder about during the experiment, their rights to a summary of the findings of the study, and the right to have their data withdrawn from the experiment up to a month after participating. The contact information of the project leader was presented so they knew who and where to contact if they needed to.

The first part of the experimental elements the participants were presented with was the questionnaire. After answering the demographic questions already mentioned the participants were informed about how the forced choice tasks were to be conducted. They were instructed to put their fingers on the “e” and the “i” on the keyboard with their thumbs resting on the spacebar. If the participant received Version A of the experiment, they were instructed to use the “e” button to answer no and the “i” button to answer yes if they thought it was possible for a “*name*” to be part of an occupation or social role. Version B which had the family block first received the similar information, but was told that they were to judge if a *family role* could be part of an occupational role. When they were ready they could start the trials. The first five trials were training trials so the participant could get a practical test to understand the concept fully before the experiment began. These trials were discarded and not used in any analysis. After completing the experimental 360 trials the participant was finished with the first block. After completion they were told to now answer yes or no on questions regarding the other

block they had not done. After completing the 5 training trials and the 360 trials in the second block the participants had completed the main body of the experiment.

The next experimental part was the IAT. The participants were informed to use the same keys as in the previous experiment. Now the “e” key was to sort the item presented to the category on the left, and the “i” key to sort the item to the category on the right. Wrong answers would be signaled with a red cross, and correct answers with a green checkmark. After completing the IAT the participants were told they’ve completed the experiment and just had a survey left. This was the ASI. The participants were presented with the questions and told to answer using a five point scale which went from “strongly disagree” to “strongly agree”.

Data preparation

In the data preparation there was both an item-by participant deselection process and a by-participant deselection process. After removing all choice trials with responses either faster than 300ms and trials where the participant failed to answer within the 5000 ms window the data was reduced by 9.7 %. The faster responses were removed because it takes longer time to read, comprehend, and then respond to tasks of this nature (Greenwald et al, 2003). On the opposite side of the spectrum, answers over 5000ms is a good limit for tasks like these where the participant is given more than enough time to evaluate the question and give a response. Answers over this limit either indicate that the participant isn’t paying enough attention or that the participant is overthinking the task. Participants who had more than 10% of their trials removed in this process were excluded from the experiment. There were 8 participants in total excluded from the dataset due to exceeding this limit. The next step in the deselection process was removing non-students. This consisted of 7 participants in total. This was our main by-participant deselection step.

To keep participants from just responding “yes” to the yes-no task filler items which had clear correct responses were added randomly throughout the experiment. Since these tasks had a correct answer, it allowed us to measure an error rate as well. By measuring how many of these responses the participants answered incorrectly compared to how many responses they answered correctly we could calculate their error rate. A quite relaxed rate of 50 % error was allowed, which meant that participants who had more error than this were excluded from the

data. There were 2 participants who were excluded in this step of the deselection process. The filler items were not used in any other analysis than this and were separated from the rest of the dataset.

The final step of the by-participant deselection process was removing a participant who was an outlier on time it took to finish the experiment. While all the other participants' completion time ranged from 24 minutes to 84 minutes this participant had a total running time of 3156 minutes. While this participant was not disqualified on error rates or response times, which as mentioned indicates that the participant was paying enough attention to participate, the total running time suggests that the participant left the test running for a whole day between sections. Because of this they did not participate on equal terms as the other participants and was therefore removed from the dataset.

There were two different analyses in the study. An analysis of choice and an analysis of response time. The choice analysis examined differences in the mean choices (yes or no). The response time analysis examined differences in the mean response time before the participant answered yes. Therefore, the trials where the participants answered no were excluded from this analysis.

Participants in the experiment

After the data preparation was completed, there were 25 participants left that were used in the experiments and analyses. The gender ratio was relatively unchanged, 19 of the participants were women (76 %), and the remaining 6 men (24 %). The age of the participants was $M = 22.48$, with a range from 20 to 27 years. It was as close to an even split as possible regarding participants who received Version A (names first), $n = 13$ (52%), and participants who received Version B, $n = 12$ (48%).

Statistical Analyses

The analyses of Choices and Response Time was done using two ANCOVAs on their respective datasets. The ANCOVA analysis was chosen since it can adjust for the effect of covariates. The two covariates chosen to adjust for were the number of trials and the results

from the ASI. The number of trials was important to account for since the participants could possibly get better at performing the task during the experiment or inversely get tired and this can affect the results. For the same reasons it was important to control for variance created by the different versions of the experiment. Participants who started with names would have a lot more experience by the time they started the family block, and this could have an effect. The ASI covariate was accounted for to see if there was a significant connection between the results and this variable, and to control for the effect this could have on the results.

Results

Choice

The results from the study found a significant Main Effect of Role Gender $F(2,10773) = 103.96, p > .001$, partial $\eta^2 = .019$. The participants responded most positively to the Neutral Roles $M = 99\%$, 95% CI [98.4%, 99.7%], followed by the Masculine Roles $M = 95.8\%$, 95% CI [95.2%, 96.5%] Mean difference; $\Delta M = 3.2\%$, 95% CI [2.1%, 4.3%], and the least positively being Feminine Roles $M = 92.4\%$, 95% CI [91.8%, 93.1%], Mean difference from Neutral; $\Delta M = 6.6\%$, 95% CI [5.5%, 7.7%]. This was the largest Main Effect observed on participants' choices.

There was a significant Main Effect of the Item Gender variable $F(1,10773) = 15.61, p < .001$, partial $\eta^2 = .001$. Participants responded significantly more positively to Female $M = 96.5\%$, CI [96%, 97%] than to Male $M = 95\%$ CI [94.5%, 95.5%]; the mean difference; $\Delta M = 1.5\%$ 95% CI [0.7%, 2.2%].

The Main Effect of Item Type was not significant $F(1,10773) = 1.17, p = .279$. The name category had a slightly more positive response rate $M = 96.0\%$, 95% CI [95.5%, 96.5%], than the family items $M = 95.6\%$, 95% CI [95.1%, 96.1%], but this difference was non-significant $\Delta M = 0.4\%$, 95% CI [-0.3%, 1.1%]

There was a significant Interaction Effect On Item Gender and Role Gender $F(2,10773) = 155.56, p < .001$, partial $\eta^2 = .028$. The results (Table 1) indicated that for female gender items the participants responded significantly more positively to neutral roles than to masculine

roles, $\Delta M = 6.5\%$, 95% CI[4.7%, 8.3%]. They also responded significantly more positively to female roles than to masculine roles, $\Delta M = 4.7\%$, 95 % CI[2.9%, 6.5%]. While for masculine gender items the participants' answers scored identical in mean with similar confidence intervals on masculine and neutral roles. The difference between the neutral/masculine roles to the feminine was significant $\Delta M = 11.4\%$, 95% CI [9.6%, 13.2%].

Table 1

Interaction effect of Item Gender and Role Gender on choices				
Item Gender	Role Gender	<i>M</i>	Lower Interval	Upper Interval
Female	Masculine	92.8%	91.9%	93.7%
	Feminine	97.4%	96.5%	98.3%
	Neutral	99.3%	98.4%	100%
Male	Masculine	98.8%	97.9%	99.7%
	Feminine	87.4%	86.5%	88.3%
	Neutral	98.8%	97.9%	99.7%

The other interaction effects were non-significant; interaction between Item Type and Item Gender $F(1,10773) = 0.50$, $p = .480$, interaction between Role gender and Item Type $F(2,10774) = 0.25$, $p = .777$, interaction between Item Type, Item Gender, and Role Gender $F(2,10774) = 0.84$, $p = .431$.

The covariates; Trial $F(1,10773) = 13.20$, $p < .001$, partial $\eta^2 = .001$, and ASI $F(1,10773) = 25.99$, $p < .001$, partial $\eta^2 = .002$, Version $F(1,10773) = 239.08$, $p < .001$, partial $\eta^2 = .022$, all had an significant effect on the choice making of the participants.

Response Time

Results from the Response Time analysis showed a Main Effect of Item Type $F(1,10316) = 219.83$, $p < .001$, partial $\eta^2 = .021$. The participants took significantly longer time to respond to the family items $M = 1227.22$, 95% CI[1210.57, 1243.87] than to the name items $M = 1052.42$, 95% CI[1035.74, 1069.11]; $\Delta M = 174.80$, 95% CI[151.20, 198.39]

There was a significant Main Effect of Role Gender $F(2,10316) = 12.36, p < .001$, partial $\eta^2 = .002$. The results showed it took significantly longer time for the participants to respond to feminine roles $M = 1169.00$, 95% CI [1148.28, 1189.72], than neutral roles $M = 1102.82$, 95% CI [1082.73, 1122.91], mean difference; $\Delta M = 66.18$, 95% CI [30.93, 101.44]. There was no significant difference in response time between feminine roles and masculine roles $\Delta M = 21.36$, CI 95% [-14.13, 56.85].

The results showed no significant difference in response time on the Item Gender variable $F(1,10316) = 1.81, p = .179$. The male items had a slightly higher average response time, $M = 1147.95$, 95% CI [1131.41, 1164.49] than the female items $M = 1131.95$, 95% CI [1115.51, 1148.39] but the difference was non-significant; $\Delta M = 16.00$, 95% CI [-7.32, 39.32].

There was a significant Interaction Effect on response time from the Item Gender and Role Gender interaction $F(2,10316) = 3.56, p = .28$, partial $\eta^2 = .001$. See table 2 for the results. When there was a female item there were no significant differences in the participants' response times. The differences from the role with the highest average response time (masculine); *masculine x feminine* $\Delta M = 6.35$, 95% CI [-50.96, 63.66], *masculine x neutral* $\Delta M = 40.42$, 95% CI [-16.76, 97.6].

While for the male items there was a significant difference when on response time comparing the role with the highest average response time (masculine) to the lowest average response time (neutral), *feminine x neutral* $\Delta M = 107.78$, 95% CI [49.84,165.73], but not when comparing with the masculine role, *feminine x masculine* $\Delta M = 53.29$, 95% CI [-4.48, 111.07]

Table 2

Interaction effect of Item Gender and Role Gender on RT				
Item Gender	Role Gender	<i>M</i>	Lower Interval	Upper Interval
Female	Masculine	1147.54	1118.48	1176.60
	Feminine	1141.19	1112.94	1169.44
	Neutral	1107.12	1079.00	1135.24
Male	Masculine	1148.35	1120.37	1176.33
	Feminine	1201.64	1171.85	1231.44
	Neutral	1093.86	1065.71	1122.01

The interaction between Item Type, Item Gender, and Role Gender was non-significant $F(2,10317) = 1.29, p = .275$. As was the interaction between Item Type and Item Gender $F(1,10317) = 0.14, p = .708$, and the interaction between Item Type and Role Gender $F(2,10317) = 0.09, p = .915$.

The Trial covariate showed a significant effect on the variance of response time $F(1,10317) = 500.97, p < .001$, partial $\eta^2 = .046$. The ASI covariate was also significant $F(1,10317) = 196.11, p < .001$, partial $\eta^2 = .019$. As was the Version covariate $F(1,10317) = 201.86, p < .001$, partial $\eta^2 = .019$.

Discussion

The first hypothesis was that there will be a higher response time for family items than for name items. This hypothesis was fully supported by the results. This result can be interpreted as an argument that family roles are seen as something completely separated from work life. By being presented with a role such as a mother or a father the participant is guided to a family setting. A name is more neutral in that sense and wouldn't need an extra step to reconsider the activation of meaning created from the family member role.

The results can also be explained by it being cognitively harder for the participants to decide if they think people in a group can also be part of another group than if a single person can be part of a group. It could be necessary for a future study to examine if this could have an effect on the results to determine if these results support the arguments presented.

It is worth noting that it did not have any significant influence on the degree participants answered yes. This could imply that it wasn't detrimental to the ability to see that these two roles could be compatible on an explicit level.

The second hypothesis was that the family items that were female gendered would have a higher average on response time than the male gendered ones. This was not supported by the results as the interaction effect of item type and item gender was found non-significant. It was actually the male items that had the highest measured response time, but it was non-significant. The mean difference between the male gendered family items and the female

gendered family items was; $\Delta M = 20.74$, 95% CI[-25.86, 67.35]. The hypothesis also implied that it was the mother role which would have the biggest impact on the response time since this is the role considered to be most family orientated. This was not the case in the results as the mother item was not a statistical outlier of any sort. This result could be explained that the effect of the family role is already strong, and the different roles do not significantly differ when the effect is already “activated”. Another explanation is that the model of parenting that is perceived as most typical in Norway could be the dual-earner model where both parents earn income in work. The studies that this hypothesis drew from were mostly in settings without the support system provided to parents in Norway which enables mothers to go to work. The percentage of preschool kids in publicly funded day care has increased from 2% to 67% from 1963 to 2006 (Ellingsæter & Gulbrandsen, 2007). These changes have greatly impacted the opportunity for families to have both parents in the workforce without this affecting the care for the children. There was already a reported change in attitude reported back in 2004, and while there was still a considerable amount of people who responded that they thought the mother should stay home with the child this number was decreasing and could possibly be a lot lower today. The small difference in average mean could be explained by a slight female bias stemming from the participants in the study having a skewed gender balance, but this difference is very small and not significant.

The third hypothesis was that there will no longer be a male bias caused by the grammatical gendering of neutral nouns. The results from the analysis showed no significant difference in response time. While the measured mean is slightly larger for the *female x neutral* than the *male x neutral* the confidence intervals are overlapping, and this indicates that the results are questionable. There was no significant difference on participants choices either, and the slight difference in measured mean was now conversely “in favor” of the feminine instead of the masculine (table for Choice item gender x role gender). The results were in support of the hypothesis but can’t fully contest the findings from the previous study as the small differences in experimental design might have slightly different effects than the ones found in this study. Future studies on this topic are required, but the results can still be said to offer some support to the hypothesis of Gabriel et al (2008).

An interesting finding was that the neutral roles both had a higher percentage of yes answers and had the lowest average response time. This makes sense considering that these occupations are supposed to be equally compatible with both masculine and feminine and

therefore not affect either. Yet if stereotypes would make people consider something stereotypical even faster than if it was just compatible the results should be that male items would be associated with male stereotyped roles faster than neutral roles, and female items would also be associated with female stereotyped roles faster than the neutral roles. The opposite of this was true, and the neutral items had consistently the lowest response time and highest positive response rate. This could be because the roles used for the neutral category by nature seem more accessible than roles that are more gender stereotyped. These results can also imply that the proscriptive effect of stereotypes is stronger than the prescriptive effect. If the prescriptive effect was stronger the gender stereotyped roles would be scoring the most favorably when matched with the congruent gender. It seems like the decision making is more affected by thoughts that “this gender can’t do this” more than “this is something this gender excels at”. If this is the case then the results would imply that it is more acceptable for a woman to be masculine than for a man to be feminine. This effect would be strongest explicitly since the difference on positive choices in the Item Gender and Role Gender interaction between the *female x masculine* and the *male x feminine* was significant $\Delta M = 5.4\%$, 95% CI[3.6%, 7.2%] are significant and considerable, while the same interaction does not have as strong effect on the response times and is inconclusive due to overlapping confidence intervals; $\Delta M = 54.10$, 95% CI[-4.76, 112.96]. This could imply that it is harder for men to step into a feminine role than it would be for a woman to step into a masculine role. This could reflect that a lot of focus on gender equality is focused on encouraging women to go into male dominated spaces, but there is less of an emphasis on encouraging men to apply to more female dominated occupations (i.e Nursing, Meadus & Twomey, 2007). The negative implications that can come from this method of making the playing field even is that it puts the expectations on women to change themselves and act more “masculine” instead of a mutual acceptance for people of both female and male gender to inhibit traits that can be outside of what is considered stereotypical today. Relating this to the childcare responsibilities between mothers and fathers, some of the efforts that have seemingly the best effect on promoting fathers to take more part of the childcare are policies which directly address men (Brendth & Kvande, 2009). The father’s quota of paternity leave which is non-transferable encourages the father to take part in the raising of children while the cash-for-care policy does not directly address the differences.

While some words might be neutral grammatically in gendering the occupation could still contain a gendered noun which can have an effect (e.g in Finnish; Gabriel et al,2018).

Examples of this in Norwegian could be “brannmann”, “barnehagetante”, and “politimann” which all were to some extent used neutrally. Recent changes in language policy is trying to change these words by switching out the gendered nouns with neutral nouns. None of the occupational roles used in this experiment had this issue which is good since it eliminates a source of “noise” that would make it harder to measure the actual strength of stereotype information. There is a lot of criticism of the usage of these words as it is believed to be linked to creating a male bias (Gabriel et al,2018). An example of similar linguistic issues that is being discouraged is the use of male pronouns as generic. Like using “han” when the gender is not known. The created “hen” is an example of measures that is being used to make the language more neutral. While changes in language won’t change overnight and can to some extent only be nudged and encouraged. These changes to the language can possibly be noticeable and research on the effect of these strategies of neutralization can be important. There has been a reported connection between perceived acceptance and compatibility of groups in an occupation and if the group considers this as an option for a career. In France which used more grammatical cues in their activation of stereotypes a text which is supposed to target both genders equally that used both forms of the gender were perceived as more compatible with women than texts which only used the masculine plural (Gygax et al, 2012) While Norwegian has a completely different source of schematic activation because of the grammatical differences, these findings can still imply that the perceived stereotypicality can have an impact on if people find social groups compatible with their own identity. This has implications for policy making to even out occupations where a gender is underrepresented and being treated worse for not fitting in with what is considered the stereotype of the occupation. Our results can imply that men in women dominated fields are the ones who would feel the least at place due to the stronger implicit and explicit bias against men in feminine roles.

In the last few years representation in the media has become a very visible issue, and there is an ongoing debate regarding more positive depictions of different races, gender, and sexualities. It is theorized that the perception of the number of a group in specific occupations can have an effect on the perceived stereotypicality (Gabriel et al, 2008). This is where the issue of lack of representation can lead to a group appearing more homogenous than what it actually is, or what it ought to be. While mediums such as TV, movies, news and other platforms have been a part of society for a long time, social media platforms are still a relatively new phenomenon. Social media also differentiates itself from the previously

mentioned mediums by being “non-centralized”. While television is limited by broadcasters deciding what to televise, social media and streaming is a lot more interactive on the user end. People more actively choose out the content they want to see, and advertising is a lot more personalized creating what has been referred to as “bubbles”. Since media could potentially have a big impact on shifting perceived stereotypicality of gender and other groupings of people (e.g Karsay & Desiree, 2019, Ward & Grower, 2020), this could be a very interesting field of topic for a cross disciplinary study with methods used in this study to gauge the effect.

Occupations in STEM fields are areas which have a high disparity of gender representation. It has been described as a “leaky pipeline” due to the high rate of women falling out and leaving the field. A study asking women who were former students why they dropped out of the field a sense of not belonging or fitting in was a common answer (Rainey et al, 2018). Something that was in common for a lot of women students that chose to stay was that they had a mentor of the same gender. This could be that having someone that they can relate to in such a setting where they feel like they do not belong the mentor can act as an example of a way they can be part of this setting. This can have implications for direct policies to reach out to groups that are underrepresented in a field. The UN officially celebrates “women in STEM Day” in hopes of encouraging women to apply to these fields of study by increasing the visibility of women existing in these fields already. Highlighting groups of people that are underrepresented without highlighting that the groups are “exceptions” in the field can be the most important and effective way to change perceptions about the groups within the field.

Limitations of the study

A limitation of scope in this study is the low number of participants, and the homogenous participant pool. As pointed out in Gyax et al, (2008), there could potentially be a cohort difference due to the changing of the grammatical rules regarding use of masculine as gender neutral. While this gradual change started in the 70s there would be a smaller group of people who grew up with those grammatical rules now than when the study was conducted in 2008, but there could still be some differences in “echoes” of those grammatical rules. And intergenerational teaching could possibly have some subtle and smaller effect. Measuring these differences would require a great sample of people from different cohorts. The

participants are mostly all part of a similar group of people being all students, all between the ages of 20 to 27 years old, and mostly being women. This reduces the generalizability of the results to the population at large. On the other hand, previous studies on this topic have also used a relatively similar participant pool (i.e. Gygax et al, 2008) and having this similarity makes the findings of the different studies easier to compare.

The sample size of the study is pretty small and could result in a lower power than needed to detect smaller differences in sources of stereotype activation. It is above a preset limit of the very least 20 participants needed for a paradigm like the one used in this study, but it is still on the lower end of acceptable number of participants.

The results from this study are also limited to a very specific setting and culture. Since stereotypes and concepts of gender roles are dependent on the culture it exists in results from studies similar to this can only tell us to what extent the stereotypes examined are used in this specific culture within this specific time. As stereotypes will change so will how they can be used to make decisions. For instance, if an occupation used to have strong ties with a gender 100 years ago, but today it is considered completely neutral stereotyped then the participants don't really use stereotype information to decide the specific task. Then participants have to rely on other information to make snap decisions. In other words, the results will be ever changing if gender roles and stereotypes are also changing. The differences could also be very different in different parts of the country, or different social groups. A lot of nuances and variance can get lost without comparisons to different demographics.

Conclusion

The results from the analyses showed considerable differences in response time when family nouns were used instead of names, but there was no significant difference between in response time between the female family items or the male family items. This could reflect a change in the perceived gender roles in the family, but future studies are required to examine this. The results did not show any effect of the historical gendering of nouns leading to a male bias, but these findings need to be examined further in a different experimental design. The neutral roles having both the fastest response time and the highest rate of positive responses even when controlling for item gender might indicate that the stereotype effect is stronger

proscriptive than prescriptive which can have a lot of implications for how to view deviation from what is gender stereotypical. Future research on this topic is needed to explore further how this can affect the acceptance of women in men dominated spaces, and men in women dominated spaces.

References

- Brandth, B., & Kvande, E. (2009). Gendered or Gender-Neutral Care Politics for Fathers? *The ANNALS of the American Academy of Political and Social Science*, 624(1), 177–189. <https://doi.org/10.1177/0002716209334119>
- Bendixen, M., Helle, S., Langbach, T., & Rasmussen, K. (2014). Voldtektsmyter og kjønnsjåvinisme blant norske lekdommere og fagdommere. *Tidsskrift for norsk psykologforening*, 51, 7-13.
- Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics*, 28(2), 277-293. doi:10.1017/S0142716407070142
- Ellingsæter, A., & Gulbrandsen, L. (2007). Closing the Childcare Gap: The Interaction of Childcare Provision and Mothers' Agency in Norway. *Journal of Social Policy*, 36(4), 649-669. doi:10.1017/S0047279407001225
- Gabriel, U., Gygas, P. M., & Kuhn, E. A. (2018). Neutralising linguistic sexism: Promising but cumbersome? *Group Processes & Intergroup Relations*, 21(5), 844–858. <https://doi.org/10.1177/1368430218771742>
- Glick, P., and Fiske, S. T., (1996) The ambivalent sexism inventory: differentiating hostile and benevolent sexism. *Journal of personality and social psychology*, 70(3): 491-512

- Gorfein, D. S., Brown, V. R., & DeBiasi, C., (2007) The activation-selection model of meaning: explaining why the son comes out after the sun. *Memory & Cognition*, 35(8): 1986-2000
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K., (1998) Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74(6): 1464-1480
- Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23(3), 464–485.
doi:10.1080/01690960701702035
- Gygax, P., Gabriel, U., Lévy, A., Pool, E., Grivel, M., & Pedrazzini, E.,(2012) The masculine form and its competing interpretations in French: when linking grammatically masculine role names to female referents is difficult. *Journal of Cognitive Psychology*, 24(4): 395-408
- Hampson, S,H, (2018) Mothers do not make good workers: the role of work/life balance policies in reinforcing gendered stereotypes, *Global Discourse*, 8:3, 510-531,
DOI: [10.1080/23269995.2018.1521129](https://doi.org/10.1080/23269995.2018.1521129)
- Karsay, K., & Schmuck, D. (2019). “Weak, sad, and lazy fatties”: adolescents’ explicit and implicit weight bias following exposure to weight loss reality TV shows. *Media Psychology*, 22(1), 60-81. doi:<https://doi.org/10.1080/15213269.2017.1396903>
- Kim, J., Gabriel, U., & Gygax, P. (2019) Testing the effectiveness of the Internet-based instrument PsyToolkit: A comparison between web-based (PsyToolkit) and lab-based (E-Prime 3.0) measurements of response choice and response time in a complex psycholinguistic task. *PLoS ONE*, 14(9): e0221802.
<https://doi.org/10.1371/journal.pone.0221802>
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of verbal learning and verbal behavior*, 13(5), 585-589.

- Meadus, R. J., & Twomey, J. C. (2007). Men in nursing: making the right choice. *The Canadian nurse*, 103(2), 13–16.
- Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., Chiarini, T., Englund, K., Hanulikova, A., Ottl, A., Valdova, J., Von Stockhausen, L., & Sczesny, S. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior research methods*, 46(3), 841–871. <https://doi.org/10.3758/s13428-013-0409-z>
- NESH. (2021, 16. December). Forskningsetiske retningslinjer for samfunnsvitenskap og humaniora. <https://www.forskningsetikk.no/retningslinjer/hum-sam/forskningsetiskeretningslinjer-for-samfunnsvitenskap-og-humaniora/>
- Shaw, S., & Frisby, W. (2006). Can gender equity be more equitable?: Promoting and alternative frame for sport management research, education, and practice. *Journal of Sport Management*, 20, 483–509.
- Shaw, S., & Hoerber, L., (2003) A strong man is direct and a direct woman is a bitch: gendered discourses and their influence on employment roles in sport organisations. *Journal of Sport Management*, 17: 347-375
- Prentice, D. A., & Carranza, E. (2002). What Women and Men Should Be, Shouldn't be, are Allowed to be, and don't Have to Be: The Contents of Prescriptive Gender Stereotypes. *Psychology of Women Quarterly*, 26(4), 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International journal of STEM education*, 5(1), 1-14.
- Redaksjonen for norsk APA-stil. (2021). Norsk APA-manual: *En nasjonal standard for norskspråklig APA-stil basert på APA 7th* (Versjon 1.7). Unit. <https://www.unit.no/tjenester/norsk-apa-referansestil>

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin*, 26(11), 1315–1328. <https://doi.org/10.1177/0146167200263001>

Ward, L. M., & Grower, P. (2020). Media and the development of gender role stereotypes. *Annual Review of Developmental Psychology*, 2, 177-199.

Appendix

The full list of occupations (these were translated into Norwegian)

Role #	Feminine	Rating	Gender Neutral	Rating	Masculine	Rating
1	Manicurists	0.12	Archivists	0.07	Soldiers	0.15
2	Beauticians	0.15	Photographers	0.01	Crane operators	0.15
3	Prostitutes	0.17	Artists	0.03	Rubbish collectors	0.17
4	Wedding planners	0.15	Meteorologists	0.05	Governors	0.14
5	Birth attendants	0.2	Biologists	-0.04	Roofers	0.2
6	Exotic dancers	0.17	Trade unionists	0.01	Astronauts	0.17
7	Strippers	0.19	Acrobats	0.03	Wrestlers	0.2
8	Child-minders	0.22	Environmentalists	0.1	Forest rangers	0.22
9	Secretaries	0.25	Physiotherapists	0.1	Factory managers	0.25
10	Clairvoyants	0.24	Violinists	0.09	Computer technicians	0.23
11	Hairdressers	0.21	Oceanographers	0.09	Drummers	0.21
12	Groupies	0.23	Athletes	-0.08	Lighthouse keepers	0.24

Full list of names

Group	Female names	Rank	Male names	Rank	Group mean
A	Nina	18	Espen	19	
	Anette	51	Mats	50	
	Mari	60	Knut	61	
	<i>Mean</i>	<i>43</i>	<i>Mean</i>	<i>43,33333333</i>	<i>43,16666667</i>
B	Sandra	28	Geir	29	
	Ida	44	Robert	43	
	Camilla	55	Jonas	56	
	<i>Mean</i>	<i>42,33</i>	<i>Mean</i>	<i>42,66666667</i>	<i>42,5</i>
<i>Overall Mean</i>				<i>42,83333333</i>	

Full list of family roles

Group	Age	Female role	Norwegian	Male role	Norwegian
A	Older	Mothers	Mødre	Fathers	Fedre
	Equal	Sisters	Søstre	Brothers	Brødre
	Younger	Daughters	Døtre	Sons	Sønner
B	Older	Aunts	Tanter	Uncles	Onkler
	Equal	Female cousins	Kusiner	Male cousins	Fettere
	Younger	Nieces	Nieser	Nephews	Nevøer

