

More Skilled and Aware

*How Non-Stereotyped Occupational Gender Portrayals Impact Cognitive
Performance and Judgements of Competence*

Authors

Gisela Gardelius

Hanna Kopelman

Supervisor

Micael Dahlen

Submitted

May 18, 2021

Abstract

This research investigated whether non-stereotyped occupational gender portrayals in advertising can impact men and women's cognitive performance and ability to form accurate judgements of competence. T-test revealed that this may well be the case, although the effects varied between genders. More specifically, women performed significantly better on a cognitive test when exposed to a non-stereotyped ad, compared to a stereotyped one. However, no such effects were found for men. For both men and women, the non-stereotyped ad led to more accurate performance estimates as a result of a decrease in men's estimates and an increase in women's performance, respectively, which in turn generated more equal estimates between men and women. The results also pointed to the ads affecting both genders' likelihood to ask for help, regardless if the task at hand was perceived as feminine or masculine, but no statistically significant conclusions could be drawn for this variable overall because of too high p-values. Finally, ad and brand attitudes were tested; the non-stereotyped ad led to increased attitudes towards the ad, but only for women. No differences were found for brand attitudes for either gender. To interpret the results, the authors draw upon theory on the Dunning-Kruger effect and stereotype threat/lift, and discuss those theories in relation to concepts such as heuristics, ad reactance and signalling-effects. The findings suggest that non-stereotyped advertising in the form of occupational gender portrayals serve as an important tool in leveling men and women's performance and performance estimates across domains. Further, the results indicate the need for marketers to work proactively to craft non-stereotyped advertising, as even seemingly subtle stereotypes can elicit negative effects such as stereotype threat.

Keywords: Advertising, femvertising, gender stereotypes, Dunning-Kruger effect, performance estimates

Preface and Acknowledgements



This post was published on the Swedish Armed Forces' social media channels on the 8th of March 2021 to mark International Women's Day. The post was met with several comments similar to the one showcased – suggesting that marketing efforts depicting female soldiers is neither necessary, nor appreciated by the general public.

Our findings suggest otherwise.

NB: This is a montage of a Facebook post and (anonymized) comment from the Swedish Armed Forces' account, posted 03-08-2021. The caption and comment have been translated from Swedish. The text on the image reads "Typical girls."

A very special thanks to

Micael Dahlen

For your tireless support, creative contributions and positive spirit, and for always replying with invaluable input to our lengthy emails and countless concerns

Norstat

And especially Andreas, for your generosity and helpfulness, and for making this study possible in the first place

Fredrik Lange

For your support and guidance throughout our years at the Stockholm School of Economics, and especially during this project

Lisa-Mari and Alexander

For your critical eyes and for preventing the authors of this thesis from falling victim to the Dunning-Kruger effect

Thank you!

Hanna & Gisela

Table of Contents

1. Introduction	1
1.1 Background	2
<i>1.1.1 Previous Research and Expected Contribution</i>	<i>4</i>
1.2 Purpose and Research Question	5
2. Literature Review and Hypotheses Development	6
2.1 The impact of Non-Stereotyped Occupational Gender Portrayals on Ad and Brand Attitudes	6
2.2 The impact of Non-Stereotyped Occupational Gender Portrayals on Actual Test Performance	8
2.3 The impact of Non-stereotyped Occupational Gender Portrayals on Performance Estimates	9
2.4 The Impact of Non-Stereotyped Occupational Gender Portrayals on Likelihood to Ask for Help	12
3. Method	13
3.1 Pre-studies	14
<i>3.1.1 Pre-study 1</i>	<i>14</i>
<i>3.1.2 Pre-study 2</i>	<i>15</i>
3.2 Results of Pre-studies	16
<i>3.2.1 Stimuli Development</i>	<i>16</i>
<i>3.2.2 Quiz Development</i>	<i>16</i>
3.3 Main Study	17
<i>3.3.1 Study Design</i>	<i>17</i>
<i>3.3.2 Survey Design</i>	<i>17</i>
<i>3.3.3 Measures</i>	<i>18</i>
<i>3.3.4 Sampling and Data Collection</i>	<i>21</i>
3.4 Reliability and Validity	21
4. Results	22
4.1 Advertisement and Brand Attitudes	22
4.2 Actual Performance	23
4.3 Estimated Performance	24
<i>4.3.1 The Dunning-Kruger Effect</i>	<i>24</i>
4.4 Likelihood to Ask for Help	27
4.5 Summary of Hypotheses	28
5. Discussion	29
5.1 Advertisement and Brand Attitudes	29
5.2 Actual Performance	30
5.3 Estimated Performance	31

5.4 Likelihood to Ask for Help	34
6. Conclusions	35
7. Implications	36
8. Limitations and Suggestions for Future Research	38
9. References	40
10. Appendices	46
Appendix A: Stimuli	46
Appendix B: Results of Pre-studies	47
Appendix C: Quiz Questions	48
Appendix D: Distribution and Performance Across Quizzes	50
Appendix E: Scenarios for H7	50
Appendix F: Age Distribution of Sample in Main Study	51
Appendix G: Excerpt of Facebook Comments	51

1. Introduction

On average, humans make 35,000 decisions every day (Krockow, 2018). Whereas only a tiny fraction of these are conscious (Iyengar, 2011), the remainder are instead facilitated by heuristics – mental shortcuts that allow for speedier decision-making as a result of reduced cognitive effort (Tversky & Kahneman, 1974). However, these shortcuts work as a double-edged sword; the same heuristics that help us get through the day may also lead to humans developing oversimplified and potentially erroneous views of themselves or others.

Forming accurate judgements of our own competencies, and knowing how our skills and abilities compare to others is a crucial component in everyday life – it helps us understand when we are capable enough to rely on our own decisions, and when we instead need to seek assistance from those around us. Unfortunately, research suggests that we are not that good at evaluating ourselves accurately. On the contrary, top performers tend to underestimate their abilities, whilst bottom performers instead overestimate their abilities to an extent that violates the laws of statistics. This phenomenon was first presented by Justin Kruger and David Dunning in their 1999 article “Unskilled and Unaware of it: how difficulties in recognizing one’s own incompetence lead to inflated self-assessments”. The phenomenon has later become known as the Dunning-Kruger effect, and it further suggests that people base their self-perceptions of performance on preconceived notions about their own skills; a type of heuristic that may not necessarily correlate with their actual performance (Kruger & Dunning, 1999). Similarly, gender stereotypes constitute a different form of heuristic that can reduce cognitive effort (Tversky & Kahneman, 1974), but also lead to inaccurate and oversimplified ideas of individuals or groups of people; ideas that can be harmful or limiting to those subject to the stereotype (Eisend, 2010). Like the Dunning-Kruger effect, gender stereotypes facilitate our daily batch of 35,000 decisions, but arguably at the expense of an accurate worldview. From a marketing perspective the question is then what role advertising plays in sustaining or breaking down these stereotypes, and what extended effects stereotyped advertising may have on people and society as a whole. Could it be that stereotyped advertising not only perpetuates inaccurate beliefs of gender, but also leads to more erroneous views of our own skills and abilities?

1.1 Background

Gender role stereotypes in advertising have been a topic for extensive research during the past few decades (De Meulenaer et al., 2018; Eisend, 2010; Grau & Zotos, 2016), and remains a relevant topic as stereotypes continue to persist in advertising – especially in regards to female portrayals (Eisend, 2010). Gender role stereotypes can be defined as “beliefs about the psychological traits, behaviors and occupational status that are regarded as differentially appropriate for men or women” (De Meulenaer et al., 2018, p. 893) and, similar to other kinds of heuristics, help us orientate ourselves in everyday life by facilitating cognitive processing. However, gender role stereotypes in advertising can lead to inaccurate and oversimplified ideas of individuals or groups of people that can be harmful or limiting to those subject to the stereotype. For instance, gender stereotyping can cause women to experience body dissatisfaction (Eisend, 2010), affect their choice of career and restrict career opportunities (Eisend, 2010; Rice & Barth, 2017; Schuster & Martiny, 2017), as well as impair their cognitive and physical abilities in a variety of settings (Hively & El-Alayli, 2014; Inzlicht & Ben-Zeev, 2000; Spencer et al., 1999).

“Unchallenged, harmful stereotypes are a root cause of gender discrimination and inequality. As long as these attitudes are perpetuated and reinforced in advertising and media culture, we cannot hope to achieve a truly equal world.”

(Unstereotype Alliance, 2021)

Many of these manifestations can in turn be explained by stereotype threat – that is, the situational risk of being judged by stereotypes (see for example Spencer et al. 1999; Steele & Aronson, 1995). Individuals not subject to stereotyping may instead experience the opposite effect – a stereotype *lift* – which implies improvements in performance due to a downward comparison with a stereotyped group. Stereotyped advertising could therefore be suggested to lead to sustained gender inequalities in society, through the reinforcement of already existing gender stereotypes.

Although research in the area remains scarce, stereotype threat has also shown to impact performance assessments, similar to how it impacts performance (Tellhed & Adolfsson, 2018). In general, people tend to base their performance estimates on preconceived notions about their own skills – notions that may not necessarily correlate with their actual

performance. These preconceptions, perhaps in part based on gender stereotypes, may then lead to wrongful estimations that have little to do with actual performance. Across several studies in a variety of fields – such as psychology, grammar, logic, and even humor – participants in the bottom performing quartile not only overestimate their own raw scores, but also tend to think that they performed better than their average peer (Dunning et al., 2003; Kruger & Dunning, 1999). This is the result of poor performers suffering from a “dual burden” – not only do they lack the expertise needed to produce correct answers, but they also lack the metacognitive ability to recognize how poorly they are performing. Top performers, on the other hand, tend to have a relatively accurate sense of how well they perform in absolute terms (i.e. their estimation of their raw score on a test), but instead consistently overestimate how well *other* people are performing on the same test. Therefore, top performers typically underestimate their percentile rank relative to other people, since they tend to think that what is easy for them must also be easy for their peers (Dunning et al., 2003). This phenomenon – that bottom performers overestimate their performance whilst top performers underestimate their performance – is referred to as the Dunning-Kruger effect (Kruger & Dunning, 1999). In practice, it implies that those who are the most confident in their level of expertise are not necessarily those who should be; an effect that may well be perpetuated by stereotypical beliefs about one’s gender.

Whether advertising “mirror” or “mold” beliefs and values in society, gender stereotypes in advertising should be put under scrutiny. While the degree of stereotyping has decreased overall in recent years, this decrease is mostly due to progress in countries such as Japan, that have high masculinity indices (i.e., a preference for social gender role distinctions between men and women (De Meulenaer et al., 2018)), whereas low masculinity countries like Sweden have seen little improvement (Eisend, 2010). Furthermore, Eisend (2010) finds that occupational stereotypes remain the most common form of stereotyping in advertising, despite the occupational and educational status of women having changed radically over the years. As an example, even in stereotypically masculine occupations such as within the Swedish Armed Forces, the number of female employees has increased from 13 percent to 21 percent within the last four years (2016-2020) (Försvarsmakten, 2016; Försvarsmakten, 2021b). This imbalance is noteworthy since gender equality has seen the most progress within the labor market, perhaps as a result of occupational equality being the primary concern of gender-related policy. In other words, it seems as if marketing efforts are lagging, and given the importance of this area to the society at large (Eisend, 2010), occupational

gender role portrayals in advertising should be of particularly high interest to marketers and policy makers going forward.

At the same time, a silver-lining can be found in the current prevalent marketing trend of ads empowering women – sometimes referred to as “femvertising”. Åkestam et al. (2017) define femvertising as “advertising that challenges traditional female advertising” (p. 795) and find support for such non-stereotyped female portrayals in ads generating positive consumer responses. Reasonably, femvertising should also have extended societal effects. For instance, is it possible that such ads could work as a positive counterforce to gender stereotyping, improving abilities and cognitive processes of recipients?

1.1.1 Previous Research and Expected Contribution

Contrary to the well-researched topic of gender role stereotypes in advertising (De Meulenaer et al., 2018; Eisend, 2010; Grau & Zotos, 2016), the effects of femvertising and non-stereotyped gender role portrayals is a more recent subject in academia. Whereas a number of studies on the topic have found positive effects for non-stereotyped gender portrayals on consumer responses (Åkestam et al., 2017; Antioco et al., 2012; Bian & Wang, 2015; Liljedal et al., 2020), many of those studies have looked exclusively at the effects on female recipients. Seeing that advertising featuring women is typically believed to target a female customer segment, this pattern may well be explained by gender stereotypes predominantly pertaining to women (Eisend, 2010), which have caused the majority of prior research on (non-)stereotyped advertising to have focused on same-gender consumers only (Åkestam et al., 2021). That being said, consumers of all genders may be exposed to a given ad, and may also constitute appropriate target groups in their role as recommenders or purchasers of gifts. Therefore, the effects of non-stereotyped gender portrayals in advertising – regardless of the intended target group – is a topic that warrants more extensive research. To date, Liljedal et al. (2020) remains the only study on the effects of non-stereotyped gender role portrayals in advertising on both men and women, specifically within the scope of occupational gender roles. It should be noted that their study tested consumer responses using fictional ads and an anonymous (blurred) brand as stimuli. This is a common approach in experimental studies in advertising research to avoid confounding effects from previous brand experiences (Dahlen et al., 2014; Liljedal et al., 2020; Peter & Ponzi, 2018). However, a recent study by Feng et al. (2019) on femvertising suggests that consumer reactions are

contingent on perceived brand congruence, which warrants research on the effects of femvertising or non-stereotyped advertising from known brands.

Extensive research can also be found on the topic of gender stereotype threat. Mainly, this research has been conducted within the setting of psychology, and the treatment has typically consisted of an explicit statement concerning men or women's abilities in a given domain (e.g., Hively & El-Alayli, 2014; Schuster & Martiny, 2017; Spencer et al., 1999). A limited number of studies have also been performed where stereotyped ads or commercials have been used as a treatment to induce stereotype threat (Davies et al., 2002).

Furthermore, the Dunning-Kruger effect has been a subject for substantial research ever since the theory was coined by David Dunning and Justin Kruger in 1999 (see for instance Kruger & Dunning, 1999; Kruger & Dunning, 2002; Dunning et al., 2003; Schlosser et al., 2013; Simons, 2013), but to the authors' knowledge, only one prior study has investigated the relationship between stereotype threat and the Dunning-Kruger effect. This study, by Tellhed and Adolfsson (2018), found evidence that stereotype threat can impair performance assessments similarly to how it affects actual performance. However, no such study has been conducted from a marketing perspective, looking into the effects of advertising featuring non-stereotyped occupational gender roles.

The present study adds to the body of research on stereotype threat and stereotypes in advertising, constituting a key contribution to the scarce literature on the effects of non-stereotyped occupational gender portrayals in advertising. In addition, our study will investigate whether these effects can be attained even when the marketing message is from a known brand, and constitutes the first-ever study to look at the effects of stereotypes in advertising linked to the Dunning-Kruger effect.

1.2 Purpose and Research Question

The purpose of this study is to examine the effects on men and women of (non-)stereotyped occupational gender portrayals in advertising from a known brand, considering a number of relevant variables. These variables include: brand and advertisement attitudes, actual performance and performance estimates, and likelihood to ask for help in various domains.

The study adds to the existing research literature by examining the effects of (non-)stereotyped gender portrayals when the messenger is known to the recipients, which is considered to be relevant from a practical viewpoint as most marketing efforts are indeed from a known brand. In addition, it provides important contributions in linking (non-)stereotyped advertising to the Dunning-Kruger effect, which entails considerable implications for marketers and policy makers in their strive to eradicate gender inequalities.

We utilize an experimental study design to answer this overarching research question:

Can (non-)stereotyped gender portrayals in advertising impact men and women's cognitive performance and judgements of competence?

2. Literature Review and Hypotheses Development

2.1 The impact of Non-Stereotyped Occupational Gender Portrayals on Ad and Brand Attitudes

While stereotyped gender role portrayals in advertising have been studied extensively, research remains in conflict on consumer responses to such stereotypes. Congruity theory supports the notion that ads depicting gender roles in a more “congruous” manner (i.e., stereotyped portrayals) should produce higher favorability (De Meulenaer et al., 2018; Orth & Holancova, 2003). Other studies have found the opposite to be true; ads that depict gender roles in a non-stereotyped way can create positive signalling-effects, whereby the “in-congruous” gender role portrayal increases ad processing and in turn generates more positive ad and brand attitudes (Liljedal et al., 2020). For women, more positive consumer responses have been found for femvertising – that is, ads challenging traditional female advertising. By reducing some of the pressure advertising typically puts on the female recipients, femvertising creates lower psychological reactance to the ad compared to traditional advertising, ultimately resulting in higher ad and brand attitudes (Åkestam et al., 2017). An additional layer of complexity is added by Feng et al. (2019), who suggest that consumers’ reaction to femvertising is contingent on brand image and the brand’s congruence with the cause (e.g., dismantling gender stereotypes). Although the marketing message might resonate well with the audience, consumers do not always react well to ads with a feministic

or anti-sexist discourse. On the one hand, non-stereotyped portrayals of women can reduce the perceived discrepancy between the consumer and the model in the ad, in turn generating more positive attitudes towards both the ad and the brand. On the other hand, consumers might perceive the message as manipulative or insincere, as the ultimate goal of any corporation or profit-driven organization is to make money (Feng et al., 2019). Therefore, consumers may suspect that organizations using femvertising as part of their marketing strategy are merely capitalizing on feminism and societal movements to build their brand and profits. In that sense, it would be reasonable to assume that ads featuring (apparent) non-stereotyped gender portrayals can create strong ad reactance as well, given that the brand is perceived as incongruent with the perceived message.

Although several perspectives on women's reaction to femvertising have been offered, very few studies have investigated men's reaction (as an example, see Liljedal et al., 2020). However, research on men's reactions to gender role stereotypes indicates positive consumer responses to femvertising across genders. While men typically are less susceptible to stereotypical portrayals of their own gender (Pillaud et al., 2015; Tellhed & Adolfsson, 2018), "cross-gender" effects have been found for men; when exposed to ads with stereotyped portrayals of women, men show lower ad and brand attitudes, mediated by their beliefs of the ad having a negative influence on women, in turn creating ad reactance (Åkestam et al., 2021). As such, femvertising's positive effects on consumer responses should likely be extendable to male consumers. Again, effects may depend on the brand and its perceived congruence with the message, hence testing an ad from an anonymous or fictional brand could prove misleading. We use a known brand to evoke more realistic responses, testing the following hypotheses:

H1: Men and women will exhibit more positive advertisement attitudes when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal

H2: Men and women will exhibit more positive brand attitudes when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal

2.2 The impact of Non-Stereotyped Occupational Gender Portrayals on Actual Test Performance

Stereotypes in advertising do not only affect attitudes and behaviors related to the brand and the ad itself, but may also have extended, societal effects. When belonging to a group that is subject to widespread prejudices and stereotypes of one's abilities – such as women facing stereotypes of being worse leaders than men (Hoyt & Murphy, 2016; Johnson et al., 2008) or Afro-Americans of having poor intellectual abilities (Steele & Aronson, 1995) – environmental cues, such as advertising, can elicit a fear of being judged or conforming to the stereotype in question. This predicament, commonly referred to as stereotype threat, can in turn impair performance in a given situation by diverting focus from the task at hand or causing excessive cautiousness (Steele & Aronson, 1995). Sequentially, this could lead to a self-fulfilling prophecy whereby an individual of a stereotyped group reinforces the stereotypes others believe to be true. That is not to say that the person subject to stereotypes believes them herself; merely being aware of them is enough to suffer from stereotype threat. (ibid).

Much of the prior research on stereotype threat has focused on the effects of negative stereotypes on women, and have found adverse effects on women's performance in a number of (typically male-dominated) domains, such as math (Spencer et al., 1999) and athletics (Hively & El-Alayli, 2014). In many of these studies, a secondary effect has also been found within the outgroup (in this case, men), whereby men experience a stereotype lift boosting their performance after having been exposed to, or reminded of, a stereotype pertaining to women within the domain (see for instance Deshayes et al, 2020; Pruyers & Blais, 2014). An important question regarding both phenomena is how easily they are evoked; in a study on women's ability in math, the adverse effect on women's performance was shown to be the strongest when it was explicitly stated that the test produced gender differences (Spencer et al., 1999), but significantly subtler cues can evoke both stereotype threat and lift. Negative stereotypes seem to be linked automatically to evaluative tests – simply describing the task at hand as diagnostic is enough to induce stereotype threat or lift (Steele & Aronson, 1995; Walton & Cohen, 2003), and only when the stereotype is declared invalid or irrelevant does the effects wear off (Spencer et al., 1999; Walton & Cohen, 2003).

Stereotype threat can also be evoked simply by priming the subject to think of their race or gender – for instance, having Afro-Americans record their race before taking a (non-diagnostic) test is enough to produce a performance drop (Steele & Aronson, 1995). Likewise, the mere suggestion of sexism in a given situation can act as a reminder of prejudicial treatment, causing a drop in performance even in the absence of actual discrimination (Adams et al., 2006). Correspondingly, ads featuring stereotyped gender portrayals can act as a cue that elicits stereotype threat or lift, in turn causing recipients to experience a drop or boost in their performance in a given task (Davies et al., 2002). Moreover, even advertising which does not feature any women should be enough to induce stereotype threat/lift, if portraying men in traditionally masculine-coded occupational roles or otherwise stereotyped ways. We test the following hypotheses in the context of practical and theoretical knowledge, and cognitive reflection:

***H3a:** Women’s actual performance will be higher when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal*

***H3b:** Men’s actual performance will be lower when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal*

2.3 The impact of Non-stereotyped Occupational Gender Portrayals on Performance Estimates

Recent research also suggests that stereotype threat may undermine performance assessments in much the same way as it affects actual performance. In a study on Swedish high-school students’ SAT performance, Tellhed and Adolfsson (2018) found that under implicit stereotype threat, women’s performance assessments showed no relationship to their actual performance, whereas a moderate-sized relationship between the two could be found in the non-stereotype threat condition. Whilst both men and women tend to demonstrate low accuracy in their performance estimates, there appears to be a gender-related difference in the *direction* of the variance in these judgements (Bordalo et al., 2019; Herbst, 2020; Stankov et al., 2009). More specifically, men tend to either overestimate or accurately estimate their abilities, whilst women tend to either accurately estimate or underestimate their abilities (Herbst, 2020). A possible explanation for this may be a phenomenon described by researchers as the gender confidence gap (Herbst 2020; Michalak et al., 2017; Papyrina et al.,

2021; Vajapey et al., 2020), which implies that men typically show higher levels of confidence than women across several domains.

To understand the underlying mechanisms of why stereotype threat may impact performance estimates, one has to become familiar with how self-estimates are formed. Researchers have found that people base their performance estimates on information and feelings that they have available at the point in time when the estimation is made – regardless if those emotional states are attributable to the task at hand (Metcalf, 1998; Tversky & Kahneman, 1974). Ideally, performance estimates should instead be based on the experience of performing the task in question – for instance, how long it took to complete the task, or certainty in how it should best be executed. However, Dunning et al. (2003) found that people tend to estimate their performance based on a top-down approach – that is, they start with preconceived notions about their skills in a specific area (e.g. “I am a good driver”), and use this preconception to estimate how well they are doing on the specific task (e.g. jump starting a car). Whilst these types of heuristics may lead to speedier decision-making as a result of reduced cognitive processing, they may also cause premature cessation of problem-solving efforts. Ultimately, this may result in a faulty, over-optimistic assessment of one’s own or other people’s ability, which simply cannot be justified by actual performance (Metcalf, 1998; Shah & Oppenheimer, 2008; Dunning et al., 2003).

Similar to the Dunning-Kruger effect, stereotyped gender beliefs may be described as a type of representativeness heuristic. This implies that people tend to estimate the likelihood of an event (e.g. a person having a high IQ) by comparing it to an existing prototype that is already present in their minds – a prototype that displays what we think is the most typical or relevant example of a particular object or event (in this case, a person having a high IQ) (Tversky & Kahneman, 1974). The probability that this person is, for instance, a male, is then assessed by the degree to which this person is similar to, or representative of, the stereotype of a person with a high IQ (Bennett, 1996). As stated earlier, this approach tends to lead to erroneous judgements, as similarity and representativeness do not actually make an event or object more likely to occur – rather, people tend to overestimate the ability of the heuristic to accurately predict the likelihood of an event or an object (Tversky & Kahneman, 1974; Goodie & Fortune, 2013)

Since both the Dunning-Kruger effect and gender role stereotypes are types of heuristics leading to potentially faulty conclusions, one could argue that a disruption in these may result in more accurate performance estimates. As previously mentioned, Liljedal et al. (2020) found that non-stereotyped occupational gender portrayals lead to increased processing of the advertisement, similar to how creative advertising affects consumers. This increased information processing may in turn incline recipients to include more factors in their performance estimates, such as the time it took for them to complete the task, rather than relying on heuristics such as the top-down approach or gender stereotypes. Ultimately, this may lead to more nuanced – and thereby more accurate – performance estimates. As a result, given that women tend to underestimate their performance in general, their performance estimates are suggested to be higher when exposed to the non-stereotyped occupational gender portrayal. Following the same logic, as men generally tend to overestimate their performance, their performance estimates are suggested to be lower under the non-stereotyped occupational gender portrayal. Therefore, the following hypotheses are formulated:

***H4a:** Women will produce higher performance estimates overall when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal*

***H4b:** Men will produce lower performance estimates overall when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal*

As the Dunning-Kruger effect is typically measured by assessing both self-estimations of performance *and* estimated performance relative to others (see for instance Kruger & Dunning, 1999), the following sub-hypotheses are formulated:

***H5a:** Women will produce higher self-estimates when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal*

***H5b:** Men will produce lower self-estimates when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal*

***H6a:** Women will produce higher estimates of their performance relative to others when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal*

H6b: Men will produce lower estimations of their performance relative to others when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal

2.4 The Impact of Non-Stereotyped Occupational Gender Portrayals on Likelihood to Ask for Help

Further gender differences are to be found within the area of help-seeking, where men typically show a lower tendency of asking for help than women (Corney, 1990; Grinstein-Weiss et al., 2005; Keohane & Richardson, 2018; Vogel, 2014). Help-seeking is commonly defined as “a communication process with others, focusing on a specific problem of emotional pain which needs to be handled and that creates enough distress to seek either formal or informal sources of help” (Grinstein-Weiss et al., 2005, p. 766). Help can be further classified into two subcategories – formal help, which is conducted by professionals such as psychologists or doctors, and informal help, which stems from family-based sources such as relatives or peer group-based sources such as friends (Grinstein-Weiss et al., 2005).

Whilst most of the current research in the area has been conducted in a psychology or health setting, possible explanations to the gender differences in willingness to seek help draw on socialization theories and gender stereotypes (Juvrud & Rennels, 2017). Female socialization is typically based on values congruent with help-seeking, such as collaboration, dependence, and emotional expression; traits that also go in line with common gender stereotypes of women being nurturing, relational, and affectionate individuals (Shea et al., 2017). Therefore, help-seeking does not impose a threat to their self-image, which may explain why women are more prone to help-seeking than males (Ang et al., 2004). Male socialization, on the other hand, is typically based on values incongruent with help-seeking, such as achievement, independence of social support, self-reliance and suppression (Grinstein-Weiss et al., 2005; Tang et al., 2014). This implies that when a man must ask for help, consideration is given to the potential threats that the act of help-seeking imposes on his self-image of masculinity. More specifically, being in distress typically causes people to feel a loss of power, autonomy and control; traits that are otherwise stereotypically viewed as masculine. Additionally, the act of self-disclosing – that is, coming forward as vulnerable and asking for help – constitutes an added threat to their sense of masculinity, further diminishing men’s likelihood to ask for

help (Keohane & Richardson, 2018; Pederson & Vogel, 2007; Tang et al., 2014; Vogel, 2014).

Given that gender stereotypes influence help-seeking behavior, a disruption in gender role portrayals could be argued to increase information processing – similar to the reasoning on performance estimates – and lead to more accurate evaluations of the need for assistance (Liljedal et al., 2020). In addition, the disruption might attenuate gender-related expectations and the perceived discrepancies between men and women, mitigating the threat that help-seeking may otherwise impose on men's self-image of masculinity (Keohane & Richardson, 2018; Pederson & Vogel, 2007; Vogel, 2014). In other words, as men's reluctance to ask for help may partially be explained by prevailing gender stereotypes, the non-stereotyped occupational gender portrayal is suggested to increase their likelihood to ask for help. Following the same logic, as women are typically more prone than men to ask for assistance, the non-stereotyped occupational gender portrayal is instead suggested to decrease their likelihood to ask for help. Therefore, the following hypotheses are formulated:

***H7a:** Women will produce lower likelihood to ask for help when exposed to the ad featuring a non-stereotyped (versus stereotyped) gender portrayal*

***H7b:** Men will produce higher likelihood to ask for help when exposed to the ad featuring a non-stereotyped (versus stereotyped) gender portrayal*

3. Method

To test the hypotheses, a quantitative study in the form of an experiment was conducted. Two different stimuli were created; one that was deemed to be a non-stereotyped occupational gender portrayal, and one that was deemed to be a stereotyped occupational gender portrayal. The effects on ad and brand attitudes were then measured through questions using a semantic differential scale. To test the effects of the treatments on actual and estimated performance, participants had to take a quiz consisting of cognitive questions relating to either practical and theoretical knowledge or cognitive reflection, and were then asked to evaluate their performance. Additionally, different scenarios were presented to the respondents to assess their likelihood to ask for help in different domains. The experimental design was considered

to be suitable for this study since our aim is to test for differences between treatment groups, in accordance with Söderlund's (2018, pp. 12-13) suggestions. Furthermore, the quantitative approach allows for statistical significance testing and the possibility to generalize the results to a larger group based on a sample.

3.1 Pre-studies

Two pre-studies were conducted with the aim to test the appropriateness of the stimuli created for the main study, and to identify questions to be included in a set of quizzes that would constitute the basis for measuring actual and estimated performance in the main study. As the gender confidence gap has been shown to be task-dependent (Bennett, 1996; Bennet, 1997; Furnham, 2005; Papyrina et al., 2021), the pre-studies aimed to identify quiz questions covering a range of topics to account for gender differences in performance and confidence across domains.

3.1.1 Pre-study 1

The first pre-study was in the form of a survey distributed through the authors' personal networks and social media channels, and collected responses between the 24th and 25th of February 2021. The respondents consisted of a convenience sample ($n=23$, 52 percent female).

The purposes of this pre-study were twofold: to establish whether the developed stimuli were indeed perceived as (non-)stereotyped occupational gender portrayals, and to develop two quizzes consisting of masculine- and feminine-coded questions that possessed the same level of difficulty as a previously developed gender-neutral cognitive reflection-test. Therefore, the first pre-study examined to what extent respondents correctly answered the questions, as well as the perceived advantage of men versus women on each question (i.e., whether the questions were perceived as masculine- or feminine-coded).

The pre-study consisted of two different mock advertisements created by the authors using stock photos from the Swedish Armed Forces' website. The first advertisement pictured a man in military uniform with the message "Apply Now". The second advertisement was from the same brand and included the same message, the only difference being that the person in

the picture was a woman in uniform (see Appendix A). The Swedish Armed Forces was chosen as the brand in question for two reasons: first, because the military occupation is typically considered to be masculine (Liljedal et al., 2020), and second, because the majority of the Swedish population were assumed to be aware of, and have a prior relation to the brand. Although many other experimental studies in advertising research make the brand anonymous to avoid confounding effects from previous brand experience (see for instance Liljedal et al., 2020), the authors' of this report intentionally chose not to do so to test the effects of stereotype threat from a known brand.

The participants in the pre-study got to see the ads one at a time, and were then asked to answer three questions: what gender they perceived that the ad targeted, whether it depicted a stereotyped portrayal of the displayed gender, and whether it depicted a stereotyped portrayal of the occupation of a soldier, using similar scales as in previous research in the area (see Liljedal et al., 2020).

Then, respondents were asked to take a quiz of ten questions developed by the authors, where each question had three response alternatives – one correct, and two incorrect options. After each question, respondents were also asked to rate whom they thought would generally perform better on the question – men or women – on a scale ranging from 0-10, where 0 equalled men and 10 equalled women. In the final part of the study, respondents were asked to state their gender, and if they had used any external sources such as Google to complete the quiz. Respondents who admitted to having used external sources were disregarded from the sample. Finally, respondents were also able to provide general feedback on the questions in free-form text.

3.1.2 Pre-study 2

The second pre-study followed the exact same design as the first pre-study, but tested ten different quiz items. Again, the respondents consisted of a convenience sample (n=30, 67 percent female), with the survey being distributed via the authors' personal networks and social media channels. The survey recorded responses between the 24th and 25th of February 2021.

3.2 Results of Pre-studies

3.2.1 Stimuli Development

	Male Soldier	Female Soldier
The ad is mainly targeted to... 0 = Men, 10 = Women	M = 1.83	M = 8.65
The picture portrays the occupation soldier in a stereotyped manner 0 = Do not agree at all, 10 = Agree Completely	M = 7.87	M = 4.96
The picture portrays men/women in a stereotyped manner 0 = Do not agree at all, 10 = Agree Completely	M = 6.13	M = 1.87

Table 1: Results of Pre-Study 1 – Stimuli Development

The first pre-study established that there were indeed differences in how the respondents perceived the two advertisements. As for the ad featuring a male soldier, the mean value clearly indicated that respondents perceived it to be targeted to men, and that the picture portrayed both the occupation and men in a stereotyped manner. As for the ad featuring a female soldier, respondents instead perceived the ad to be targeted to women, and depict females in a non-stereotyped way. Although it is noted that the mean value of perceived stereotyped occupational portrayal for this ad was fairly close to the midpoint of 5.5, there is still a clear difference in how the respondents rated the two different ads in terms of stereotyped occupational portrayals (mean difference: +2.91). Despite the sample being too small to test for statistical significance, the results provided strong indications of the intended effects, why the ads later were used as stimuli in the main study. Hereafter, the term “non-stereotyped ad” will refer to the ad featuring the female soldier, whereas the term “stereotyped ad” will refer to the ad featuring the male soldier.

3.2.2 Quiz Development

The joint purpose of pre-study 1 and 2 was also to identify four questions that participants generally believed men would perform better at, and four questions where the same held true for women. An additional criterion for the questions was a difficulty level that was on par with that of the gender-neutral cognitive reflection-test – around 44 percent correct answers

on average (Primi et al., 2016). The full results for all 20 questions that were tested can be found in Appendix B.

The questions that were ultimately chosen for the feminine and masculine quiz in the main study were those that showed a strong perceived advantage for men and women, respectively. The chosen feminine-coded questions were on average answered correctly by 49.64 percent of the respondents, whereas the chosen masculine-coded questions were answered correctly by 48.81 percent of the respondents. Although there is a slight deviation in average performance from that of the gender-neutral cognitive reflection-test, this was deemed to be acceptable as the most important factor for the quiz items was that there would be room for participants to answer incorrectly to be able to assess the Dunning-Kruger effect. All three quizzes can be found in Appendix C.

3.3 Main Study

The main study aimed to examine the effects of the two different treatments on men and women's advertisement and brand attitudes (H1-H2), actual and estimated performance on a random quiz (H3-H6), and likelihood to ask for help in different scenarios (H7).

3.3.1 Study Design

The study employed an experimental approach in the form of a between-subjects design, where each participant was randomly exposed to one of the two treatments, to be able to find causal effects between the two advertisements and subsequent dependent variables. To test the effect of respondents' gender as an additional causal variable, a 2x2 factorial design was employed.

3.3.2 Survey Design

Broadly speaking, the questionnaire consisted of three parts – an introduction, treatment, and the subsequent questionnaire. Since the study was conducted in Sweden and the majority of the respondents were expected to be native Swedish speakers, the survey was formulated in Swedish to avoid any confusion in language.

At the very start of the survey, respondents were informed of the layout and estimated response time of the questionnaire, and were notified that all responses were recorded anonymously. Respondents were prompted to answer the upcoming quiz to the best of their knowledge without seeking assistance from others, and were informed that no specific prerequisites were needed to participate in the survey.

Then, respondents were randomly presented with one of the two treatments tested in Pre-study 1, and asked to answer a set of questions on their attitudes to the advertisement and brand. Thereafter, they completed a random quiz (feminine-, masculine- or neutrally-coded questions), after which they were asked to estimate their performance, and then answer questions regarding their likelihood to ask for help in a set of different scenarios. At the end of the survey, respondents were informed that the advertisement they had seen was fictive and created for the sole purpose of the study.

3.3.3 Measures

Advertisement and Brand Attitudes

These two variables were measured through a 10-point Osgood scale and measurements recommended by Holbrook (1987). The questions used were “What is your attitude towards the advertisement?” (Bad-Good, Dislike-Like, Negative attitude-Positive attitude; Cronbach’s $\alpha = 0.942$), and “What is your attitude towards the brand behind the advertisement?” (Bad-Good, Dislike-Like, Appealing-Not appealing, Cronbach’s $\alpha = 0.929$).

Actual Performance

This variable was defined in accordance with previous studies on the Dunning-Kruger effect as the percentage of correct responses on the quiz taken, and was used to determine respondent’s actual percentile ranking (0-100 percent) for the Dunning-Kruger analyses (Kruger & Dunning, 1999; Schlosser et al., 2013). Actual performance was measured by assigning each respondent one of the three quizzes related to either masculine-, feminine- or neutrally-coded domains. The quizzes were randomly assigned (see Appendix D) to eliminate any potential gender-bias and/or advantage for each gender (e.g. men performing better in masculine-coded domains). Although there were differences in how well men and women performed on the masculine- and feminine-coded quizzes respectively, the differences offset each other on an aggregate level. In other words, there were no significant differences between the genders in actual performance overall (see Appendix D). Therefore, unless

otherwise specified, *Actual Performance* will hereafter refer to the respondents' percentage of correct answers on the quiz completed, regardless if feminine, neutral or masculine.

For the feminine- and masculine-coded quizzes, the questions previously tested in the pre-studies were used. For the neutral quiz, questions consisted of the cognitive reflection test (CRT) – a three-question test measuring peoples' ability to reflect upon a problem to find the correct answer, rather than choosing the answer that seems right intuitively but is in fact incorrect (Frederick, 2005). One additional question was added to the CRT to match the number of questions in the other two quizzes, and to allow for greater variation in actual performance between respondents. The fourth question was retrieved from a longer version of the CRT-test, developed by Primi et al. (2016). Although there is a slight tendency for men to perform better on the CRT-test (Frederick, 2005), likely due to its mathematical content, no gender differences were found in terms of actual performance in our study (see Appendix D). As previously mentioned, all three quizzes can be found in Appendix C.

Estimated Performance

This measure constituted two metrics: *self-estimates* and *performance relative to others*, reported by the participants after completing the quiz. *Self-estimates* is defined as estimations of one's own performance in absolute terms, in line with Schlosser et al. (2013), and was measured by the question "How many correct answers do you think you got on the quiz you just took (out of 4 questions)?", with the response alternatives 0, 1, 2, 3, and 4. For comparability purposes, the raw score estimate was later converted into a percentage by the authors (e.g. an estimated score of 2 was converted to 50 percent). *Performance relative to others* is defined as estimations of one's own performance in relative terms, in line with Schlosser et al. (2013). This metric was measured by the question "Compared to other people from the Swedish population that have taken this quiz, how would you rate your performance? E.g. a rank of 25% implies that you outperformed 25% of the Swedish population that have taken this quiz.". It should be noted that this measurement constitutes both a type of self-estimate and peer-estimate. The response alternatives were in the form of a slider scale ranging from 0 (worse than all other participants) to 100 (better than all other participants). Both of these questions are commonly used for measuring the Dunning-Kruger effect (Kruger & Dunning, 1999; Dunning et al., 2003; Schlosser et al., 2013). The variables were both measured separately, and combined as an index (Cronbach's alpha = 0.715).

Likelihood to Ask for Help

This variable measured the respondent's likelihood to ask for help in different scenarios, by employing a 10-point Likert scale where only the endpoints were labelled (1=Not likely at all, 10 = Very likely) (Likert, 1932). As previous research in this area has mainly been conducted in the field of medicine and psychology, the question for this particular section was developed by the authors together with faculty members at the Stockholm School of Economics. The question read "How likely is it that you would ask for help with the following?", followed by a clarification to include both formal and informal help: "*To ask for help could for instance mean to ask a relative for advice or employ a professional*". The question was followed by six scenarios, three of which were determined to be masculine-coded, and three of which were feminine-coded. This allowed for the subsequent results to be analyzed both on an aggregate level – thereby accounting for overall differences related to the gender confidence gap – and separated by domain to identify potential task-dependent differences between men and women. The scenarios were developed using the 20 questions from the pre-studies and their perceived level of masculinity/femininity as a basis. As an example, respondents believed men to perform better on the question "What order should be followed when jump starting a car?", a question which was then transformed into the scenario "How likely is it that you would ask for help when jump starting a car?". The six scenarios can be found in Appendix E.

The order in which the six scenarios were listed was randomized in the survey. No indices were created of the masculine-coded and feminine-coded scenarios respectively, due to Cronbach's alpha being too low (Cronbach's $\alpha_{\text{masculine domains}} = 0.492$, Cronbach's $\alpha_{\text{feminine domains}} = 0.673$), and the results were instead reported separately.

Demographic Questions

The study also collected demographic information, to ensure the sample's representativeness and identify respondent gender for the hypotheses testing. Gender was assessed using a multiple choice-question with the options "Male", "Female", "Other/Prefer not to answer", and age was assessed through a multiple-choice question containing seven brackets (18-25, 26-35, 36-45, 46-55, 56-65, 66-75, 75+).

3.3.4 Sampling and Data Collection

In collecting the responses, Norstat Sverige AB's services were used – a market research company specializing in data collection through paid survey respondents. The nature of these services may result in respondents replying to numerous surveys during a brief period of time, which could increase the risk of respondents skimming through the survey in an inattentive manner. To mitigate this risk, a control question was included in the middle of the survey ("Choose the number 7": 10, 7, 2). All of the respondents in the sample answered this question correctly, and no respondents were therefore disqualified on the basis of failing the control question.

The sample was intended to be representative of the Swedish population, given the study's between-groups design. Norstat were therefore instructed to gather respondents aged 18 and older. The survey was published on the 11th of March 2021, and collected a total of 258 valid answers until the 29th of March 2021. It should be noted that all of the questions in the survey were mandatory to answer, hence there were no missing values in our dataset. Out of the 258 respondents, 59.7 percent were women and 40.3 percent were men. The ages of the respondents varied from 18 to 75+, with the median age bracket being 46-55 years of age. For more information regarding the age distribution, see Appendix F.

3.4 Reliability and Validity

To ensure reliability, measures that have been established in previous research were used whenever possible. Further, multi-item measures were used to measure psychological reactions such as attitudes to avoid any potential sources of errors. To ensure that the results were coherent across questions, Cronbach's alpha was also computed when applicable. When Cronbach's alpha > 0.7, the measurements were presented as an index, otherwise the results were reported separately (Söderlund, 2018, pp. 135-136).

In line with Söderlund (2018, pp. 173-174), external validity was ensured by the use of a relatively large sample, deemed to represent the Swedish population. To ensure internal validity, the treatments were made identical to the largest possible extent except for the gender of the person featured; the setting, clothing, facial expression, and stance were the same in both advertisements to avoid any confounding effects unrelated to the portrayed

soldier's gender (Söderlund, 2018, pp. 172-173). To avoid arriving at inaccurate conclusions regarding the causal effects of the treatments, a manipulation check was carried out (Söderlund, 2018, p. 87) in pre-study 1, which verified that the advertisements were indeed perceived as (non-)stereotyped.

In the quizzes used to test the dependent variable “actual performance”, four questions were included in each; whereas standard cognition tests are usually more extensive, it was not viable for the scope of this thesis to include a larger number of questions. Furthermore, for the feminine and masculine quizzes, the included questions covered a variety of domains to eliminate the effects of a specific respondent being very good in a certain domain. To further diminish potential sources of errors such as straightlining, both the order of the quiz items and their corresponding response alternatives were randomized. All in all, we argue that the reliability and validity of the study is adequate.

4. Results

In the following section, the results of the study will be presented in order of the previously presented hypotheses. Independent samples t-tests were employed to test the hypotheses; mean values, standard deviations and statistical significance of mean differences are reported in the tables below. All results that are deemed to be relevant – regardless if significant or not – will be presented in this section and subsequently commented in the discussion.

4.1 Advertisement and Brand Attitudes

	Women			Men		
	Stereotyped n=79	Non- Stereotyped n=75	Sig.	Stereotyped n=55	Non- Stereotyped n=49	Sig.
Advertisement Attitudes	M = 6.85* SD = 2.05	M = 7.39* SD = 2.24	t = -1.56 p = 0.06	M = 7.72 SD = 1.98	M = 7.73 SD = 2.03	t = -0.05 p = 0.48
Brand Attitudes	M = 7.26 SD = 1.99	M = 7.12 SD = 2.10	t = 0.43 p = 0.33	M = 8.06 SD = 1.76	M = 8.03 SD = 1.89	t = 0.09 p = 0.46

* p<0.10, ** p<0.05, *** p<0.01

Table 2: Mean Comparisons of Advertisement and Brand Attitudes

H1 and H2 suggested that both men and women will exhibit more positive advertisement and brand attitudes when exposed to the non-stereotyped ad, compared to the stereotyped ad. The results for H1 and H2 are presented in Table 2.

For female respondents, there was a significant increase in advertisement attitudes towards the non-stereotyped ad, compared to the stereotyped ad. No significant differences were found for women in terms of brand attitudes, although it should be noted that the attitudes towards the non-stereotyped ad were slightly lower than the stereotyped ad (mean difference -0.14). For male respondents, no significant or noteworthy differences between the two treatments were found.

4.2 Actual Performance

	Women			Men		
	Stereotyped n=79	Non-Stereotyped n=75	Sig.	Stereotyped n=55	Non-Stereotyped n=49	Sig.
Actual Performance	M = 43.04* SD = 29.67	M = 50.00* SD = 32.36	t = 1.39 p = 0.08	M = 48.64 SD = 30.21	M = 48.98 SD = 26.98	t = 0.06 p = 0.48

* p<0.10, ** p<0.05, *** p<0.01

Table 3: Mean Comparisons of Actual Performance (in %)

H3a and H3b predicted that women's (men's) actual performance would be higher (lower) when exposed to the non-stereotyped ad, compared to the stereotyped ad. The results for H3a and H3b are presented in Table 3.

For female respondents, actual performance was significantly higher in the non-stereotyped condition, compared to the stereotyped condition (mean difference: +6.96 percentage points). For male respondents, there were no significant differences in actual performance, although the respondents exposed to the non-stereotyped ad on average produced slightly higher results than those exposed to the stereotyped ad (mean difference: +0.34 percentage points).

4.3 Estimated Performance

	Women			Men		
	Stereotyped n=79	Non-Stereotyped n=75	Sig.	Stereotyped n=55	Non-Stereotyped n=49	Sig.
Estimated Performance (index)	M = 52.52 SD = 23.75	M = 55.97 SD = 23.51	t = 0.91 p = 0.18	M = 64.73*** SD = 20.32	M = 55.39*** SD = 18.59	t = -2.44 p = 0.01
Self-Estimations	M = 61.08 SD = 28.52	M = 66.00 SD = 27.46	t = 1.09 p = 0.14	M = 75.45** SD = 23.32	M = 67.35** SD = 24.58	t = -1.72 p = 0.05
Estimated Performance Relative to Others	M = 43.96 SD = 24.75	M = 45.95 SD = 23.42	t = 0.51 p = 0.31	M = 54.00*** SD = 23.49	M = 43.43*** SD = 21.27	t = -2.40 p = 0.01

* p<0.10, ** p<0.05, *** p<0.01

Table 4: Mean Comparisons of Estimated Performance (in %)

H4-H6 expected that women (men) would produce higher (lower) performance estimates when exposed to the non-stereotyped ad, compared to the stereotyped ad – both overall (H4a, H4b), as well as in terms of self-estimates (H5a, H5b) and estimated performance relative to others (H6a, H6b). The results for H4-H6 are presented in Table 4.

For women, there were no significant differences in estimated performance – neither overall, nor in terms of self-estimates or estimated performance relative to others. The results still indicate that women estimated their performance as higher when exposed to the non-stereotyped ad, especially in terms of self-estimations (mean difference: +4.92 percentage points). It should however be noted that women did in fact also produce higher *actual* performance (H3a) in the non-stereotyped condition. Men, on the other hand, displayed significant differences in estimated performance on all parameters – overall, men exposed to the non-stereotyped ad estimated their performance 9.34 percentage points lower than when exposed to the stereotyped ad, a pattern which also held true across self-estimates (mean difference: -8.10 percentage points) and estimated performance relative to others (mean difference: -10.57 percentage points).

4.3.1 The Dunning-Kruger Effect

By plotting the participants' estimated performance as a function of their actual performance on the test, the Dunning-Kruger effect per gender and treatment can be observed in the below graphs.

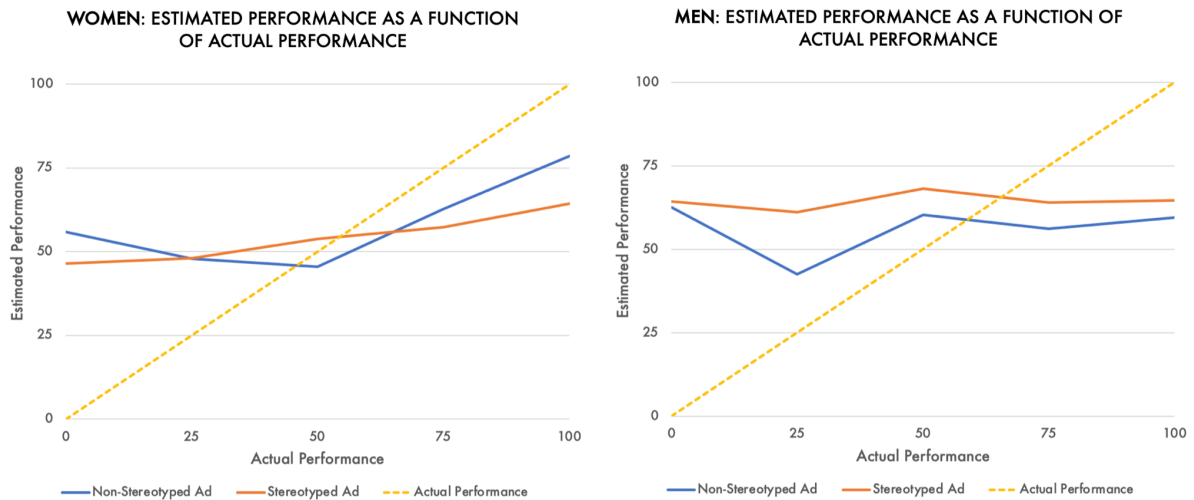


Figure 1: Estimated Performance as a Function of Actual Performance

As can be seen, evidence for the Dunning-Kruger effect was indeed found in our sample for both genders. Regardless of treatment, bottom performers (actual performance ≤ 25 percent) consistently overestimated their performance (46th to 64th percentile rank), whereas top performers (actual performance ≥ 75 percent) generally underestimated their performance (56th to 78th percentile rank). There were also differences between the genders depending on treatment; both bottom- and top-performing women generally rated their performance as higher when exposed to the non-stereotyped ad, whereas men – regardless of actual performance – consistently rated their performance lower across actual performance percentiles in the non-stereotyped condition compared to the stereotyped condition.



Figure 2: Self-Estimates as a Function of Actual Performance

For self-estimates in particular, women’s ratings were generally not too affected by the different treatments. For men, however, bottom performers generally rated themselves significantly lower (i.e. more accurately) when exposed to the non-stereotyped ad. Under the stereotyped condition, the Dunning-Kruger curve was fairly flat for men – regardless of actual performance, most men thought that they performed around the 65th-80th percentile.

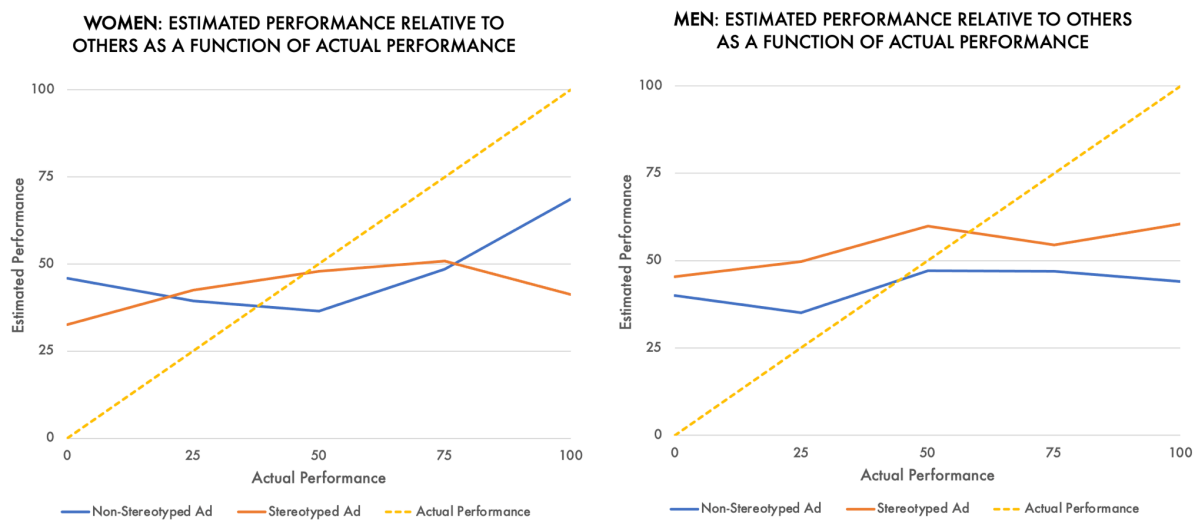


Figure 3: Estimated Performance Relative to Others as a Function of Actual Performance

In terms of estimated performance relative to others, female bottom- and top-performers rated themselves higher when exposed to the non-stereotyped ad, whereas average performing females (actual performance = 50 percent) rated themselves lower when exposed to the non-stereotyped ad. Men consistently rated themselves lower when exposed to the non-stereotyped ad across actual performance ranks (around 11.4 percentage points lower compared to when seeing the stereotyped ad).

4.4 Likelihood to Ask for Help

How likely is it that you would ask for help with the following?	Women			Men		
	Stereotyped n=79	Non-Stereotyped n=75	Sig.	Stereotyped n=55	Non-Stereotyped n=49	Sig.
... jump start a car	M = 6.54 SD = 3.53	M = 6.27 SD = 3.39	t = -0.50 p = 0.31	M = 3.64 SD = 3.19	M = 3.74 SD = 3.25	t = 0.16 p = 0.44
... connect a laptop to a large screen display	M = 4.35 SD = 3.47	M = 4.19 SD = 3.23	t = -0.31 p = 0.38	M = 3.33 SD = 3.12	M = 3.80 SD = 3.43	t = 0.73 p = 0.23
... invest on the stock market	M = 5.92 SD = 3.10	M = 5.76 SD = 3.06	t = -0.33 p = 0.37	M = 5.40 SD = 3.54	M = 5.49 SD = 2.98	t = 0.14 p = 0.45
... remove a severe stain from a piece of clothing	M = 3.42 SD = 2.93	M = 3.24 SD = 2.80	t = -0.38 p = 0.35	M = 4.15*** SD = 2.74	M = 5.51*** SD = 2.78	t = 2.52 p = 0.01
... take up a pair of pants that are too long	M = 3.87 SD = 3.29	M = 3.55 SD = 3.48	t = -0.60 p = 0.28	M = 7.13 SD = 3.29	M = 6.71 SD = 3.08	t = -0.66 p = 0.26
... wash a down jacket	M = 2.96 SD = 2.74	M = 3.23 SD = 2.95	t = 0.58 p = 0.28	M = 5.02 SD = 3.06	M = 5.35 SD = 3.26	t = 0.53 p = 0.30

* p<0.10, ** p <0.05, *** p <0.01

Table 5: Mean Comparisons of Likelihood to Ask for Help

H7a and H7b hypothesized that women (men) would produce lower (higher) likelihood to ask for help when exposed to the non-stereotyped ad, compared to the stereotyped ad. The results for H7a and H7b are reported in Table 5.

For women, no significant differences between the conditions were found. However, the results on five out of six scenarios point in the same direction – when exposed to the non-stereotyped ad, females report a lower likelihood to ask for help compared to when seeing the stereotyped ad. When it comes to washing a down jacket, however, the results for women instead point in the opposite direction, with female respondents reporting a slightly higher likelihood to ask for help when exposed to the non-stereotyped ad. Male participants showed a similar pattern in five out of six cases, where they reported a higher likelihood to ask for help when exposed to the non-stereotyped ad compared to the stereotyped ad, although mean differences were non-significant with the exception of removing a severe stain from a piece of clothing. The only deviating result was in the instance of men taking up a pair of pants – in this case, they showed a lower likelihood to ask for help after being exposed to the non-stereotyped ad compared to the stereotyped ad.

4.5 Summary of Hypotheses

Below is a summary of all hypotheses and whether they were supported or unsupported in terms of statistical significance ($p < 0.1$).

H1: Men and women will exhibit more positive advertisement attitudes when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal	Partially Supported
H2: Men and women will exhibit more positive brand attitudes when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal	Unsupported
H3a: Women's actual performance will be higher when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal	Supported
H3b: Men's actual performance will be lower when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal	Unsupported
H4a: Women will produce higher performance estimates overall when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal	Unsupported
H4b: Men will produce lower performance estimates overall when exposed to the ad featuring a non-stereotyped (versus stereotyped) occupational gender portrayal	Supported
H5a: Women will produce higher self-estimates when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal	Unsupported
H5b: Men will produce lower self-estimates when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal	Supported
H6a: Women will produce higher estimates of their performance relative to others when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal	Unsupported
H6b: Men will produce lower estimations of their performance relative to others when exposed to the non-stereotyped (versus stereotyped) occupational gender portrayal	Supported
H7a: Women will produce lower likelihood to ask for help when exposed to the ad featuring a non-stereotyped (versus stereotyped) gender portrayal	Unsupported
H7b: Men will produce higher likelihood to ask for help when exposed to the ad featuring a non-stereotyped (versus stereotyped) gender portrayal	Unsupported

5. Discussion

Our results suggest that (non-)stereotyped gender portrayals in advertising can in fact impact cognitive performance and judgements of competence, although the effects differ between genders. The following section will discuss the results in detail in relation to theory on, inter alia, stereotype threat/lift and increased information processing as a result of non-stereotyped occupational gender portrayals.

5.1 Advertisement and Brand Attitudes

H1 and H2 stipulated that ad and brand attitudes would be higher for both men and women in the non-stereotyped condition. In regards to attitudes towards the ad, women exhibited a significant increase in the non-stereotyped condition. This supports previous findings on femvertising generating higher advertising attitudes among women (Åkestam et al., 2017), and provides strong evidence against congruity theory in this setting, as military jobs are traditionally highly male-dominated (Försvarsmakten, 2021), and also perceived as such (Liljedal et al., 2020). Possible explanations to the increase in advertisement attitude for women after being exposed to the non-stereotyped ad draw upon reactance and signalling theories; by evoking positive feelings such as empowerment, the ad may have led to lower psychological reactance, in turn increasing advertisement attitudes among women, in accordance with Åkestam et al.'s findings (2017). Alternatively, the perceived uniqueness of the ad could have contributed to the increase in attitude, in line with Liljedal et al. (2020). It should be noted that the brand context may have impacted the results; the Swedish Armed Forces is a government body with high public trust among large parts of the Swedish population (Försvarsmakten, 2021a), which may have eliminated possible skepticism towards the sincerity of the message and prevented ad reactance.

For men, there were no differences in attitudes neither towards the ad, nor the brand. These results suggest that the findings on cross-gender effects by Åkestam et al. (2021) cannot be extended to femvertising. Although men respond poorly to ads which they perceive as harming to women, the opposite could not be proven. In other words, ads that empower women by challenging traditional female portrayals are not necessarily favored over traditional male portrayals by men. These results also stand in contrast to the findings by

Liljedal et al. (2020), which showed that both men and women exhibited more positive consumer responses when ads featured non-stereotyped occupational gender portrayals. A difference between their study and ours is that the former tested ads using blurred brands, whereas the latter used a visible logo in the ads. A possible explanation to the conflicting results could therefore relate to the respondents' prior attitudes to the brand used in this study – namely, the Swedish Armed Forces.

Interestingly, there were no significant differences in brand attitudes for women between-groups. Again, seeing that we used a visible brand that also happens to be very well-known among the Swedish population, prior brand attitudes could have affected the results. While female respondents may have received the message in the non-stereotyped (versus stereotyped) ad better, it was perhaps not enough to change their prior attitudes towards the brand. H1 was in other words only partially supported, whereas H2 could not be supported.

5.2 Actual Performance

In accordance with H3a, women's actual performance was significantly higher in the non-stereotyped condition compared to the stereotyped condition, which aligns with previous research on stereotype threat (for instance, Steele & Aronson, 1995). In addition, the results provide further evidence that stereotype threat can be elicited even from subtle cues – in this case, from an ad that did not feature, or even mention, women. Moreover, in the non-stereotyped condition, female respondents' performance was higher than those of male participants in either condition. Although beyond the scope of this study, this might indicate that non-stereotyped occupational gender portrayals not only counteract stereotype threat, but potentially leads to an additional increase in performance as well, if crafted in an empowering femvertising-manner. H3a was nonetheless fully supported.

On the contrary, H3b was not supported, as men's actual performance was not significantly lower in the non-stereotyped condition. Possibly, the stereotyped treatment was too subtle to induce any stereotype lift on men; seeing that men are typically less susceptible to gender-related stereotype threat cues (Pillaud et al., 2018; Tellhed & Adolfsson, 2018), it is possible that more explicit cues also are required to produce any stereotype lift. Although

female respondents were primed to think of their gender and expected inferiority in a variety of fields after having seen a man portrayed within a male-dominated field, men were likely not primed in the same way to think of women. Subsequently, there could be no downward comparison with women, which in turn prevented the occurrence of stereotype lift.

5.3 Estimated Performance

On average, both men and women overrated themselves in relation to their actual performance, which is in line with previous research on the Dunning-Kruger effect regarding performers in the lower actual performance percentiles (see for instance Kruger & Dunning, 1999), but partially contradicts previous research on the gender confidence gap which suggests that women underestimate or accurately estimate themselves (Herbst, 2020). The results showed a pattern that was in line with the hypothesized outcome for H4a, H5a, and H6a, but none of the differences were significant. In other words, no significant increase in women's estimated performance could be seen in the non-stereotyped condition; neither overall, nor in terms of self-estimations or estimated performance relative to others. In terms of the Dunning-Kruger effect (Figure 1-3), the magnitude remained rather consistent for females in both conditions when considering self-estimates. For estimated performance relative to others, however, both top-performing and bottom-performing women seemed to produce higher estimations when exposed to the non-stereotyped ad. We further theorized that the non-stereotyped ad would produce more accurate performance estimates; on average, this also proved to be correct (Figure 4). However, seeing that women also exhibited the Dunning-Kruger effect, higher accuracy as a result of increased performance estimates was naturally only true for top-performers when separating females by actual performance. The increase in performance estimates also for bottom-performers may indicate that femvertising or non-stereotyped occupational gender portrayals could have a confidence-boosting (but sometimes misleading) effect – leading bottom-performing women to estimate their performance in similar terms as men (43rd-63rd percentile when actual performance $\leq$ 25 percent).

A possible explanation for the non-significant results relates to women's actual performance increase in the non-stereotyped condition. Contrary to H4a, H5a and H6a, women in the non-stereotyped condition did *not* rank themselves (significantly) higher, but rather more

accurately in relation to their actual performance. This finding may be explained by the same mechanisms as presented in the hypothesis – given that the Dunning-Kruger effect is caused by heuristics, the non-stereotyped ad may have led to increased information processing, ultimately resulting in women producing performance estimates that better correspond with their actual performance. It could also be argued that the non-stereotyped ad may have counteracted stereotype threat, similar to how it is hypothesized to have affected actual performance. More specifically, the non-stereotyped ad may have led to women producing more accurate performance estimates through mitigating the diversion of focus that stereotype threat may otherwise impose, in line with previous findings by, for instance, Steele and Aronson (1995).

Consistent with the predictions of H4b, H5b, and H6b, males showed significantly lower performance estimates in the non-stereotyped (versus the stereotyped) condition in absolute terms. As there were no significant differences in actual performance for men between the two conditions, it can be concluded that the non-stereotyped occupational gender portrayal does indeed impact performance assessments, leading men to judge themselves more accurately (i.e., closer to the actual average result of around 50 percent). In terms of the Dunning-Kruger effect, as seen in Figure 2, bottom-performing men (actual performance $\leq$ 25 percent) showed a remarkable drop in self-estimations in the non-stereotyped condition. All in all, these findings support our hypothesis and may well be explained by the notion that non-stereotyped occupational gender portrayals increase information processing which leads to more accurate performance estimates. This was especially apparent in estimated performance relative to others, where men who were exposed to the non-stereotyped treatment consistently ranked themselves 10-12 percentage points lower than those exposed to the stereotyped treatment across performance percentiles. The lower ratings in the non-stereotyped condition support our hypothesis, but only resulted in more accurate estimates for bottom-performing men, whereas top-performing men were led to underrate their relative performance further – causing them to estimate their performance in similar terms as women in the stereotyped condition (40th-50th percentile, when actual performance $\geq$ 75 percent). Similarly to women, however, men overall rated themselves more accurately in the non-stereotyped condition (Figure 4).

In other words, our findings are in line with previous research by Tellhed and Adolfsson (2018) on how stereotype threat can confuse performance assessments, but also add an

additional aspect to their study in regards to men’s assessments. We find that even when compared to a condition that does not elicit any stereotype threat to them, men too increase their accuracy in a non-stereotyped condition. Therefore, there must be an additional mechanism to explain the increased accuracy in performance estimations for both genders, which is why we argue that the results could in part be explained by the previously discussed breach of the Dunning-Kruger effect.

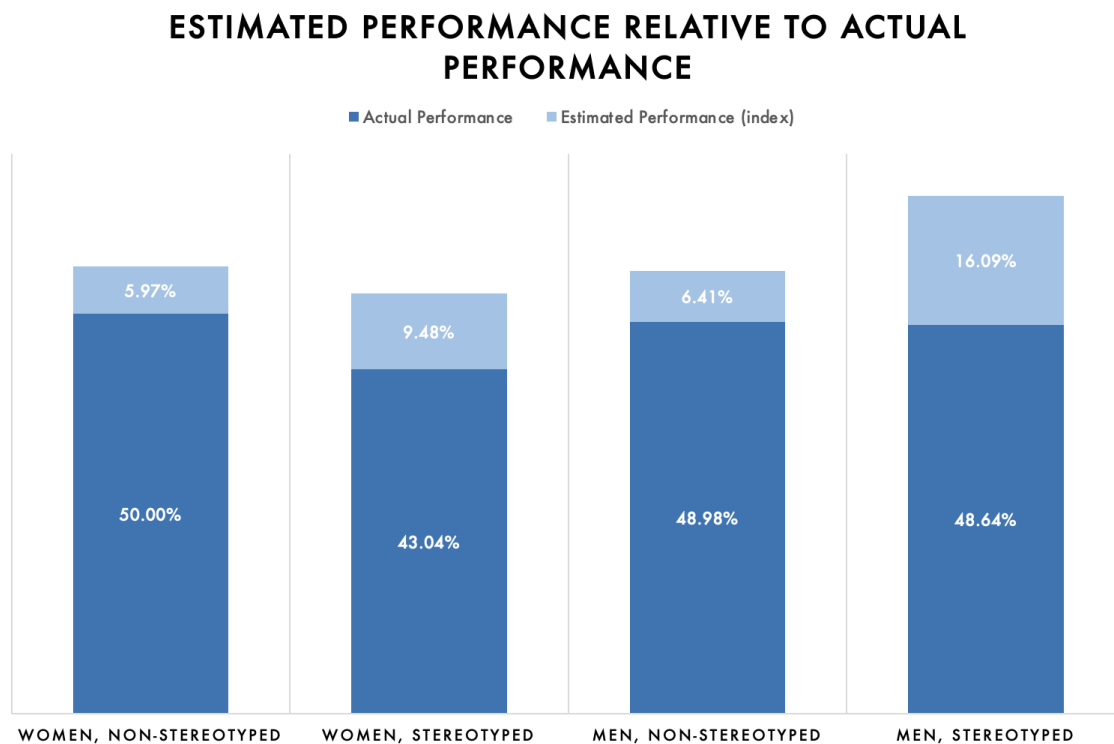


Figure 4: Estimated Performance Relative to Actual Performance

By comparing estimated performance to actual performance (Figure 4) divided by condition and gender, another interesting observation can be made – namely that the non-stereotyped treatment seems to result in women producing performance estimates that are more equal to those of men. In the stereotyped condition, men overrated themselves by 16.09 percentage points, compared to 9.48 percentage points for women. However, in the non-stereotyped condition, men instead overrated themselves by only 6.41 percentage points; and seeing that women’s actual performance increased in this condition, their de facto overestimation dropped to 5.97 percentage points. Although the drop in estimated performance differed in magnitude for men and women in the two conditions, their performance estimates overall were close to identical in the non-stereotyped condition. Our findings therefore add a dimension to the findings by Liljedal et al. (2020), and yet another dimension to those of

Tellhed and Adolfsson (2018) – namely that men and women’s performance estimates both become more accurate *and* more equal in the non-stereotyped condition.

Another interesting notion when it comes to the variable *estimated performance relative to others*, is that both genders were fairly modest in their estimations, and consistently underestimated themselves in the non-stereotyped condition relative to others. This indicates that the major effects in terms of the Dunning-Kruger effect in our sample, representative of the Swedish population, were attributable to erroneous self-estimates. Although beyond the scope of this study, research has shown that cultural differences may exist in the better-than-average effect (which is a part of the metric *estimated performance relative to others*). For instance, Americans tend to display the better-than-average effect more than Koreans (Lee, 2012). These cross-cultural differences have also been found to be mediated by a modesty norm that, in this case, exists in Eastern culture which prescribes downplaying one’s own accomplishments and showing moderation in self-presentation (ibid). Although research on this topic has not yet been conducted in a Swedish or Scandinavian context, our results show a tendency of a similar modesty norm, suggesting that Swedes tend to estimate their performance as either worse (non-stereotyped condition) or just slightly better compared to peers (stereotyped condition).

5.4 Likelihood to Ask for Help

H7a hypothesized that women would have lower likelihood to ask for help in the non-stereotyped condition, whereas H7b hypothesized that men would have higher likelihood to ask for help in the same condition. The majority of the results showed a pattern that was similar to the one hypothesized for each gender. However, most differences were non-significant, which may well be due to the survey design – the questions pertaining to H7 were included at the end of the questionnaire, meaning that they were the furthest away from the treatment and the immediate effects might have started to wear off when respondents arrived at this section. It should also be noted that the treatments in this study were rather subtle in their stereotyped portrayals, in the sense that they did not display outright sexist or degrading gender role depictions. All in all, these two factors may have contributed to reducing the effects on the genders’ likelihood to ask for help.

More specifically, the results on five out of six questions for H7a indicated a slightly lower likelihood for women to ask for help in the non-stereotyped condition, although all of the results were non-significant. The results for H7b were similar; five out of six questions followed a pattern similar to the one hypothesized, but the results were only significant in terms of likelihood to ask for help in removing a severe stain from a piece of clothing. Overall, the effects were similar for men and women across domains; in other words, men exhibited a *higher* likelihood to ask for help under the non-stereotyped condition regardless if the domain was coded as feminine or masculine. Likewise, women exhibited a *lower* likelihood to ask for help in the same condition, regardless of domain. The observed patterns indicate increased information processing as a result of the non-stereotyped occupational gender portrayal, leading to a more accurate assessment of one's need for help – adding to the findings by Liljedal et al. (2020). An alternative or parallel mechanism could be that the stereotyped ad may evoke a form of stereotype threat and lift; reminding women of their expected inferiority, and men of their expected superiority across domains – in line with previous research in the area by, inter alia, Steele and Aronson (1995). However, as the results were overall non-significant, no statistically robust conclusions regarding these hypotheses can be drawn.

6. Conclusions

Our findings suggest that non-stereotyped occupational gender portrayals lead to increased advertisement attitudes for women, whilst no significant differences between the stereotyped and non-stereotyped ads were found in advertisement attitudes for men. For brand attitudes, no significant differences were found for either gender. Additionally, we find that the non-stereotyped ad led to increased actual performance for women, but had no such effect on men. We conclude that the non-stereotyped ad leads to lower, and in other words more accurate, performance estimates for men both overall, as well as in terms of self-estimates and estimated performance relative to others. However, no significant differences were found in this variable for women. When comparing performance estimates to actual performance in both conditions, we find that the non-stereotyped ad leads to a more even performance for men and women, whilst at the same time causing both genders to estimate their performance more accurately (and as a result, more equal to each other). We also conclude that overall, the patterns shown in the results for likelihood to ask for help indicate that men display a higher

likelihood to ask for help when exposed to the non-stereotyped ad, and women show a lower likelihood to ask for help when exposed to the same ad. Nonetheless, these findings were overall statistically unsupported, and no conclusions can therefore be drawn with regards to these variables. All in all, the study provides evidence for non-stereotyped occupational gender portrayals in advertising having positive effects on certain consumer responses, as well as on the ability to perform cognitive tasks and form accurate performance assessments. In other words, we conclude that the non-stereotyped ad led to women becoming more skilled, and both genders becoming more aware of their own skills and abilities.

7. Implications

For practitioners, our findings suggest that portraying women in male-dominated occupations leads to increased advertisement attitudes for women, whilst at the same time not significantly jeopardizing brand or advertisement attitudes for men, nor brand attitudes for women. The ads used in this experiment were from the Swedish Armed Forces, which impacts the ability to generalize our results from a theoretical perspective. However, from a managerial perspective, the results have important implications since the brand behind an ad – especially in the case of public authorities – is usually known from a customer viewpoint. Our results therefore contribute to managers by testing the effects on different variables for a brand that large parts of the population have a prior relation to, and proves that non-stereotyped advertising can have positive effects also for a (real) known brand.

In the specific case of the Swedish Armed Forces, it has been noted that after posting ads with females, their comment section on social media is usually met by comments (mostly by men) complaining about the excessive inclusion of women (see Appendix G for an excerpt). Although this suggests that men may be negatively impacted in terms of brand and advertisement attitudes after seeing an ad with a female soldier, our results show no such effects. Rather, our results imply that brands can break down occupational gender stereotypes while also increasing women's advertisement attitudes, which may lead to reduced barriers for women applying to otherwise male-dominated jobs, and vice versa.

Our results also imply that non-stereotyped occupational gender portrayals lead to women performing better; on a level that is equal to, or better than, that of men. Additionally,

non-stereotyped occupational gender portrayals seem to make both genders more accurate and equal in their performance assessments. In other words, our findings suggest that the Dunning-Kruger effect is reduced (in overall terms) when people are exposed to ads featuring women in otherwise male-dominated occupations. Knowing the extent of one's own skills and abilities can be an important component in making everyday life more manageable, but may also constitute an important factor in bigger life decisions such as which education to pursue, or what jobs to apply to. By leveling men and women's performance assessments and in turn closing the confidence gap somewhat, more equal premises can perhaps be achieved for both genders in professional and educational spheres. Moreover, granting women the opportunity to perform at their best by eliminating stereotype threat in a given situation will further benefit female individuals, and in extension also lead to a higher level of societal gender equality by allowing women to compete on the same terms as men. Our findings contribute to a better understanding of the reactions to non-stereotyped occupational gender portrayals in advertising, and we therefore argue that it can lead to a society that is more mindful in how to combat gender inequalities.

Beyond the positive societal effects, it should be in most marketers' interest to create advertising that does not perpetuate stereotypes due to the risk of offending or alienating potential consumers (Eisend, 2010). As cultural and societal views on gender and equality change, advertising will have to be adjusted to remain relevant. In the light of social media, consumers can now easily voice their dislike for a certain brand or marketing message in the public (Åkestam et al., 2021), which makes it all the more important for brands to use considerate and well-thought-through messages in their marketing. Our results indicate that even subtle stereotypes in the form of occupational gender portrayals can elicit negative effects – in other words, avoiding outright sexist or degrading gender role portrayals is not enough to avoid the negative effects found in this study. Rather, marketers will have to work actively to incorporate non-stereotyped and empowering gender portrayals in their ads going forward if they aim to be progressive.

8. Limitations and Suggestions for Future Research

This study only tested for the effects of a stereotyped ad compared to a non-stereotyped ad, which is why no control group (i.e., a part of the respondents that do not receive any

treatment) was included in this study. It could, however, be argued that including a control group in future studies could be of interest to assess whether the effects found in our study are a result of a decrease or increase in different variables, rather than assessing the effects of the ads only in relation to one another. For instance, we found that women's performance was slightly higher than men's in the non-stereotyped condition, and including a control group could add some clarity to whether women perform at "normal capacity" or at a higher level when exposed to an empowering female portrayal. Furthermore, our study only tested whether there was indeed a difference between the two conditions. Although we theorize that our results can be explained by previous research on ad reactance, stereotype threat/lift, and increased information processing, *inter alia*, a topic for future research would be to examine to what extent these mechanisms can explain the effects found in our study, and how they interact.

Contrary to most previous research on the subject, our study tested ads from a known brand to investigate the response in a more realistic setting and account for factors such as perceived brand congruence. However, as respondents' underlying attitudes and previous brand experiences with the Swedish Armed Forces cannot be assumed to hold for brands regardless of industry or brand image, the results of our study are generalizable only to a limited extent. A suggestion for future research would be to produce similar stimuli but for a variety of brands, to assess if the effects found in this study hold across other domains and brands.

Furthermore, *actual performance* in this study was assessed by randomly assigning each participant either a masculine-coded, feminine-coded or neutral-coded test. Unfortunately, the groups were too small to test for between-group differences in terms of the quizzes for the Dunning-Kruger effect ($n < 30$ for the groups). In future research, such a comparison would be interesting to investigate, especially since the gender confidence gap has been shown to be task-dependent. For instance, men tend to estimate their abilities as significantly higher than women in logical and quantitative domains, as well as for computer skills and technical abilities (Bennett, 1996; Bennett, 1997; Furnham, 2005; Papyrina et al., 2021), whereas women report higher confidence than males in interpersonal skills, organizational ability, and information literacy (Papyrina et al., 2021; Michalak et al., 2017). In addition, Brown (2012) has found that the better-than-average effect (which is part of the Dunning-Kruger effect), is stronger for attributes that are viewed as important to one's self-image. In other words, since

it could be argued that domains that are viewed as stereotypically masculine (feminine) are more important for males (females), the better-than-average effect should increase when a person is asked to judge their performance relative to others in a gender-congruent domain (Brown, 2012).

Moreover, our study investigated non-stereotyped occupational gender portrayals pertaining to women, why investigating whether these effects hold true for similar portrayals of men could be of value (e.g. a male nurse or primary school teacher). Finally, this study has looked at the effects of gender role portrayals within the scope of occupations; an interesting topic for future research would be whether other types of gender stereotypes produce similar effects, and to what extent.

9. References

- Adams, G., Garcia, D. M., Purdie-Vaughns, V., & Steele, C. M. (2006). The detrimental effects of a suggestion of sexism in an instruction situation. *Journal of Experimental Social Psychology, 42*(5), 602-615. doi:10.1016/j.jesp.2005.10.004
- Åkestam, N., Rosengren, S., & Dahlen, M. (2017). Advertising “like a girl”: Toward a better understanding of “femvertising” and its effects. *Psychology & Marketing, 34*(8), 795-806. doi:10.1002/mar.21023
- Åkestam, N., Rosengren, S., Dahlén, M., Liljedal, K. T., & Berg, H. (2021). Gender stereotypes in advertising have negative cross-gender effects. *European Journal of Marketing, ahead-of-print(-)* doi:10.1108/EJM-02-2019-0125
- Ang, R. P., Lim, K. M., & Tan, A. (2004). Effects of gender and sex role orientation on help-seeking attitudes. *Current Psychology, 23*(3), 203-214. doi:10.1007/s12144-004-1020-3
- Antioco, M., Smeesters, D., & le Boedec, A. (2012). Take your pick: Kate Moss or the girl next door? - the effectiveness of cosmetics advertising. *Journal of Advertising Research, 52*(1), 15-30. doi:10.2501/JAR-52-1-015-030
- Bennett, M. (1996). Men's and women's self-estimates of intelligence. *Journal of Social Psychology, 136*(3), 411-412. doi:10.1080/00224545.1996.9714021
- Bennett, M. (1997). Self-estimates of ability in men and women. *Journal of Social Psychology, 137*(4), 540-541. doi:10.1080/00224549709595475
- Bian, X., & Wang, K. (2015). Are size-zero female models always more effective than average-sized ones?: Depends on brand and self-esteem. *European Journal of Marketing, 49*(7-8), 1184-1206. doi:10.1108/EJM-08-2013-0414
- Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about gender. *American Economic Review, 109*(3), 739-773. doi:10.1257/aer.20170007
- Brown, J.D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin, 38*(2), 209-219. doi:10.1177/0146167211432763
- Corney, R. H. (1990). Sex differences in general practice attendance and help seeking for minor illness. *Journal of Psychosomatic Research, 34*(5), 525-534. doi:10.1016/0022-3999(90)90027-2

- Dahlen, M., Rosengren, S., & Smit, E. (2014). Why the marketer's view matters as much as the message: Speaking down to the consumer speaks badly to a brand's image. *Journal of Advertising Research*, 54(3), 304-312. doi:10.2501/JAR-54-3-304-312
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality & Social Psychology Bulletin*, 28(12), 1615-1628. doi:10.1177/014616702237644
- De Meulenaer, S., Dens, N., De Pelsmacker, P., & Eisend, M. (2018). How consumers' values influence responses to male and female gender role stereotyping in advertising. *International Journal of Advertising*, 37(6), 893-913. doi:10.1080/02650487.2017.1354657
- Deshayes, M., Zory, R., Seitchik, A. E., Chalabaev, A., & Clément-Guillotin, C. (2020). Can the stereotype threat and lift phenomenon be applicable to a muscular endurance task? *Research Quarterly for Exercise and Sport*, 91(2), 354-360. doi:10.1080/02701367.2019.1668518
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83-87. doi:10.1111/1467-8721.01235
- Eisend, M. (2010). A meta-analysis of gender roles in advertising. *Journal of the Academy of Marketing Science*, 38(4), 418-440. doi:10.1007/s11747-009-0181-x
- Feng, Y., Chen, H., & He, L. (2019). Consumer responses to femvertising: A data-mining case of Dove's "campaign for real beauty" on YouTube. *Journal of Advertising*, 48(3), 292-301. doi:10.1080/00913367.2019.1602858
- Försvarsmakten. (2016). *Kvinnan och försvaret*. Retrieved from <https://www.forsvarsmakten.se/sv/information-och-fakta/var-historia/artiklar/kvinnan-i-forsvaret/>. Date Retrieved: May 10, 2021.
- Försvarsmakten. (2021a). *Förtroendet för försvarsmakten i topp bland myndigheter*. Retrieved from <https://www.forsvarsmakten.se/sv/aktuellt/2021/03/fortroendet-for-forsvarsmakten-i-topp-bland-myndigheter/>. Date Retrieved: May 10, 2021.
- Försvarsmakten. (2021b). *Personalsiffror*. Retrieved from <https://www.forsvarsmakten.se/sv/organisation/om-var-organisation/personalsiffror/>. Date Retrieved: May 10, 2021.
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4), 25-42. doi:10.1257/089533005775196732
- Furnham, A. (2005). Gender and personality differences in self and other ratings of business

- intelligence. *British Journal of Management*, 16(2), 91-103.
doi:10.1111/j.1467-8551.2005.00434.x
- Goodie, A., & Fortune, E. (2013). Measuring cognitive distortions in pathological gambling: Review and meta-analyses. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, 27(3), 730-743. doi:10.1037/a0031892
- Grau, S. L., & Zotos, Y. C. (2016). Gender stereotypes in advertising: A review of current research. *International Journal of Advertising*, 35(5), 761-770.
doi:10.1080/02650487.2016.1203556
- Grinstein-Weiss, M., Fishman, G., & Eisikovits, Z. (2005). Gender and ethnic differences in formal and informal help seeking among israeli adolescents. *Journal of Adolescence*, 28(6), 765-779. doi:10.1016/j.adolescence.2005.01.002
- Herbst, T. H. H. (2020). Gender differences in self-perception accuracy: The confidence gap and women leaders' underrepresentation in academia. *SA Journal of Industrial Psychology*, 46 doi:10.4102/sajip.v46i0.1704
- Hively, K., & El-Alayli, A. (2014). "You throw like a girl:" The effect of stereotype threat on women's athletic performance and gender stereotypes. *Psychology of Sport and Exercise*, 15(1), 48-55. doi:/10.1016/j.psychsport.2013.09.001
- Holbrook, M. B. (1987). Mirror, mirror, on the wall, what's unfair in the reflections on advertising? *Journal of Marketing*, 51(3), 95-103. doi:10.1177/002224298705100307
- Hoyt, C. L., & Murphy, S. E. (2016). Managing to clear the air: Stereotype threat, women, and leadership. *The Leadership Quarterly*, 27(3), 387-399.
doi:10.1016/j.leaqua.2015.11.002
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11(5), 365-371. doi:10.1111/1467-9280.00272
- Iyengar, S. (2011). *Sheena Iyengar: How to make choosing easier*. Retrieved from https://www.ted.com/talks/sheena_iyengar_choosing_what_to_choose/transcript?language=en.
Date Retrieved: May 10, 2021.
- Johnson, S. K., Murphy, S. E., Zewdie, S., & Reichard, R. J. (2008). The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes*, 106(1), 39-60.
doi:10.1016/j.obhdp.2007.12.002
- Juvrud, J., & Rennels, J. L. (2017). "I don't need help": Gender differences in how gender

- stereotypes predict help-seeking. *Sex Roles*, 76(1), 27-39.
doi.org/10.1007/s11199-016-0653-7
- Keohane, A., & Richardson, N. (2018). Negotiating gender norms to support men in psychological distress. *American Journal of Men's Health*, 12(1), 160-171.
doi.org/10.1177/1557988317733093
- Krockow, E. M. (2018, September 27). How many decisions do we make each day? . *Psychology Today*, Retrieved from <https://www.psychologytoday.com/us/blog/stretching-theory/201809/how-many-decisions-do-we-make-each-day>. Date Retrieved: May 10, 2021.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
doi.org/10.1037/0022-3514.77.6.1121
- Kruger, J., & Dunning, D. (2002). Unskilled and unaware - but why? A reply to Krueger and Mueller (2002). *Journal of Personality and Social Psychology*, 82(2), 189-192.
doi:10.1037/0022-3514.82.2.189
- Lee, J. (2012). Trait desirability and cultural difference in the better-than-average effect. *Asian Journal of Social Psychology*, 15(4), 261-272.
doi.org/10.1111/j.1467-839x.2012.01381.x
- Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22(140), 1-55.
- Liljedal, K. T., Berg, H., & Dahlen, M. (2020). Effects of nonstereotyped occupational gender role portrayal in advertising: How showing women in male-stereotyped job roles sends positive signals about brands. *Journal of Advertising Research*, 60(2), 179-196. doi.org/10.2501/jar-2020-008
- Metcalf, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality and Social Psychology Review*, 2(2), 100-110.
doi.org/10.1207/s15327957pspr0202_3
- Michalak, R., Rysavy, M. D. T., & Wessel, A. (2017). Students' perceptions of their information literacy skills: The confidence gap between male and female international graduate students. *The Journal of Academic Librarianship*, 43(2), 100-104.
doi:10.1016/j.acalib.2017.02.003
- Orth, U. R., & Holancova, D. (2003). Consumer response to sex role portrayals in advertisements: Effects of incongruity and prejudices on emotions and attitudes. *Journal of Advertising*, 32(4), 77-89. doi:10.1080/00913367.2003.10639144

- Papyrina, V., Strebel, J., & Robertson, B. (2021). The student confidence gap: Gender differences in job skill self-efficacy. *Journal of Education for Business*, 96(2), 89-98. doi.org/10.1080/08832323.2020.1757593
- Pederson, E., & Vogel, D. (2007). Male gender role conflict and willingness to seek counseling: Testing a mediation model on college-aged men. *Journal of Counseling Psychology*, 54(4), 373-384. doi:10.1037/0022-0167.54.4.373
- Peter, C., & Ponzi, M. (2018). The risk of omitting warmth or competence information in ads: Advertising strategies for hedonic and utilitarian brand types. *Journal of Advertising Research*, 58(4), 423-432. doi:10.2501/JAR-2018-005
- Pillaud, V., Rigaud, D., & Clémence, A. (2015). The influence of chronic and situational social status on stereotype susceptibility. *PloS One*, 10(12), e0144582. doi:10.1371/journal.pone.0144582
- Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, 29(5), 453-469. doi:10.1002/bdm.1883
- Pruysers, S., & Blais, J. (2014). Anything women can do men can do better: An experiment examining the effects of stereotype threat on political knowledge and efficacy. *The Social Science Journal*, 51(3), 341-349. doi:10.1016/j.soscij.2014.05.005
- Rice, L., & Barth, J. M. (2017). A tale of two gender roles: The effects of implicit and explicit gender role traditionalism and occupational stereotype on hiring decisions. *Gender Issues*, 34(1), 86-102. doi:10.1007/s12147-016-9175-4
- Schlosser, T., Dunning, D., Johnson, K. L., & Kruger, J. (2013). How unaware are the unskilled? empirical tests of the “signal extraction” counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, 39, 85-100. doi:10.1016/j.joep.2013.07.004
- Schuster, C., & Martiny, S. E. (2017). Not feeling good in STEM: Effects of stereotype activation and anticipated affect on women’s career aspirations. *Sex Roles*, 76(1), 40-55. doi:10.1007/s11199-016-0665-3
- Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, 134(2), 207-222. doi:10.1037/0033-2909.134.2.207
- Shea, M., Wong, Y. J., Nguyen, K. K., & Baghdasarian, S. (2017). College women’s subjective femininity stress, gender solidarity, and psychological help-seeking intentions: *The Counseling Psychologist*, 45(3), 438-461. doi:10.1177/0011000017701820

- Simons, D. J. (2013). Unskilled and optimistic: Overconfident predictions despite calibrated knowledge of relative skill. *Psychonomic Bulletin and Review*, 20(3), 601-607. doi:10.3758/s13423-013-0379-2
- Söderlund, M. (2018). *Experiments in Marketing*. Lund: Studentlitteratur.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1), 4-28. doi:10.1006/jesp.1998.1373
- Stankov, L., Lee, J., & Paek, I. (2009). Realism of confidence judgments. *European Journal of Psychological Assessment*, 25(2), 123-130. doi:10.1027/1015-5759.25.2.123
- Steele, C., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology*, 69(5), 797-811. doi:10.1037/0022-3514.69.5.797
- Tang, M. O. T., Oliffe, J. L., Galdas, P. M., Phinney, A., & Han, C. S. (2014). College men's depression-related help-seeking: A gender analysis. *Journal of Mental Health*, 23(5), 219-224. doi:10.3109/09638237.2014.910639
- Tellhed, U., & Adolfsson, C. (2018). Competence and confusion: How stereotype threat can make you a bad judge of your competence. *European Journal of Social Psychology*, 48(2), 189-197. doi:10.1002/ejsp.2307
- Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131. doi:10.1126/science.185.4157.1124
- Unstereotype Alliance. (2021). *About the unstereotype alliance*. Retrieved from <https://www.unstereotypealliance.org/en/about>. Date Retrieved: May 10, 2021.
- Vajapey, S. P., Weber, K. L., & Samora, J. B. (2020). Confidence gap between men and women in medicine: A systematic review. *Current Orthopaedic Practice*, 31(5), 494-502. doi:10.1097/BCO.0000000000000906
- Vogel, D. L. (2014). Referring men to seek help: The influence of gender role conflict and stigma. *Psychology of Men & Masculinity*, 15(1), 1-8. doi:10.1037/a0031761
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39(5), 456-467. doi:10.1016/S0022-1031(03)00019-2

10. Appendices

Appendix A: Stimuli



Mock advertisements created to resemble authentic advertising from the Swedish Armed Forces. The images are stock photos retrieved from the [Swedish Armed Forces' database](#).

Appendix B: Results of Pre-studies

N.B.: The below questions are translated from Swedish

Question	Perceived Advantage: Men (0), women (10)	% correct answers
In a football game: Which one within their team is Player 1 allowed to pass the ball to? 	1.87	38.10%
Which of the following is not a typical time for penalty in ice hockey?	2.07	36.67%
What type of screw head is a Phillips-screwdriver used for?	2.52	57.14%
What does USB stand for?	2.74	61.90%
What order should be followed when jump starting a car?	2.77	43.33%
When investors talk about a "bear market", what do they mean?	3.13	56.67%
What does it mean to short-sell a stock?	3.50	56.67%
What is an HDMI cable used for?	3.90	60.00%
Person A has 16 shares in Company X, and each share is worth 12 SEK. Company X then announces a stock split 4-to-1. What happens to Person A's shares?	3.96	66.67%
What combination of car lights and headlights are permitted to use in traffic under Swedish regulation?	4.70	28.57%
What does it mean to tack?	8.04	57.14%

Which of the following products should <i>not</i> be used for removal of blood stains?	7.90	53.33%
What do the following laundry symbols mean? 	7.10	50.00%
Which of the following garments can be line-dried?	7.07	53.33%
What color is associated with Amethyst?	6.91	57.14%
What ingredient is typically not included in traditional Swedish pancakes?	6.53	76.67%
What do aquarelle, acrylic and gouache all have in common?	6.48	38.10%
What cooking time is stated on the package of Kungsörmen's Idealmakaroner?	5.91	80.95%
What characteristics besides weight is normally used to determine the quality of a diamond?	5.78	71.43%
What movement did the artist Pablo Picasso belong to?	5.53	50.00%

Appendix C: Quiz Questions

N.B.: The questions are translated from Swedish

The **neutral quiz** consisted of the following questions (correct answers in **bold**):

1. A ball and a ball pump cost 110 SEK together. The ball pump costs 100 SEK more than the ball. How much does the ball cost? (5, 10, **15**)
2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (5, 10, **100**)
3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (5, 24, **47**)
4. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are there in the class? (15, **29**, 30)

The **masculine-coded** quiz consisted of the following questions (correct answers in **bold**):

1. *In a football game: Which one within their team is Player 1 allowed to pass the ball to? (Appendix B for picture); (**Only player 3 as player 2 is offside**; None, as both player 2 and 3 are offside; Both, since all three players are within the opponent's penalty area, so no one is offside)*
2. *What type of screw head is a Phillips-screwdriver used for? (**Cross Slot**; Slotted; Torx)*
3. *What order should be followed when jump starting a car? (**Plus to plus and minus to earthed point on the car**; Minus to minus and plus to earthed point on the car; Plus to minus and minus to earthed point on the car)*
4. *When investors talk about a "bear market", what do they mean? (**When the stock market is open and stock prices face a steep decline**; When the stock market is open and the stock prices are on the rise; When investors' confidence in the stock market is low)*

The **feminine-coded** quiz consisted of the following questions (correct answer in **bold**):

1. *Which of the following products should not be used for removal of blood stains? (**Hot water**; Cold water; Salt)*
2. *What do the following laundry symbols mean? (Picture in Appendix B); (Do not tumble dry, do not dry clean, cold iron; **Do not tumble dry, do not bleach, cold iron**; Do not dry clean, do not bleach, hot iron)*
3. *What does aquarelle, acrylic and gouache all have in common? (**They are all water-based colors**; They are all in fluid form; They are all commonly used on canvas)*
4. *What does it mean to tack? (**To sew together two pieces of fabric with long loose stitches**; To sew together two pieces of fabric using zig-zag stitches; To knit using loose stitches)*

Appendix D: Distribution and Performance Across Quizzes

	Females	Males	Sig.
Feminine Questions	M = 55.51** SD = 28.64 n=59	M = 45.00** SD = 25.76 n=30	t = -1.69 p = 0.04
Neutral Questions	M = 33.33 SD = 35.85 n=48	M = 37.50 SD = 31.90 n=36	t = 0.55 p = 0.29
Masculine Questions	M = 48.40*** SD = 24.12 n=47	M = 62.50*** SD = 21.55 n=38	t = 2.81 p < 0.01
Actual Performance (aggregate, overall)	M = 46.43 SD = 31.11	M = 48.80 SD = 28.60	t = 0.62 p = 0.27
* p < 0.10, ** p < 0.05, *** p < 0.01			

Appendix E: Scenarios for H7

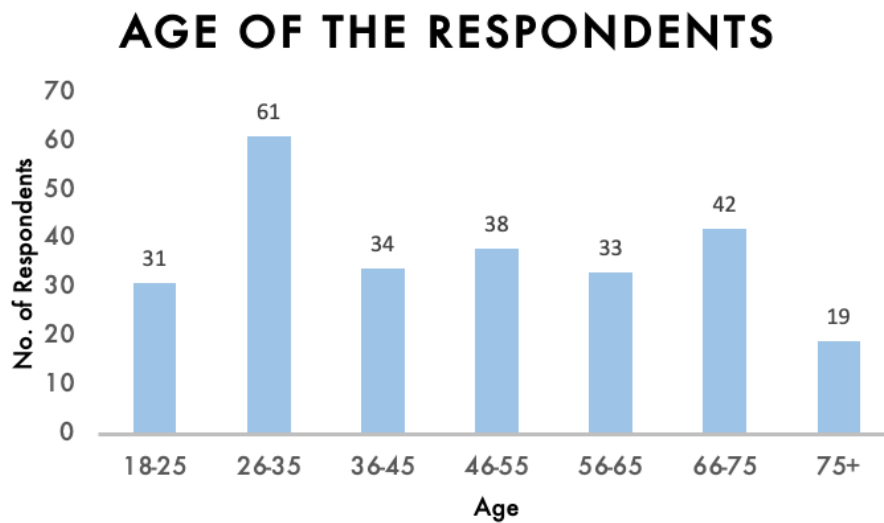
N.B.: The question and scenarios are translated from Swedish

How likely is it that you would ask for help with the following?

To ask for help could for instance mean to ask a relative for advice or employ a professional.

- ... jump start a car
- ... connect a laptop to a large screen display
- ... invest on the stock market
- ... remove a severe stain from a piece of clothing
- ... take up a pair of pants that are too long
- ... wash a down jacket

Appendix F: Age Distribution of Sample in Main Study

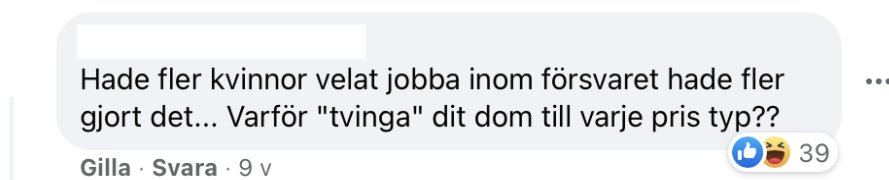


Appendix G: Excerpt of Facebook Comments

Source: The Swedish Armed Forces' Facebook post on the 8th of March 2021. Date Retrieved: May 15, 2021.



Translation: "If there's a war..and the childrens' mom and dad participate...orphans then"



Translation: "If more women had wanted to work in the armed forces then wouldn't they have done so... Why "force" them there at like every cost??"

There's no way this chick can carry all that equipment for longer than half a KM...

Gilla · Svara · 9 v

Militären ska INTE fokusera på kön, sexuell läggning, hudfärg, vad sägs om att Militären endast gör det som militären ska göra och ignorerar allt PK skit?

Gilla · Svara · 9 v · Redigerad



↪ 11 svar

Translation: "The Military should NOT focus on gender, sexual orientation, skin color, how about the Military only does what the military are supposed to and ignore all political correctness shit?"

Om kön inte spelar någon roll varför bryr ni er då om ifall det är "för få kvinnor"?

Det hela är såklart ologiskt och orsaken är att er ambition styra av ideologi, inte målet (att försvara Sverige på bästa möjliga sätt).

Jag inser att detta stärker er position hos landets regerande politiker men det gör det inte hos svenska folket i gemen.

Gilla · Svara · 9 v · Redigerad



↪ 42 svar

Translation: "If gender doesn't matter then why do you care whether there are "too few women"? The whole thing is of course illogical and the reason is that your ambition is built on ideology, not the objective (to defend Sweden in the best possible way). I realize that this strengthens your position with the country's politicians but it doesn't with the Swedish people overall."