

Not only blonde hair: possible effects of different styles of make-up on gender-science stereotype

Abstract

Women are affected from the so-called Gender Science Stereotype (GSS) according to which they are always considered as less competent than men in scientific fields for the only reason of being women, hence being obstructed and discriminated while pursuing that career. The scope of this study is to test whether some well-known stereotypes, such as the ‘‘dumb blonde’’ (i.e., blonde women are considered dumb or naïve), could explicitly affect the GSS and how they interact with each other. To investigate this, we manipulated pictures of a woman (Exp.1) and a man (Exp.2), by modifying (i) hair colour; (ii) hair length; (iii) glasses and (iv) make-up (Exp.1) or beard (Exp.2), obtaining in this way 36 pictures for each experiment (3x3x2x2). Each picture was presented twice, once associated with a scientific text and once with a novel. Here we are presenting preliminary data, where participants (n=45) were asked to evaluate, on a 10-points Likert scale, the authorship’s likelihood. Results show a negative effect of blonde hair (‘‘dumb blonde’’) for both genders and heavy make-up for women, as well as a positive effect of glasses for both genders and of beard for men. Interestingly, although no gender-science stereotype has been found for women (Exp.1), this emerged in Exp.2. The interaction between scientific genre and make-up for women show that this stereotype is present when enhanced by other factors, such as the ‘‘highly’’ feminising effect of make-up. These results are discussed considering both perceptual and social theories.

Keywords: *Social perception, Gender-science stereotype, Make-up, Blonde hair.*

Introduction

According to the gender-science stereotype (GSS) men are more competent in technology and science as compared to women, who in turn are more competent in arts and humanities (Nosek, et al., 2002). This stereotype is particularly insidious, given that scientists are generally considered as more “intelligent” than the rest of the population (e.g. Dutton & Lynn, 2014; Bian et al., 2017; Diaz et al., 2020). However, recent findings from studies conducted in the western culture suggest that this bias is no longer present, given that women are seen as equally competent as men: Eagly and colleagues (2019), in a meta-analysis integrated 16 nationally representative U.S. public opinion polls showing that belief in competence equality increased over time, thus reaching gender equality in competence in recent years. These recent findings are in apparent contradiction with the observation that women are under-represented in fields of science, technology, engineering, and mathematics (STEM), as well as in Technological Industry, where not only they are underrepresented (only 35% worldwide, according to the UNESCO Global Education Agenda of 2017) but also are promoted at a slower rate than men. Of course it could also be the case that these percentages are still, at least partially, influenced by stereotypization of the past years. Whatever is the case, we think that it is worth investigating whether gender-science stereotype are still present, albeit maybe in a covert way.

We believe that the gender-science stereotype, as well as gender stereotypes in general (such as women being considered as more communal and emotive than men, and men more aggressive and agentic, Prentice, 2002), are still present, albeit possibly weaker or masked by the social desirability bias (Edwards, 1953). According to the social desirability bias, people respond in a way that is considered socially acceptable, even if not representing their real opinion, to avoid negative judgments from others. Thus, given a massive education and social pressure towards gender equality in our western society, people respond in a way that fits social norms of acceptability, so as to fulfil the desirability of gender equality and do not go against social consensus.

To preliminary test this hypothesis, we focus on social perception studies. According to these studies, we form an impression, an immediate idea, of the individual we are interacting with (or simply facing at) based on their physical appearance. For instance, when we see a person with an angry expression we consider the person to be untrustworthy (whereas happiness is linked to trustworthiness; Todorov et al., 2008), meaning that we also infer personality traits through a simple glance at a face. In particular, when judging another individual’s face, we automatically extract information related to gender (O’Toole et al.,

1998), age (e.g. Slessor et al., 2010, Ciardo et al., 2014) social status (Ciardo et al., 2021), ethnicity (e.g. Losin et al., 2012, Capellini et al., 2016) and emotions (Ekman, 1992), besides the information regarding gaze direction (Actis-Grosso & Ricciardelli, 2017) which could modulate other automatic processing related to face perception, such as others' intentions and emotions (Frischen et al., 2007).

These pieces of information are beyond our cognitive control, meaning that they are "stimulus driven" and thus follow a bottom-up route (Mellers, 1990). Of course an interaction with more cognitive (top-down) variables is expected, which ultimately leads us to what could be called "first impression" of somebody else (Fiske et al., 1987). Each of these features (e.g. gender, age, ethnicity) could be subjected to some stereotypes or bias, which means that we build a prototypical image of a certain category to quickly access it, in future interactions with the same category and in doing this, we overgeneralise (Zebrowitz, 2012) some stereotypical aspects of the category to each individuals. Thus we can consider stereotypes as cognitive heuristics (Bodenhausen, 1985) that we deploy to simplify our daily effort to make sense of the world we live in and, in particular, of the people we meet. As social perceivers we have limited cognitive resources, time, and motivation to afford in-depth analyses of each and every situation. The other side of the coin is that these stereotypes are affecting the judgement of the others in a categorical and not veridical way already as a first impression. For example, in automatic inferences black faces presenting stereotypical black features are associated with criminality more than any other face (Eberhardt et al., 2004).

Focusing on the GSS, we think that some elements, described below, might enhance the activation of a gender bias. In particular, by presenting a face of (a) a female and (b) a male, in association (i.e. presented on top) with a written text of scientific (thus stereotypically associated to men) or literary (thus stereotypically associated to women) nature (i.e. a scientific abstract and a novel's plot), we think that a requested judgement on the likelihood of text's authorship in relation to the face visualized might be modulated by the presence or absence of some specific features on the presented face. If this is true, we could directly test whether GSS is still present and start investigating the factors that modulate this bias.

Regarding the elements that we think might modulate the activation of the GSS, a central role is played by those elements that could enhance the perceived femininity or masculinity, such as make-up for females and beard for males. In fact the use of make-up is associated with an increased femininity (e.g. Workman & Johnson, 1991: A female model wearing heavy cosmetics was rated as significantly more feminine than the model wearing no cosmetics), while the presence of a beard on a male face is associated with an increased

masculinity (Dixson, 2005). Thus, if, as we hypothesise, the GSS is still present (although possibly weaker), this could emerge in a study comparing female faces with or without make-up and male faces with or without beard.

This reasoning led us to a second consideration: studies on the relation between the use of cosmetics for women and related judgements on their features (as attractiveness, competence, morality etc.) usually (e.g. Bernard, 2020) compared models with no make-up with models with (a) a moderate amount or (b) a high amount of makeup. However, visual perception studies tell us that the perception of high vs. moderate makeup is highly influenced by the colour contrast (i.e. the same spectral stimulus can appear to be differently colored when viewed against different chromatic backgrounds, see for example Hering, 1864, Lotto & Purves, 1999, 2000) between both cosmetics colour (such as lipstick and eyeshadows) and the colour of skin, hair and eyes of the model. For this reason it is important to compare different colour combinations of at least the same makeup with different hair colours (thus leading to a different contrast). This idea, being based on visual perception studies (and never investigated before) requires also different hair-length to be introduced as variables, given that the percentage of blonde/black hair presented in the picture, influences the perceived contrast. The idea of comparing different hair colours is also aimed at testing the possible interaction between GSS (if present) and another well known stereotype which seems to be related with both gender and hair color: the dumb blonde effect (Weir, 1989) according to which women are perceived as more naïve or ultimately dumb, simply because of their hair colour (to the best of our knowledge this stereotype has never been tested with men). Regarding the GSS we think that the dumb blonde effect would facilitate the emergence of a difference between science and humanities for women but not for men. In other words we hypothesise that the difference between science and humanities subtends a difference in the perceived “intelligence” (i.e. women are stereotypically considered as less intelligent than men): if this is true blonde women, being considered as “dumber”, should also be considered as less likely to be a scientist.

Besides, we reasoned that our hypothesis could be better tested by introducing another variable that could reverse the effect of blond hair. For this reason we chose to also introduce faces wearing glasses, based on previous literature that shows that wearing glasses increases the perceived competence (Hellstrom, 1994). In this case the reasoning was the same as above, but reversed: if the person wearing glasses is considered more “intelligent”, then it could be that both male and female wearing glasses will be considered more competent (then the evaluated probability of authorship will be higher) but it is also possible that women

wearing glasses, being more ‘‘intelligent’’ would be considered more likely to be a scientist (compared to a novel writer).

We are aware that this hypothesis subtends different assumptions) but our aim is to preliminary test whether these assumptions could provide a fertile ground to pursue this path in face perception. In this perspective this study should be considered preliminary, as it is mainly based to test whether some perceptual feature of a face could modulate gender bias. Furthermore, with think that with this experimental paradigm it would be possible to test and possibly expand the study by Pinna & Deiana (2019), according to which a face can be considered as a mosaic of juxtaposed independent components, hence a whole that is equal to the sum of its parts.

We then decided to design two experiments in which the picture of a young woman in Experiment 1 was modified so to have three different versions for different use of make-up (i.e. no make-up, moderate and heavy), while in Experiment 2 the picture of a young man was modified so to have three different versions for different ‘‘styles’’ of beard (i.e. shaved, stubble, full beard). In this case our reasoning was related to the possible enhancing effect of gender-related features, where make-up is supposed to enhance stereotypical features related to women, and beard is supposed to enhance stereotypical features related to men. In both experiments we also decided to modify the hair colour (i.e. brown, black, blonde) with our reasoning based on the perceptual observation that different make-up has different perceptual contrast depending not only on the colour of the skin, but also on the colour of the hair. For this reason, we might surmise that the same ‘‘heavy’’ make-up could have a different moderating effect on blonde-haired as compared to black-haired. For the same reason (i.e., the amount of perceptual contrast) we presented modified versions of the pictures with (a) long- or (b) short-hair. Both male and female were compared with their plain version (i.e., no make-up for women and no beard for men). All pictures were also modified so as to have, for both female and male characters, a version wearing glasses.

Each picture was then associated once with a scientific text (an abstract of a scientific paper) and once with a literary text (the plot of a novel). The literary text was written by reinterpreting existing thriller/detective plots, as we wanted to avoid stereotypically feminine books such as romances and love stories. Participants were then asked to evaluate the likelihood of authorship of the text presented below every picture.

Thus both experiments underwent the same 3x3x2x2x2 experimental design: Make-up/Beard (3 levels, i.e. absent, light, heavy for make-up and shaved, stubble, full for beard) x

Hair Colour (3 levels, i.e. black, brown, blonde) x Hair Length (2 levels, i.e. short and long) x Glasses (2 levels, i.e. absent and present) x Genre (science and literature).

Summing up the aims of the present research, we want to investigate the presence of GSS in the Italian academic context, both for women and for men. We hypothesize that the social desirability bias could possibly cover the GSS for women, and for this reason we introduce some variables that could stereotypically enhance perceived femininity (i.e. make-up) and masculinity (i.e. beard). By doing this, we also hypothesize (and test) the role of colour contrast in enhancing the dumb blonde stereotype both in men and women.

Method

Ethical Statements

All participants gave a written informed consent before testing. The study was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki (World Medical Association, 2013) and fulfilled the ethical standard procedure recommended by the Italian Association of Psychology (AIP). The study was specifically approved by the local Ethics Committee of Milano-Bicocca University (protocol number RM-2022-547).

Experiment 1

In the first experiment (with only female faces stimuli) we expect, in relation to the GSS, a significant difference favouring humanities and arts, with a more probable association of the female gender to a novel's plot rather than a scientific abstract. We expect a heavy use of make-up to lower the perceived competence in both scientific and literary domain. We also expect the face wearing glasses to be associated with a higher probability to both texts. We then expect a major effect of the variable 'hair colour', with the face with blonde hair associated with a lower probability to the authorship of both texts (novel and abstract). Moreover, we expect an interaction between 'hair colour' and 'hair length' variables, with long hair accentuating the blonde colour, thus enhancing the 'dumb blonde' stereotype, lowering the overall probability of authorship.

Participants

45 participants (25 females) took part in the experiment, being either undergraduate students from the University of Milano-Bicocca who received course credits for their participation in the study or personal contacts of the experimenter recruited by word-of-

mouth. The participants' age ranged from 21 to 31 years of age ($M=25.6$, $SD=2.71$). Sample size was estimated through a power analysis, calculated using the statistical software G*Power (ver. 3.1.9.4) planning a repeated measures ANOVA with a small effect size ($f=.10$, $\alpha=.05$, $power=.80$).

All 45 participants had normal or corrected-to-normal vision and were unaware of the study's purpose. They signed the informed consent before partaking in the experiment. None of them withdrew and none of the data have been discarded.

Stimuli

Stimuli consisted of photographs of a female face modified by hair colour (brown, black, blonde), make-up (no make-up, heavy make-up, light make-up), hair length (short, long), glasses (present, absent). In Figure. 1, some examples of stimuli are reported. The original photograph has been taken from the database of the FaceApp application, which allows users to freely use them, and modified using the graphic tools present in the same software. Overall, from the interaction of the aforementioned variables we obtained 36 photographs ($3 \times 3 \times 2 \times 2$). Each photograph was then associated and presented once with an abstract of a scientific paper (adapted and translated from Benci & Fortunato, 1988) and once with a plot of a novel raising the total number of presented stimuli to 72. (The English translations of the two texts are reported in the appendix section).

Procedure

Participants were tested individually in a dimly lit room. They were sitting approximately 60 cm in front of the computer monitor (14 inches display, 1920x1080 pixels, 60Hz refresh rate) where the 72 trials were presented in a randomised order at the centre of the computer monitor with either the abstract or the novel's plot below (Fig. 2). Participants were asked to look at the photograph and, after reading the text below the photograph, to rate on a 10-point Likert scale the probability that the person depicted in the photograph was the author of the text. Responses were given verbally and recorded by the experimenter. After each response was given, participants moved to the next stimulus by pressing the spacebar. No time constraints were given. The total duration of the experiment was approximately 10 minutes.

Results

A preliminary t-test was run to compare results obtained from male and female participants. No differences were found; thus, we collapsed data from the whole pool of participants.

A Repeated Measures ANOVA (3x3x2x2x2) was run to test our hypotheses. In Table 1 all the significant main effects and interactions are reported, together with the effect size.

The three main effects were in line with expectations, with the presence of make-up lowering the probability of being judged as the authors of both texts, the presence of glasses increasing the same judged probability and the blond hair colour lowering the perceived competence.

In particular, regarding make-up, pictures of women wearing no make-up are always scored as the highest ($M=6.97$, $SE=0.124$), while light make-up received a lower score ($M=6.09$, $SE=0.151$) and heavy make-up the lowest score ($M=5.49$, $SE=0.167$), while regarding glasses, pictures of women wearing glasses are evaluated as more likely to be authors of a novel or of a scientific abstract ($M=6.57$, $SE=0.126$) as compared to pictures of women with no glasses ($M=5.79$, $SE=0.138$).

Regarding Hair Colour, pictures of women with black hairs are evaluated with the highest scores ($M=6.46$, $SE=0.132$), brown hairs receives lower scores ($M=6.33$, $SE=0.132$) while blonde hair gets the lowest scores ($M=5.75$, $SE=0.132$). However, post hoc comparison (Tukey correction) show no difference between black and brown hair colour ($p=0.240$).

The interaction Make-up X Genre ($F=48.76$, $p <.001$) is showed in figure 3, where it is possible to see how the effect of make-up is higher on Science as compared to Literature: wearing no make-up increases the scores for Science more than it does for Literature ($M=7.40$, $SE=0.138$ for science and $M=6.54$, $SE=0.131$ for literature, $p <.001$ with Tukey correction), while heavy make-up lowers the score for science ($M=5.11$, $SE=0.206$) more than it does for Literature ($M=5.87$, $SE=0.165$) with a Tukey correction $p <.001$.

The interaction between Make-up X Hair Colour (fig.4) is significant ($F=9.06$, $p <.001$), with brown-haired wearing no make-up getting the highest scores ($M= 7.39$, $SE=0.126$) while blonde women wearing heavy make-up getting the lowest scores ($M=5.25$, $SE=0.180$). Not wearing any make-up enhances the likelihood of authorship for all different hair colours.

This interaction could be better understood in light of the three-way interaction Make-Up X Hair Colour X Genre ($F=9.23$, $p <.001$) reported in Figure 5. As it can be seen, the effect of wearing no make-up is modulated by hair colour: for both genres blonde women

($M= 4.85$, $SE=0.213$ for Science and $M=5.66$, $SE=0.207$ for Literature) are judged as less competent (Tukey correction, $p <.001$) than both black- and brown-haired women, with no difference between these last two ($p=1.000$ and $p=0.984$, respectively for Science and Literature). As far as Science is concerned, this difference between hair colour is no more present when make-up is used, independently on whether it is light or heavy. Instead, for judgements on Literature, the presence of light make-up has no effect for black-haired women ($M= 6.73$, $SE=0.182$), but it influences the judgements for brown-haired ($M= 6.24$, $SE=0.184$), and blonde ($M= 5.37$, $SE=0.202$), thus causing a significant difference between the three hair colour ($p <.001$); on the contrary, heavy make-up lowers the perceived competence of black- ($M= 6.06$, $SE=0.171$) and brown-haired ($M=5.89$, $SE=0.12$), but unexpectedly increases the perceived competence of blonde women ($M= 5.66$, $SE=0.202$), thus resulting in a general equality between the three.

Another effect is between Hair Length and Hair Colour ($F=8.31$, $p <.001$, reported in figure 6), with long blonde hair receiving the lowest score ($M=5.61$, $SE=0.147$). Interestingly, only for blondes, short hair increases the scores ($M=5.89$, $SE=0.135$, Tukey correction $p <.001$). For black and brown-haired long or short hair makes little to no difference.

The interaction between Make-up X Glasses X Hair Colour ($F=7.85$, $p <.001$) in figure 7 shows that wearing glasses and no make-up is “the best combination” for all Colour conditions. Brown-haired wearing glasses and no make-up is the most plausible scientist or novelist with $M=7.63$, $SE=0.132$ against the $M=7.53$, $SE=0.146$ of black-haired and the $M=6.42$, $SE=0.156$ of blondes (with Tukey correction $p <.001$ for blonde while a $p=1.000$ for black and brown, showing no difference), once again being the least plausible authors. Interestingly, the negative effect of heavy make-up, which was high for both brown- and black-haired wearing glasses ($M=6.06$; $M=5.96$, respectively) was practically absent for blonde women, for which wearing glasses with the three different styles of make-up was judged as almost equally ($M=6.42$, $SE=0.156$; $M=6.19$, $SE=0.178$; $M=5.99$, $SE=0.192$ respectively no make-up, light and heavy).

The interaction Glasses X Genre ($F= 8.10$, $p =.007$) reflects the aforementioned result where the moderating effect of glasses on genre raises the average for both scientific and literary texts, with $M=5.70$, $SE=0.153$ for science when not wearing glasses against $M=6.68$, $SE=0.162$ when wearing glasses. Same improvement for Literature with $M=5.88$, $SE=0.148$ without glasses and $M=6.47$, $SE=0.127$ with glasses.

The higher order interactions between Make-up X Glasses X Hair Length X Hair Colour ($F=4.98, p <.001$) and Make-up X Glasses X Hair Colour X Genre ($F=9.15, p <.001$) could be easily understood by looking at the three levels interactions reported above.

Discussion

Paradoxically, the first interesting result of this study emerges from the absence of a main effect for Genre. This put in evidence the absence of peculiar differences in scientific or literary culture when associated with a woman, all else being equal. Hence, from these data we cannot confirm the GSS, intended as women being more likely associated with humanities and arts than sciences (cfr. Nosek, et al., 2002) and being instead in favour of recent findings (Eagly, et al., 2019), confirming that stereotypes could be flexible to changes in social roles. However the interaction between Make-up and Genre tells us that this effect still remains to some extent, considering that women wearing heavy make-up are more likely considered novel writers than scientists. This effect, if considered in light of the fact that the use of make-up enhances perceived femininity, is in favour of the hidden presence of GSS as a gender bias.

We found a main effect of ‘‘Make-up’’, which greatly affects the likelihood of authorships of both genres, with the dark make-up being associated with the lowest scores, while not wearing any seems the desirable choice when evaluating the likelihood, where the light make-up is between the previous two, confirming that wearing no make-up increases the perceived competence.

Consistently with the previous literature (Hellstrom, 1994) we obtained a main effect of ‘‘Glasses’’, increasing the overall scores when present.

Another expected result comes from the main effect of ‘‘Hair Colour’’, which is consistent with the ‘‘dumb blonde’’ stereotype, with blonde women receiving lower scores than brown- or black-haired women.

Differently from what expected, blonde women are not likely to be a novel writer as compared to other hair-coloured women (i.e. no interaction between hair colour and genre has been found), although hair colour presents some interesting interactions with other variables.

As for the interactions Make-up X Genre and Make-up X Hair-Colour X Genre it is notable to point out that, although the presence of make-up is detrimental in every situation, the association of dark make-up with blonde hair is the combination that leads to the lowest scores, especially with the scientific paper, while these negative effects are less severe for

literature. Not wearing any make-up highly enhances the likelihood of authorship, again, especially for the scientific text.

As mentioned before, we can see the positive effect of glasses in both genres, confirming pre-existing literature (Hellstrom, 1994), and, regarding the interaction between make-up, glasses and hair colour, we can see that make-up always affects and moderates the positive effect of the glasses. More interestingly, however, we can see how only when associated with blonde hair the positive effects of wearing glasses somehow cancel the negative effects of make-up: not wearing any make-up or wearing dark or light make-up makes little to no difference when glasses are present. One of the possible explanations is of perceptual nature: the high contrast of dark make-up for blonde women and that of the glasses (with a dark frame) are equally perceived and elicit the same level of colour contrast perception, levelling them out, with no notable positive or negative effect ultimately elicited. Being the first time that this “perceptual” hypothesis has been tested, this explanation should be furtherly tested with more experiments explicitly addressing it, also in light of the fact that stereotypes, being a very complex issue, are likely to be formed based on a variety of different characteristics.

Lastly, it is worthwhile mentioning another interaction which has a perceptual valence, which is the Hair Length X Hair Colour: while hair length does not make any meaningful difference between black and brown hair, it has a notable effect on blonde hair, with short hair on blonde individuals highly increasing the likelihood of authorship. We believe this can have a perceptual meaning, of course combined with the dumb blonde stereotype: shorter hair equals to less amount of perceived colour, hence to a lower activation of the dumb blonde effect (which remain present).

Experiment 2

In this second experiment with male face stimuli we expect, in relation to the gender-science stereotype, a significant difference favouring the scientific culture, with a more probable association of the male gender with the scientific abstract rather than the novel’s plot. However, if, as we hypothesize, the dumb blonde effect is gender related, this effect should not be present with male faces as stimuli. Despite current literature hasn’t investigated the “dumb blonde” stereotype on male population, we expect lower scores for blonde men on the probability of authorship of both texts (novel’s plot and abstract). We expect the face wearing glasses to be associated with a higher probability to both texts. As detailed in the Introduction section, we replaced the variable “make-up” with “beard” and we expect an

effect although we cannot predict in which direction: a short, groomed beard might be assimilated to the ‘heavy make-up’ variable (speculatively speaking, as in a more vain attitude, hence less competent) or on the contrary it could enhance the perceived masculinity and be considered as more probable to be a scientist, according to the gender-science stereotype.

Participants

The same 45 participants (25 females) from Experiment 1 took part in the experiment, being either undergraduate students from the University of Milano-Bicocca who received course credits for their participation in the study or personal contacts of the experimenter recruited by word of mouth. The participants’ age ranged from 21 to 31 years of age ($M=25.6$, $SD=2.71$). Sample size was estimated through a power analysis, calculated using the statistical software G*Power (ver. 3.1.9.4) planning a repeated measures ANOVA with a small effect size ($f=.10$, $\alpha=.05$, $power=.80$).

The order of presentation of the two experiments has been counterbalanced between subjects.

All 45 participants had normal or corrected-to-normal vision and were unaware of the study’s purpose. They signed the informed consent before partaking in the experiment. None of them withdrew and none of the data have been discarded.

Stimuli

Stimuli consisted of photographs of a male face modified by hair colour (brown, black, blonde), beard (shaved, stubble, full beard), hair length (short, long), glasses (present, absent). In Figure 8, some examples of stimuli are reported. The original photograph has been generated by ‘This Person Does Not Exist’, which is a random face generator based on an Artificial Intelligence freely available on the web (<https://thispersondoesnotexist.com>) and it was consequently modified using the graphic tools present in the FaceApp software. Overall, from the interaction of the aforementioned variables we obtained 36 photographs (3x3x2x2).

Each photograph was then associated and presented once with an abstract of a scientific paper (adapted and translated from Benci & Fortunato, 1988) and once with a made-up plot of a novel, raising the total number of presented stimuli to 72.

Procedure

Participants were tested individually in a dimly lit room. They were sitting approximately 60 cm in front of the computer monitor (14 inches display, 1920x1080 pixels,

60Hz refresh rate) where the 72 trials were presented in a randomised order at the centre of the computer monitor with either the abstract or the novel's plot below. Participants were asked to look at the photograph and, after reading the text below the photograph, to rate on a 10-point Likert scale the probability that the person depicted in the photograph was the author of the text. Responses were given verbally and recorded by the experimenter. After each response was given, participants moved to the next stimulus by pressing the spacebar. No time constraints were given. The total duration of the experiment was approximately 10 minutes.

Results

A preliminary t-test was run to compare results obtained from male and female participants. No differences were found; thus, we collapsed data from the whole pool of participants. A Repeated Measures ANOVA ($3 \times 3 \times 2 \times 2 \times 2$) was run to test our hypotheses. In Table 2 all the significant main effects and interactions are reported, together with the effect size.

The main effect of Beard shows that pictures of men without beard receive the lowest scores ($M=5.96$, $SE=0.177$), men with a stubble beard have overall lower scores ($M=6.33$, $SE=0.147$) and lastly full beard received the highest scores ($M=6.88$, $SE=0.125$).

The main effect of Glasses shows that men wearing glasses received significantly higher scores ($M=6.75$, $SE=0.124$) than men not wearing glasses ($M=6.04$, $SE=0.143$).

The main effect of Genre, tells us how Science receives higher scores ($M=6.64$, $SE=0.130$) than Literature ($M=6.14$, $SE=0.136$).

Lastly, the main effect of Hair Colour shows black hair associated with the highest scores ($M=6.78$, $SE=0.123$), followed by brown hair with lower scores ($M=6.39$, $SE=0.133$) and lastly with blonde hair receiving the lowest scores of all ($M=6.01$, $SE=0.140$).

In figure 9 we can see the interaction Beard X Genre ($F=5.03$, $p=.009$) with a full beard significantly increasing the authorship of scientific texts with $M=7.29$, $SE=0.143$ against $M=6.46$, $SE=0.155$ for the stubble and $M=6.18$, $SE=0.171$ for shaved men. With the Literature genre the result follows the same trend but with overall lower scores than science.

In figure 10 is reported the interaction between Beard and Hair Colour ($F=4.45$, $p=.002$) with black hair and a full beard being associated with higher probability of authorship ($M=7.31$, $SE=0.139$) independently on the text genre, against the $M=6.59$, $SE=0.156$ for black

hair and stubble, and $M=6.43$, $SE=0.187$ for black hair shaved beard. The same trend is present for brown hair as well, but blonde hair with shaved beard received the lowest scores ($M=5.42$, $SE=0.193$).

We can see the interaction between Hair Length X Hair Colour ($F= 11.35$, $p <.001$) where we can see that long blonde hair receive the lowest scores ($M=5.79$, $SE=0.152$) but at the same time blonde is the only colour that shows an increase with short hair ($M=6.23$, $SE=0.151$), while for other two colours there is almost no difference between long and short hair.

We can see, same as in Experiment 1, a strong effect of Glasses X Genre ($F=17.16$, $p <.001$), showing an increase in scores for both genres when wearing glasses.

Another result is shown in figure 11, with the interaction between Beard X Glasses X Hair Colour ($F=6.97$, $p <.001$). More specifically we see that black haired men with a full beard wearing glasses are the ones getting the highest scores with $M=7.66$, $SE=0.164$, against the fully shaved blonde men wearing no glasses ($M=5.18$, $SE=0.225$). Overall, wearing glasses raises every score within every interaction, with blonde hair and shaved beard being the lowest in every condition.

In the last graph (fig. 12) we reported the Glasses X Hair Colour X Genre ($F=5.53$, $p =.005$) interaction, showing that men wearing glasses with black hair are associated more to science ($M=7.52$, $SE=0.147$) than literature ($M=6.83$, $SE=0.139$), while blonde men with no glasses received the lowest scores for science ($M=5.63$, $SE=0.169$) and for literature ($M=5.73$, $SE=0.173$).

Discussion

With the results from Experiment 2 we can notice the presence of a main effect of ‘Genre’ here, which has not been found in Exp.1, where only female faces were shown. Men seem to be generally perceived as more likely to be scientists than novel writers, in line with the GSS. Moreover, the main effect of Beard tells us that a full beard has a strong impact in positively evaluating the likelihood of both genres, while not having any beard is the least optimal condition and wearing a stubble beard is between the two conditions. This result is in line with the study by Bakmazian (2014), showing an association between the presence of beard and perceived trustworthiness: here we found a similar association between beard and perceived competence. Furthermore the presence of a full beard has a higher impact on the

likelihood of science as compared to literature, meaning that wearing a beard (which is known as increasing the perceived masculinity, as reported by Dixson, 2005) enhances the GSS. Same as before, the main effect of ‘‘Glasses’’ confirms previous literature (Hellstrom, 1994) and, more interestingly, the main effect of ‘‘Hair Colour’’ confirms the dumb blonde effect to be present in men as well.

According to this result we are led to believe that the ‘‘dumb blonde’’ bias is not gender-related, differently from what we expected. As for the interactions we can see more in detail how ‘‘Beard’’ and ‘‘Genre’’ interact with one another. Although for both of them a full beard is the best solution to receive higher scores, we can notice that ‘‘Literature’’ is always lower than ‘‘Science’’, in line with the main effect of ‘‘Genre’’, but the increasing curve for science is steeper, meaning that a full beard increases the likelihood of being a scientist more than it increases the likelihood of being a novelist. We believe that this result shows the strength of the stereotypical scientist as being a man with a beard, enhancing this stereotype because of increasing masculine appearance. For ‘‘Hair Length’’ X ‘‘Colour’’ it’s important, as mentioned above, to notice how the results are almost identical for the two experiments, showing how a possible explanation is a strong perceptual effect of hair length for the dumb blonde stereotype activation, which we think is perceptually based.

General discussion

In conclusion, even if we did not find a GSS for women from Experiment 1, the effect was present for men in Experiment 2 and somehow present even in Exp.1. Indeed, it is worthwhile noticing that also in Exp.1 the gender-science stereotype persists when associated with certain features that carry strong biased characteristics and strongly feminine, such as make-up or blonde hair. Although it would be interesting to analyse and compare data from the two experiments, the power analysis ($f=.10$, $\alpha=.05$, $power=.80$) tells us that we need more participants ($n=82$) to run a mixed Anova on both experiments. For this reason, we limit our comparison, for the moment, to a qualitative analysis of the differences between the two experiments, keeping in mind that our discussion is corroborated by statistical analyses separately run for each experiment. It is worth reminding that existing literature usually addresses this issue by studying only female gender, thus the fact that we could compare two genders, at least qualitatively, is a plus and not a disadvantage. Furthermore, the presence of the GSS in Experiment 2 (and in Exp. 1 for blond women) substantiate our reasoning.

Women are generally considered less competent than men in science ($M=6.19$ in science for women and $M=6.64$ for men) but slightly more competent in humanities ($M=6.17$ in literature for women and $M=6.14$ for men) according to the mean results shown. However, this difference needs further investigation, also in light of the fact that men with a full beard are more likely to be considered “scientists”, which could be due to the “higher masculinity” of bearded men but also to a possible (and not tested here) effect of a sort of “good scientist stereotype”, according to which the scientist is: a) male b) old and c) bearded. This alternative hypothesis is not present in the literature and needs at least a dedicated experiment.

Thus, in light of our results, we cannot say that we have reached equity between men and women in western society (as reported by Eagly, 2019): our study shows how women are always receiving lower scores than men ($M=6.39$ for male, $M=6.18$ for female) and start from a disadvantageous position, with more severe detrimental effects of both blonde hair and make-up. Furthermore, men are still considered as more inclined towards science than liberal arts. Nonetheless, some components or traits of the gender-science stereotype might be changing alongside culture and society. Another interesting point about this stereotype is that through the Make-up X Genre interaction we can see how the academic bias still remains with women wearing heavy make-up (hence higher contrast and higher perceived femininity) less penalised when associated to Literature, meaning that to a certain extent a woman is more of a novelist than a scientist, despite her looks. However, if our hypothesis that the gender-science stereotype has at least changed in its components, we must also be aware of what direction the stereotype is headed, so that we can understand what elicits our reactions to be able to avoid acting on our stereotypes. Of course, gender stereotypes, as any other kind of stereotype, are still present and we cannot ignore the fact that every society has a way of depicting men and women and these depictions are culturally generated or moderated (Cuddy, et al., 2015), but other origins are perceptual and extracted automatically and this will probably persist through generations. It is true that perception has also an environmental and cultural basis, but it still works as an immediate source of information, and it is emblematic how participants of this study, especially those in Exp.1 complained about how bad and superficial they felt for giving women with make-up very low probabilities of authorship. On the other hand, when judging male faces in Exp.2 they were complaining about the difficulty in giving scores to faces that appeared all equally probable, showing in this way the presence of the stereotype (no difficulty was reported in Exp.1).

For more considerations on the perceptual nature of this study the most important and unexpected result is that the combination of glasses and make-up on blonde hair elicited a sort

of “levelling out” on the stereotypical activation. As mentioned above we believe this is due to a strong perceptual component, and it’s most probably due to the black frame of the glasses which has a high contrast, somehow assimilated to the high contrast of heavy make-up. However, to further test this supposed prevalence of perception over stereotype it would be useful, in future studies, to use glasses with a white or transparent frame to see if the change in make-up will elicit any changes in evaluation. However, we must not demote the “cultural” stereotypes because in Exp.2, while our original intention was to compare the perceptual contrast of the make-up with the presence of a beard on men (i.e. heavy make-up equals full beard), we found opposite results, suggesting that the perceptual factor is not the only possible explanation. Results that are in line not only with the stereotypical image of a bearded scholar, but also with our original idea that increasing the perceived femininity (or masculinity in the case of the second experiment) will enhance the activation of the related gender stereotypes. So, it’s not only being a woman or a man that influences our judgement of them, but also “how much” women (i.e. feminine) or men (masculine) they are.

However, we must not forget that this is the first attempt at studying how all these variables and characteristics interact with one another so it would be beneficial to continue in this path. Moreover, this research used explicit measures, meaning that we directly asked participants to evaluate the pictures with no time constraints, which of course could lead to some suppression of the stereotypes and perceptions due to the social desirability bias. It is then even more interesting that, in a context of high social desirability (participants were sitting next to the experimenter), we obtained these gender biased results; with an implicit measure we would probably obtain the same results, if not stronger. Of course, this statement should be tested with a dedicated experimental protocol. Moreover, a suggestion for future studies would be to investigate how different fields of science and literature elicit different stereotypes (e.g., romance novels can be stereotypically associated to women more easily than crime stories or documentaries).

An ending note worth of mention is that we have found no difference in participants’ results based on their gender, meaning that these stereotypes seem to be equally present in men and women.

We can conclude by saying that we will always have to deal with stereotypes because as mentioned above they are a way to understand the world we live in, in a timely manner and help us in our daily life. What this research proposes to do is to offer a means to reflect on some features that trigger these gender stereotypes and try to avoid acting on them. And most importantly this is not and does not want to be an incentive for women to adapt and conform

to these stereotypes. On the contrary, we do believe that gender stereotypes could be better avoided when better known, instead of denying their existence.

References

- Actis-Grosso, R., Ricciardelli, P. (2017). Gaze and Arrows: The Effect of Element Orientation on Apparent Motion is Modulated by Attention. *Vision*, 1, 21.
- Bakmazian, A. (2014). The man behind the beard: Perception of men's trustworthiness as a function of facial hair. *Psychology*, 5, 185-191.
- Benci, V., Fortunato, D. (1998). An eigenvalue problem for the Schrödinger-Maxwell equations. *Topological methods in Nonlinear Analysis*, 11, 283-293.
- Bernard, P., Content, J., Servais, L., et al. (2020). An initial test of the cosmetic dehumanization hypothesis: heavy makeup diminishes attribution of humanness-related traits to women. *Sex roles* 83, 315-327.
- Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355(6323), 389-391.
- Bokova, I.G. (2017). Cracking the code: girls' and women's education in science, technology, engineering, and mathematics (STEM). *UNESCO, Global Education Agenda 2030*.
- Capellini, R., Sacchi, S., Ricciardelli, P., & Actis-Grosso, R. (2016). Social threat and motor resonance: when a menacing outgroup delays motor response. *Frontiers in Psychology*, 7, 1697.
- Ciardo, F., Marino, B. F., Actis-Grosso, R., Rossetti, A., & Ricciardelli, P. (2014). Face age modulates gaze following in young adults. *Scientific reports*, 4(1), 1-7.
- Ciardo, F., De Angelis, J., Marino, B., Actis-Grosso, R., & Ricciardelli, P. (2021). Social categorization and joint attention: Interacting effects of age, sex, and social status. *Acta psychologica*, 212, 103223.
- Cuddy, A. J. C., Wolf, E. B., Glick, P., Crotty, S., Chong, J., & Norton, M. I. (2015). Men as cultural ideals: Cultural values moderate gender stereotype content. *Journal of Personality and Social Psychology*, 109(4), 622-635.

- Diaz, V., Runyon, K., & Kroehler, C. J. (2020). Are scientists smart? Kindergarteners' gendered understanding and use of descriptors about science and intelligence. *Science Communication*, 42(4), 538-554.
- Dixson, A., Dixson, B., Anderson, M. (2005). Sexual selection and the evolution of visually conspicuous sexually dimorphic traits in male monkeys, apes, and human beings. *Annual Review of Sex Research*, 16, 1-19.
- Dutton, E., & Lynn, R. (2014). Intelligence and Religious and Political Differences Among Members of the US Academic Elite. *Interdisciplinary Journal of Research on Religion*, 10.
- Eagly, A. H., Nater, C., Miller, D.I., Kaufmann, M., Sczesny, S. (2019). Gender Stereotypes Have Changed: A Cross Temporal Meta-Analysis of U.S Public Opinion Polls from 1946 to 2018. *American Psychologist*.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37(2), 90–93.
- Ekman, P. (1992). Facial expressions of emotion: an old controversy and new findings. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 335(1273), 63-69.
- Fiske, S.T., Neuberg, S.L., Beattie, A.E., Milberg, S.J. (1987). Category-based and attribute-based reactions to others: Some informational conditions of stereotyping and individuating processes. *Journal of Experimental Social Psychology*, 23, 399-427.
- Fiske, S.T., Taylor, S.E. (2017). *Social Cognition: From Brains to Culture*. Sage.
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, 133(4), 694.
- Hellstrom, A., Tekle, J. (1994). Person perception through facial photographs: Effects of glasses, hair, and beard on judgments of occupation and personal qualities. *European Journal of Social Psychology*, 24, 693-705.
- Hering, E. (1964). *Outlines of a Theory of the Light Sense* (Hurvich, L.M. & Jameson, D., Trans.) Harvard Univ. Press, Cambridge, MA. (Original work published 1864).

- Losin, E. A. R., Iacoboni, M., Martin, A., Cross, K. A., Dapretto, M. (2012). Race modulates neural activity during imitation. *Neuroimage*, 59(4), 3594-3603.
- Lotto, R. B., & Purves, D. (1999). The effects of color on brightness. *Nature neuroscience*, 2(11), 1010-1014.
- Lotto, R. B., & Purves, D. (2000). An empirical explanation of color contrast. *Proceedings of the National Academy of Sciences*, 97(23), 12834-12839.
- Mellers, B., Wyers, R.S. (1990). Review of Advances in Social Cognition: Volume 1. A Dual Process Model of Impression Formation. *The American Journal of Psychology*, 103, 124-127.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of personality and social psychology*, 83(1), 44.
- O'Toole, A.J., Deffenbacher, K.A., Valentin, D. *et al.* The perception of face gender: The role of stimulus structure in recognition and classification. *Memory & Cognition* 26, 146–160 (1998).
- Pinna, B., & Deiana, K. (2019). When the whole is equal to the sum of its parts: A new approach to study face and body perception and representation. *Vision Research*, 157, 252-263.
- Slessor, G., Laird, G., Phillips, L. H., Bull, R., & Filippou, D. (2010). Age-related differences in gaze following: Does the age of the face matter?. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 65(5), 536-541.
- Weir, S., Fine-Davis, M. (1989). 'Dumb Blonde' and 'Temperamental Redhead': The effect of hair colour on some attributed personality characteristic of women. *The Irish Journal of Psychology*, 10 (1), 11-19.
- Workman, J.E, Johnson, K.K. (1991). The role of cosmetics in impression formation. *Clothing and Textiles Research Journal*, 10, 63-67.
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama*, 310(20), 2191-2194.
- Zebrowitz, A.L. (2012). Ecological and social approaches to face perception. *Oxford handbook of face perception*. Oxford University Press.

