# How are gender counter-stereotypical adolescents evaluated by their peers? Investigating the backlash effect in French schools

Ethan Meimoun[1] · Virginie Bonnot[1] · Julia Berenguer[1] · Cristina Aelenei[1]

**Abstract**

Research on the *backlash* effect has long studied the social sanctions that individuals impose on those who do not conform to stereotypical expectations. Specifically, research has focused on reactions to gender stereotype violations. Studies have shown that adults, adolescents, and even children will sanction others when they fail to conform to gender stereotypes. The present pre-registered research (https://tinyurl.com/4epypv3a) studied this gender *backlash* among adolescents in French middle and high-schools. For this study, we were also interested in exploring the moderating role of participants' socioeconomic status in their expression of *backlash*. We created four profiles of adolescents presenting their candidacy for a school representative role. The profiles varied in the personality traits and behaviors they contained (stereotypically feminine vs. masculine) and the gender of the target (boy vs. girl). A sample of 840 participants between the ages of 13 and 18 rated the target adolescent on measures intended to capture *backlash*, as well as gender stereotypicality items. Using multilevel models to analyze our data, we did not find evidence of *backlash* as hypothesized. However, unexpectedly, on a measure of perception of *arrogance,* we found that participants judged a counter-stereotypical target more positively than a stereotypical target. Furthermore, we also found evidence that, compared to stereotypical targets, counter-stereotypical targets were judged more competent, mostly by high SES adolescents. Possible methodological and theoretical explanations for these results are discussed, as well as implications for future research on the *backlash* effect, particularly among adolescents.

**Keywords** *Backlash* · Gender stereotypes · Adolescence · Socio-economic status

✉ Ethan Meimoun
   ethan.meimoun@etu.u-paris.fr

1   Université Paris Cité, Laboratoire de psychologie sociale, 92100 Boulogne-Billancourt, Paris, France

⚘ Springer

## 1 Introduction

Gender stereotypes and their consequences have been frequently studied in social psychology, and their prevalence in adolescence has been the subject of much research (Brown & Stone, 2016; Verniers et al., 2015). A large amount of work in developmental psychology has also investigated questions regarding gender stereotypes, such as their acquisition and use through childhood and adolescence (Blakemore, 2003; Charafeddine et al., 2020; McGuire et al., 2021). Researchers have also shown interest in how adolescents perceive others in relation to gender stereotypes. While they may claim that they would challenge these stereotypes, adolescents still expect their peers not to, perhaps because they are aware of the consequences of not conforming to gender norms (Mulvey & Killen, 2014; Mulvey et al., 2015). In fact, research shows that adolescents not only believe that their peers commonly adhere to gender stereotypes, but they also adopt these stereotypes themselves, with boys describing themselves in a more stereotypically masculine manner than girls, and conversely girls describing themselves as more feminine than boys (Klaczynski et al., 2020). Most interestingly, some research has highlighted that endorsement of gender stereotypes may be higher during adolescence than at other ages (Hill & Lynch, 1983; Klaczynski et al., 2020). Individuals between 12 and 17 years old have been found to report more gendered self-concepts and use gender stereotypes to predict their peers' behavior more than younger children or young adults (Klaczynski et al., 2020).

One of the social phenomena that could lead to this increased reliance on gender stereotypes is the important role that peer influence plays during adolescence (Blakemore & Mills, 2014). Behaviors that go against gender norms are then especially salient during this time, and have repercussions (Gordon et al., 2018; Navarro et al., 2015). A line of research that has extensively studied the phenomenon of reactions to gender norm violations is research on the *backlash* effect (Rudman, 1998), which describes how individuals sanction those who deviate from stereotypes as a way to preserve existing social hierarchies (Rudman et al., 2012a).

In the present research, we followed the theoretical and methodological framework of prior work on the *backlash* effect to study the social sanctions and negative social evaluations that adolescents impose on peers who violate gender stereotypic expectations.

### 1.1 Reactions to gender counter-stereotypicality in adolescence

#### 1.1.1 Research on adolescents sanctioning

Several studies over the years have investigated adolescents' reactions to their gender counter-stereotypical peers. For example, research with participants from the United States has highlighted the consequences of gender norm violations during adolescence in school contexts through cross-sectional studies linking gender counter-stereotypicality with school victimization (Jewell & Brown, 2013). Some studies

have found that, compared to adolescents who behave in ways consistent with gender norms, adolescents who do not conform to gender norms are more likely to be involved in a fight during the school year, to miss school due to bullying, or in general be victims of bullying (Gordon et al., 2018; Navarro et al., 2015).

These cross-sectional studies as well as other experimental works have also consistently found that the sanctioning against counter-stereotypical adolescents is stronger for boys than it is for girls (Sullivan et al., 2018; Yu et al., 2017). Although not capturing direct sanctions, Mulvey and Killen (2014) have identified that adolescent boys and girls from the USA tend to expect that a target peer who violates gender norms will be excluded from a group, and especially so if the target was a boy. Another study that specifically examined boys' reactions to counter-stereotypical boys found evidence of social sanctions against these "atypical" male targets (Lobel et al., 2004). In this study, Israeli adolescents and young adults were presented with an adolescent target who was a candidate for his class's representative election. This adolescent target was described either as masculine and very competent, masculine and average in competence, or feminine and very competent, using traits and behavior descriptions. On several measures such as ratings of likeability, election choice regarding this target candidate, and perceived popularity of this target, adolescent participants rated the counter-stereotypical target more negatively compared to the stereotypical targets. Moreover, supporting the idea that counter-stereotypicality is viewed particularly negatively during adolescence, these negative judgments of the feminine boy candidate were only expressed by adolescents and not young adults.

However, although still relevant regarding the evaluation of gender stereotypes during adolescence, past research (e.g., Horn, 2007; Lobel et al., 2004) using a methodology that involves comparing behaviors stereotyped as 'feminine' and 'masculine' may be limited in addressing the specific question of how counter-stereotypicality is evaluated. Indeed, we argue that this type of comparison is problematic because it leads to a comparison of behaviors in addition to a judgment on counter-stereotypicality. For example, a girl displaying stereotypical characteristics might be judged more positively than a girl behaving counter-stereotypically (with stereotypically masculine traits) simply because feminine characteristics are generally viewed more positively (Eagly et al., 1991). In this case, this might not be evidence of sanctions due to counter-stereotypicality, but rather of a preference for feminine characteristics (Iacoviello et al., 2021). Similarly, when comparing a feminine boy to a masculine boy, the feminine boy will likely be preferred, yet that would not reflect a preference for counter-stereotypicality, but, again, a preference for feminine traits. To deal with this potential issue—that is, the inherent valence of gender stereotypes—studies on the *backlash* effect have largely adopted the practice of comparing targets of different genders who display the same behaviors. For example, comparing the perception of a stereotypically feminine girl to that of a stereotypically feminine boy eliminates the confounding described above. Although this method might lead to another confounding variable, we believe that comparing targets of different genders and of different names will involve a much less severe confound than that associated with between behaviors comparisons. Thus, by adopting the *backlash* framework and its methodology in our study, we avoid the previously mentioned

confounding problem and are also able to recontextualize the findings of previous research within this theoretical framework.

### 1.1.2 Applying the backlash framework

The *backlash* effect framework details the sanctions that individuals impose on counter-stereotypical targets, as well as the psychological mechanisms at play in this sanctioning. Most studies on this effect have been carried out on adult gender norm violations—specifically women behaving in an agentic manner—in order to understand the phenomena that keep women out of leadership positions. However, recently, research has shown that the *backlash* framework can be relevant for younger targets (Sullivan et al., 2018). According to Rudman et al. (2012a), who first discussed the *backlash* effect, individuals sanction those who deviate from gender stereotypes as a way to uphold the current gender hierarchy. This reasoning is based on *system justification theory* and its conceptualization of stereotypes (Jost & Banaji, 1994; Verniers et al., 2015). Indeed, according to this view, stereotypes play a role in legitimizing social inequalities. In the case of gender, the gender hierarchy reflects an unequal system in which men have an advantageous position over women and benefit from higher social status and power than women. Gender stereotypes, which assign characteristics such as assertiveness, competitiveness, and ambition to men and characteristics such as sociability, warmth, and sensitivity to women, contribute to this gender hierarchy (Rudman et al., 2012b). Indeed, the characteristics which are stereotypically associated with men are labeled *agency* and reflect higher status, while the characteristics associated with women are labeled *communion* and reflect lower status (Rudman et al., 2012b). The *backlash* effect thus builds on *system justification theory* and suggests that by displaying traits or behaviors which are incongruent with their assigned status as men or women, counter-stereotypical individuals threaten the gender hierarchy. *Backlash* against those who are counter-stereotypical may thus be one of the ways through which individuals keep social hierarchies intact (Rudman et al., 2012a).

This theoretical perspective allows us to understand previous findings of greater sanctions toward counter-stereotypical boys compared to girls (Mulvey & Killen, 2014; Sullivan et al., 2018). This difference in sanction severity may occur because counter-stereotypicality from boys is seen as more threatening to the gender hierarchy—men (boys) have higher status and thus risk more (i.e. in terms of status loss) by behaving counter-stereotypically (Sirin et al., 2004). Relatedly, it may be expected that, since men (boys) represent a higher status group compared to women (girls), they may be more inclined to defend the gender hierarchy and thus sanction those who threaten it. Indeed, some authors have suggested that high status individuals should be the ones most likely to be motivated to justify existing systems (Owuamalam et al., 2018). However, here we follow the theoretical rationale of the *backlash* effect and *system justification theory,* which predicts—and mostly finds— no gender differences in the expression of sanctions against those who behave counter-stereotypically. This perspective rests on the idea that gender stereotypes are complementary and award both men and women with characteristics that are positively and negatively valenced (Jost & Kay, 2005), which contributes to creating

a perception of "gender fairness" that does not need to be challenged. Additionally, gender relations are characterized by a strong intimate interdependence due to the largely heterosexual functioning of society (Glick & Fiske, 2001). It follows that both men and women are motivated to maintain a certain level of positive contact and relational harmony through following the complementary gender guidelines (Lemus et al., 2010). Thus, both men and women (as well as boys and girls) should have an interest in policing gender counter-stereotypicality because it threatens the gender system that appears beneficial for all and on which "harmonious" gender relations rests.

In the present study, however, we instead focused on studying the role of socio-economic status, which has not yet received attention from past research on reactions to gender norm violations, although some previous studies have suggested that this variable may act as a potential moderator.

## 1.2 Examination of the role of perceiver's SES

Although previous *backlash* research has not examined the impact that individuals' socio-economic status (SES) might have on their reaction to gender counter-stereotypicality, some studies have gathered data which provide useful information regarding this question. Firstly, Hoffman et al. (2019), in a study of perceived pressure to conform to gender norms, showed that French boys of North African (Maghrebian) origins reported feeling more pressure to conform to gender norms than French boys of European backgrounds. As the authors explain, this result could be due to participants' SES rather than cultural background, as the two are largely confounded in French society, with members of cultural minority groups having on average lower socioeconomic status (National Institute of Statistical and Economic Studies, 2012). Furthermore, a study of Mexican American mothers and fathers by Leaper and Valin (1996) examined the relationship between several predictors and egalitarian gender attitudes. They found that participants' level of education significantly predicted egalitarian and progressive attitudes, including attitudes towards the acceptability of counter-stereotypical gender roles. Specifically, participants with lower education levels held less egalitarian gender attitudes than participants with higher education levels.

One possible explanation for these findings could be differences in how high and low SES individuals are socialized, such that high SES individuals might be socialized in environments that place a stronger emphasis on egalitarian values, or at least the social desirability of these egalitarian values (An, 2015). Indeed, parents can exercise a large influence on the gender development of children and adolescents (Xiao et al., 2022), and their general gender attitudes can impact those of adolescents. Some research has shown that stronger felt pressure to conform to gender norms coming from parents predicted more agreement with sexist remarks among children and preadolescents (Schroeder & Liben, 2020). Another possible explanation of SES differences could relate to the differences in self-construal between low and high social class individuals. It has been proposed that social class is related to self-construal, with individuals from lower

class backgrounds being more interdependent, while individuals from higher class backgrounds are more independent (Goudeau et al., 2017; Stephens et al., 2007). According to this perspective, interdependent individuals are more reliant on their social environment and have a stronger need for interpersonal harmony (Stephens et al., 2014), and thus might view adherence to gender roles as one way to preserve this harmony. Conversely, independent individuals tend to be more prone to seek distinctiveness and thus might value individuality (i.e., divergence from gender norms) in themselves or others (Meijs et al., 2015; Stephens et al., 2014).

Thus, the current study examined the link between adolescents' socioeconomic status—measured through their parents' educational level—and their reactions to gender stereotype violations. As Antonoplis (2022) points out, education level (one's own and parents') is only one component of socio-economic status, and researchers interested in SES should select the component that is most relevant to their research question (see also: Goudeau et al., 2017). Leaper and Valin's (1996) results suggest that education level might be the more relevant component of SES with regards to gender attitudes/perception of counter-stereotypicality, over, for example, income level, which they did not find to be a significant predictor of egalitarian gender attitudes.

## 1.3 Hypotheses and current study

The present research was intended to test previous findings on adolescent gender norm violations using the standard methodology of research on the *backlash* effect. Thus, for this study, we presented adolescent participants with a profile of an adolescent boy or girl who displayed traits and behaviors that were stereotypically feminine or masculine in a $2 \times 2$ between-participants design. Our study had three hypotheses which were based on the research detailed previously. First, we expected that adolescents would exhibit *backlash* and sanction counter-stereotypical adolescents regardless of their gender (H1) (Mayeux & Kleiser, 2019; Rudman et al., 2012a). Secondly, counter-stereotypical boys have been shown to receive stronger sanctions and more negative social evaluations than counter-stereotypical girls when they do not conform to gender expectations (Mulvey & Killen, 2014; Sullivan et al., 2018). Thus, we expected that the magnitude of sanctions and negative social judgments expressed by participants would be larger for the counter-stereotypical boy than the counter-stereotypical girl (H2). Finally, based on the aforementioned research that hints at stronger adherence to traditional gender norms and stereotypes among lower SES individuals (Leaper & Valin, 1996), a novel aim of the present research was to explore the moderating role of participants' SES on their expression of *backlash*. Our prediction was that lower SES adolescents would express stronger sanctions and more negative social evaluations than higher SES adolescents (H3). For reasons previously discussed, we did not expect that adolescents' gender would moderate their expression of *backlash* against counter-stereotypical targets.

## 2 Method

### 2.1 Participants

This study was carried out in middle and high schools in France during the 2022 school year. Schools were either contacted through standardized emails explaining the goal of the study or through personal contacts with the school directors. In the end, the study took place in 5 schools, and 34 classes were recruited with on average 24.7 students in each class. A majority of the data collection was done in person via pen and paper questionnaires; however, a portion of the data collection was also done online using Qualtrics. Participants were excluded from individual analyses based on our pre-registered criteria (Cooks' D and studentized deleted residual values; see https://tinyurl.com/4epypv3a). Participants who did not indicate their class names and who could not be identified as belonging to a certain class were also excluded from analyses. Because a portion of the sample (209 participants) was recruited online, we also excluded participants with a response time smaller than 3 min,[1] which we had not pre-registered. The final sample consisted of 840 participants. Participants' age ranged from 13 to 18 years old with a mean age of 14.9 years. The sample included 413 girls (49.2%), 394 boys (46.9%), 8 non-binary students (0.9%), 13 students identifying as "other" (1.5%), and 12 students who did not respond (1.4%).[2] We categorized participants' socio-economic status (SES) using their parents' education level according to our pre-registered criteria which we detail below. Our final sample included 537 students belonging to a "high" SES group (63.9%) and 248 (29.5%) belonging to a "low" SES group; the remaining participants either chose not to answer or did not supply enough information for us to categorize them (6.5%).

### 2.2 Sensitivity power analysis

We did not carry out a sample size calculation based on a power analysis before data collection. Our pre-registered analyses included the use of multilevel models (mixed models), and power analysis for these analyses requires prior knowledge of several features of the data such as the ICC (intra class correlation) as well as the variance components of the random effects. Since this information was not available to us either through a pilot study or prior research, we pre-registered no a-priori sample size calculations and collected as much data as possible within the available timeframe, and planned to run a sensitivity power analyses before conducting hypothesis tests. This sensitivity power analysis based on our data allowed us to identify the smallest detectable effect size for our design. This sensitivity analysis was carried out in *R* using the *simglm* package which allows to carry power analyses for multilevel models, using simulations. Carrying out this analysis, we find that our design

---

[1]  We identified this as the smallest feasible response time for our study.

[2]  Results did not vary when data from participants who did not identify as boy or girl were excluded.

has 80% power to detect interaction standardized regression coefficients of 0.4. This effect corresponds to an eta squared of 0.01, which is smaller than some previous research on gender norm violations, with Mulvey and Killen (2014) reporting an interaction effect of $\eta^2_p = 0.06$ and Sullivan et al. (2018) reporting interaction effects of 0.007 and 0.04. The code for this analysis can be found on the OSF page for this research (https://tinyurl.com/48x3546a).

## 2.3 Materials

For this study, two profiles were created which were intended to convey either masculine or feminine stereotypes. The profiles were presented as a fictional adolescent's candidacy speech to be elected school representative. These profiles were created based on gendered characteristics specific to adolescents, identified in a past study (Koenig, 2018). The masculine profile included traits such as *competitive*, *strong personality*, and *able to lead*. Conversely, the feminine profile included traits such as *communicative, empathetic,* and *emotional*. Both profiles also included an activity that the adolescent practiced, either *football* (soccer) for the masculine profile or *dance* for the feminine profile (Plaza et al., 2016). Finally, the descriptions of the adolescents included a candidacy proposal at the end of their speech which was intended to summarize the stereotypicality of the profiles. Masculine targets declared that they wanted to "organize sports tournaments between the classes" of their schools. The feminine targets proposed to "create discussion groups so that students could talk about their emotions and problems." Finally, targets were presented as being either boys or girls. Therefore, we had a $2 \times 2$ between-participants design, creating four conditions ("stereotypically feminine" boy or girl or "stereotypically masculine" boy or girl). All materials used can be found at the following link (https://tinyurl.com/48x3546a).

### 2.3.1 Measures

Nine self-report measures were administered to participants following the reading of the adolescent target's profile.

*Stereotypicality check items*. Four items were included in order to measure the stereotypicality of the profiles, as well as a way to measure facets of negative social evaluations. The first two questions were measures of masculine stereotypes with the items *aggressiveness* and *arrogance*. Past research has shown that these two traits were associated with masculine stereotypes (Koenig, 2018). For the other two items, participants were asked to rate the extent to which the profile characters possessed the stereotypically feminine traits, *naive* and *weak*. These four traits were previously used in another *backlash* study (Iacoviello et al., 2021). These items were chosen because they reflected gender proscriptive traits for boys and girls and thus allowed us to verify the gender stereotypicality of profiles while also capturing negative social evaluations of traits proscribed for each gender. For example, feminine boys might be evaluated as more naive and weaker than feminine girls because their counter-stereotypicality would intensify the evaluation of proscribed

traits. Similarly, counter-stereotypical girls may be evaluated as more aggressive and arrogant than typical boys. Participants had to rate how well the traits described the target they had read about. All items were rated on 7-point Likert scales (*1-Not at all* to *7-Very much*).

*Backlash*. We included three items used in past research in order to capture social preferences regarding (counter)stereotypical targets (Lobel et al., 2004; Mulvey & Killen, 2014). First, we asked participants to rate on a 7-point Likert scale how willing they would be to elect the target they had read about as school representative (*1-Not at all agree* to *7-Very much agree*). A second item asked participants to rate on a 7-point Likert scale how *likeable* they thought the target was (*1-Not at all* to *7-Very much*). Finally, in order to measure another facet of *backlash*, we asked participants to rate on a 7-point Likert scale how willing they would be to include the target in their friend group (*1-Not at all* to *7-Very much*).

*School popularity*. Another dependent measure was a scale intended to measure the perceived *popularity* of the target. This measure was intended to capture a sort of "school *backlash*" through participants' wider perception of how the target would fare in the school context. This item also allowed us to measure a different facet of social evaluation more linked with power and status, rather than simply social preference (van den Berg et al., 2020). The measure's instructions and presentation were based on the Subjective Social Status Scale (Giatti et al., 2012). Participants saw an image of a ladder with ten rungs and the following instructions: "The scale below represents the place that students occupy in their school. At the top of this scale (10) are the students who are the most *popular*, and therefore, the ones who have the most friends, influence the styles that other students adopt, and choose the activities to engage in and the people to include. At the bottom of the scale (1) are the students who are the least *popular*, that is, the students who have the fewest friends and the least influence concerning style, choice of activities, or people to be included in activities." We then asked participants to indicate where they would place the target on this ladder. These instructions were written using the typical formulation of the Subjective Social Status Scale and changing key terms to include the definition of school *popularity* from Brown (2004).

*Competence*. The questionnaire also included a measure of perceived *competence,* asking participants to rate how *competent* they thought the target was on a 7-point Likert scale (*1-Not at all* to *7-Very much*). This item is often included in *backlash* research on adults and usually shows no *backlash* effect against counter-stereotypical targets (Rudman et al., 2012b; Sullivan et al., 2018).

*Socioeconomic status*. Participants indicated whether their parents had any post high-school education as well as their profession. We provided adolescents with broad answer choices: "no high-school diploma", "a high-school diploma" or "a post high-school diploma", because we anticipated that demanding precise information regarding the type of diploma obtained by their parents would be difficult for many adolescents. Participants also had the option to check an "I do not know" answer option. We planned to categorize participants as high SES if either of their parents had one form of post high-school education. On the contrary, if both participants' parents had no post high-school education, we categorized them as low SES (previous research used similar categorization, Harackiewicz

et al., 2014; Jury et al., 2018; Sommet et al., 2015). For participants who did not report their parents' education level, we used the education level necessary for their parents' profession and categorized them accordingly. If no clear information was reported by participants, we did not categorize them in either group.

## 2.4 Procedure

Data collection took place in-person in four schools in France and online in one French school. Since the study did not include any type of intervention (the students were only asked to give their impressions of a fictional student's profile presented to them), the majority of schools informed parents of the general nature of the study (social evaluation of peers in adolescence) and parents could retract consent for their children. Moreover, for every in-class data collection session and for the online data collection, participants were reminded of their rights as participants in psychological research and were free to not participate in the study or to stop at any point they wished. For in-person data collection, presentation of the study to participants and supervision was carried out by the first author and two master's students. Participants who responded online did so during a school hour where they had access to an individual computer and were supervised by their teacher. Participants were all (in person and online) informed that they were asked to take part in a study of adolescents' social judgments of their peers and specifically of an adolescent who wants to be elected school representative. Students were informed that they were free to participate or not in the study and that their responses would be completely anonymous and analyzed at the collective level. For in-person data collection, in order to avoid students looking at their class neighbors' questionnaires and thus becoming aware of our experimental manipulation, we distributed the same questionnaire to each student in a given row. However, the condition to be distributed to each row was chosen at random by experimenters. After reading the information about the study on the first page of the questionnaire, students gave their consent and were asked to turn over the page and read the presentation of the target candidate (*Léa* or *Théo* presented with a feminine or masculine profile, based on participants' experimental condition). Following this presentation, participants responded to our dependent and socio-demographic measures. For in-person data collection, after all participants in a classroom had finished responding, the experimenters thanked students for their participation and carried out a debriefing which took the form of a small 30-min discussion presenting the aim of the study, as well as giving students information about gender stereotypes and their consequences. Students could also ask questions and discuss the study with each other during this time. After the discussion, participants were once again thanked and dismissed. For online data collection, after completing the study, participants were thanked for their participation and received a debriefing form describing the objective of the study and its results at a later point in order to avoid communication between students who had not taken part in the study at the exact same time.

# 3 Results

## 3.1 Analysis plan

Considering our design and the likely non-independence of our data due to students being nested in classrooms, we planned to use multilevel modeling to analyze our data. As pre-registered, our analysis plan followed the recommendation of Sommet and Morselli (2021). For each dependent variable, we first ran empty models using the *lmer* function in *R* in order to calculate our models' Design EFFect (DEFF) and Intra Class Correlation (ICC). Their recommendations suggested treating DEFFs close to (or larger than) 1.5 and ICCs larger than 0.01 as evidence of the need to use multilevel modeling. However, differing from our pre-registration and to remain consistent in our analyses and their interpretability, we decided to use multilevel modeling in every case, even when the DEFFs or ICCs were lower than the values suggested by Sommet and Morselli (2021). The values of the ICCs and DEFFs can be found on the OSF in our analysis code (https://tinyurl.com/48x3546a). For our analyses, our models included a random intercept for participants' classes. As fixed effects, we included our two manipulated independent variables, the type of profile (coded −0.05 for *feminine* and 0.5 for *masculine*), the gender of the target (coded −0.05 for *girl* and 0.5 for *boy*), participants' SES as our moderator (coded −0.5 for *low* and 0.5 for *high* SES), as well as the three two-way interactions and the three-way interaction. As mentioned in the introduction, we did not predict any moderating role of participants' gender on our effect of interest (i.e., the interaction of the two manipulated variables). However, in preliminary analyses, we included this moderator with our manipulated variables in our models. In accordance with past *backlash* research, we did not find any moderation of relevant effects.[3]

## 3.2 Manipulation check items

Regarding item aggregation, we first entered our four stereotypicality items into an exploratory factor analysis to determine whether our items could be indexed. As can be seen in Table 1, we found that two factors emerged, one with the *masculine* items and the other with the *feminine* items. As a second step, we checked correlations between the items and found that the *aggressiveness* and *arrogance* items were positively correlated and that this correlation was of medium strength ($r = 0.43$, $p < 0.001$; Cohen, 1992). Similarly, for the feminine stereotype items of *naiveté* and *weakness*, there was a positive correlation of medium strength ($r = 0.41$, $p < 0.001$). These correlations were in our opinion too low to allow for aggregation of the measures. Furthermore, we observed notable divergences between individual and aggregated analyses. Thus, we report below all four items analyzed separately.

---

[3] Including participant's gender as a single moderator or with participant's SES in the same model did not reveal any relevant effects (moderation of our two manipulated variables' interaction or four-way interaction).

**Table 1** Exploratory factor analysis of the four stereotypicality check items

| Factor loadings | Factor | | |
|---|---|---|---|
| | 1 | 2 | Uniqueness |
| Arrogance | 0.638 | | 0.547 |
| Naivete | | 0.629 | 0.509 |
| Aggressiveness | 0.687 | | 0.556 |
| Weakness | | 0.682 | 0.564 |

EFA uses a "Principal Axis" extraction method to deal with the non-normal distribution of some of the measures and a "promax" rotation to allow correlation between the factors

*Masculine stereotypes.* First, in line with the masculine gender stereotype content, on the *arrogance* measure, we found a main effect of the type of profile, showing that masculine targets ($M=3.03$, $SD=1.72$) are seen as more *arrogant* than feminine targets ($M=2.45$, $SD=1.60$),[4] $t(774)=4.51$, $p<.001$, $ß=0.37$, 95% CI [0.21, 0.53], $\eta^2_p=0.03$.[5] We did not find an effect of the target's gender ($ß=0.06$, $p=.46$, $\eta^2_p<0.001$). Interestingly, we also found an unexpected interaction between the type of profile and the gender of the target, $t(769)=2.37$, $p=.018$, $ß=0.39$, 95% CI [0.07, 0.70], $\eta^2_p=0.009$. This interaction shows that gender counter-stereotypical targets were seen as less *arrogant* than the stereotypical targets. We then exploratorily looked into the simple effects of the target's gender on each profile. The effect of the target's gender in the feminine profile was not significant ($ß=-0.13$, $p=.26$, $\eta^2_p=0.003$), while it was significant for the masculine profile, $t(769)=2.26$, $p=.024$, $ß=0.25$, 95% CI [0.03, 0.47], $\eta^2_p=0.005$ and showed that masculine girls ($M=2.91$, $SD=1.74$) were rated less *arrogant* than masculine boys ($M=3.14$, $SD=1.67$). The cell means for these effects are displayed in Fig. 1. Regarding the moderating role of participants' SES, we found no effect of SES ($ß=0.12$, $p=.15$, $\eta^2_p=0.002$), no interactions between SES and target's gender ($ß=-0.29$, $p=.07$, $\eta^2_p=0.004$) or profile ($ß=-0.15$, $p=.35$, $\eta^2_p=0.002$) as well as no interaction between the three ($ß=-0.22$, $p=.49$, $\eta^2_p=0.001$).

On the *aggressive* measure, we found similar results, showing that masculine targets ($M=2.47$, $SD=1.54$) were seen as more *aggressive* than feminine targets ($M=1.44$, $SD=0.93$), $t(776)=10.16$, $p<.001$, $ß=0.61$, 95% CI [0.50, 0.73], $\eta^2_p=0.13$. The effects of the target's gender or its interaction with the type of profile were not significant (respectively: $ß=-0.07$, $p=.25$, $\eta^2_p=0.002$; $ß=—0.09$, $p=.46$, $\eta^2_p<0.001$). On this measure, we also found an effect of participant's SES showing that higher SES students ($M=2.08$, $SD=1.42$) found the target more *aggressive* than low SES students ($M=1.75$, $SD=1.27$), $t(661)=4.71$, $p<.001$, $ß=0.28$, 95% CI [0.17, 0.40], $\eta^2_p=0.02$. The two effects were qualified by an interaction, $t(780)=2.87$, $p=.004$, $ß=0.35$, 95% CI [0.11, 0.58], $\eta^2_p=0.006$. This interaction

---

[4] Means and standard deviations reported are not standardized.

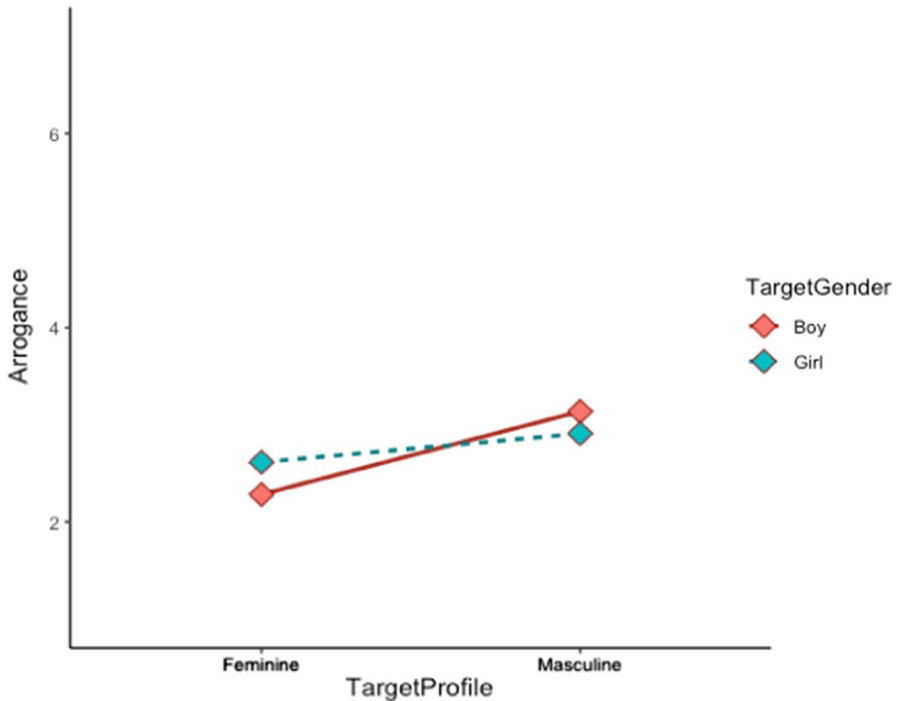[5] All regression coefficients presented are standardized.

**Fig. 1** Perception of the target's arrogance, based on their profile and gender

showed that the effect of the type of profile previously identified (i.e. masculine targets being seen as more aggressive than feminine targets) was stronger for high SES adolescents. Neither the interaction between SES and target's gender nor the three-way interaction were significant (respectively: ß $= -0.11$, $p = .37$, $\eta^2_p < .001$; ß $= 0.19$, $p = .77$, $\eta^2_p < 0.001$).

*Feminine stereotypes*. Conversely, but again in line with the content of gender stereotypes, on the *naiveté* measure, we found an effect of the type of profile showing that feminine targets ($M = 3.15$, $SD = 1.65$) were seen as more naive than masculine targets ($M = 2.60$, $SD = 1.55$), $t(779) = -4.29$, $p < .001$, ß $= -0.36$, 95% CI [$-0.52$, $-0.19$], $\eta^2_p = 0.026$. The effects of target's gender or the interaction between the two were not significant (respectively: ß $= 0.04$, $p = .67$, $\eta^2_p < 0.001$; ß $= -0.02$, $p = .88$, $\eta^2_p < 0.001$). Regarding the moderating role of participants' SES, we found an interaction between participants' SES and the target's gender, $t(779) = -2.41$, $p = .016$, ß $= -0.40$, 95% CI [$-0.72$, $-0.08$], $\eta^2_p = 0.006$. This interaction showed that high SES participants' rating of the target as naïve was higher for the girl target, while for low SES adolescents, it was higher for the boy target.

For the *weakness* measure, we similarly found a main effect of the type of profile, showing that feminine targets ($M = 3.32$, $SD = 1.85$) were seen as weaker than masculine targets ($M = 2.06$, $SD = 1.30$), $t(783) = -9.37$, $p < .001$, ß $= -0.70$, 95% CI [$-0.85$, $-0.56$], $\eta^2_p = 0.12$. The effects of the target's gender and the interaction between our manipulated variables were not significant (respectively: ß $= 0.01$,

**Table 2** Exploratory factor analysis of the four *backlash* items

| Factor loadings | Factor | |
|---|---|---|
| | 1 | Uniqueness |
| Election | 0.537 | 0.712 |
| Likeability | 0.847 | 0.283 |
| Socialinclusion | 0.705 | 0.502 |
| Popularity | | 0.995 |

EFA uses a "Principal Axis" extraction method to deal with the non-normal distribution of some of the measures and a "promax" rotation to allow correlation between the factors

$p = .84$, $\eta^2_p < 0.001$ and $\beta = 0.16$, $p = .29$, $\eta^2_p = 0.001$). Similarly, the effects of participants' SES, as well as its interaction with target's gender, profile, or both were not significant (respectively: $\beta = 0.04$, $p = .57$, $\eta^2_p < 0.001$; $\beta = -0.14$, $p = 0.37$, $\eta^2_p < 0.001$; $\beta = -0.11$, $p = .48$, $\eta^2_p < 0.001$; $\beta = 0.39$, $p = .20$, $\eta^2_p < 0.001$).

### 3.3 Hypotheses tests

### 3.3.1 Backlash index

For the four *backlash* items, entering these items into an EFA led to the identification of one factor which included the *likeability*, *social inclusion* and *election* items, but did not include the *popularity* measure, as can be seen in Table 2. As a second step, checking the reliability of a *backlash* index (including items loading onto the factor previously mentioned), we found that internal consistency was satisfactory (alpha = 0.72) and thus we aggregated these three items into a *backlash* index. Below, we report results for this index and results for the *popularity* measure independently.

For this measure, because the assumption of normality was violated (i.e., the Shapiro–Wilk's W test on unstandardized residuals was significant, $p = .001$), we used a bootstrap estimation of our model parameters (Berkovits et al., 2000). On this *backlash* index, we found a main effect of the type of profile showing that feminine targets ($M = 4.69$, $SD = 1.16$) were more likable, more likely to be elected as school representative, and more likely to be included in one's friend group compared with masculine targets ($M = 4.47$, $SD = 1.15$), $p = .01$, $\beta = -0.21$, 95% CI [$-0.34$, $-0.05$], $\eta^2_p = 0.009$. We found no effect of the target's gender ($\beta = -0.009$, $p = .91$, $\eta^2_p < 0.001$), and the interaction between the two was also not significant ($\beta = -0.21$, $p = .19$, $\eta^2_p = 0.002$). No effects of SES or of its interaction with the type of profile, gender of the target, or the two together, were identified (respectively: $\beta = 0.05$, $p = .51$, $\eta^2_p = 0.002$; $\beta = 0.16$, $p = .27$, $\eta^2_p = 0.001$; $\beta = 0.09$, $p = .53$, $\eta^2_p < 0.001$; $\beta = 0.12$, $p = .68$, $\eta^2_p < 0.001$).

### 3.3.2 Popularity

On our *popularity* measure, we found a main effect of the type of profile, showing that targets with a masculine profile ($M=6.44$, $SD=2.20$) were perceived as more *popular* than targets with a feminine profile ($M=5.56$, $SD=1.86$), $t(737)=5.39$, $p<.001$, ß$=0.39$, 95% CI [0.25, 0.53], $\eta^2_p=0.037$. The effect of the gender of the target (ß$=-0.07$, $p=.38$, $\eta^2_p<0.001$) and the interaction were, however, not significant (ß$=0.03$, $p=.81$, $\eta^2_p<0.001$). We also found no effect of participants' SES, or its interaction with the target's gender or profile or both (respectively: ß$=0.06$, $p=.41$, $\eta^2_p=0.001$; ß$=-0.05$, $p=.76$, $\eta^2_p<0.001$; ß$=0.20$, $p=.162$, $\eta^2_p=0.003$, ß$=0.19$, $p=.50$, $\eta^2_p=0.001$).

### 3.3.3 Competence

Again, due to the violation of the normality assumption ($p<.001$), we used bootstrap estimation of our model parameters. On this *competence* rating item, we found no main effect of target's profile or gender, as well as no interaction between the two (respectively: ß$=-0.09$, $p=.22$, $\eta^2_p=0.002$; ß$=-0.04$, $p=.53$, $\eta^2_p<0.001$; ß$=-0.16$, $p=.27$, $\eta^2_p<0.002$). However, there was a significant three-way interaction between our two manipulated variables and participants' SES, $p=.046$, ß$=-0.56$, 95% CI [−1.16, −0.01], $\eta^2_p=0.005$. To further understand this three-way interaction based on our hypothesis, we looked at the two-way interaction between our two manipulated variables for low and high SES participants. The two-way interaction for low SES adolescents was not significant, ß$=0.10$, $p=.67$, $\eta^2_p<0.001$. However, the interaction for high SES participants was significant, $p=.004$, ß$=-0.45$, 95% CI [−0.76, −0.13], $\eta^2_p=0.009$. This showed that the higher *competence* rating for counter-stereotypical targets was mostly driven by high SES students. Figure 2 shows high and low SES participants' ratings on the *competence* measure based on targets' profile and gender.

## 4 Discussion

The point of this study was to investigate the social sanctions targeting adolescents who do not conform to gender stereotypes. Based on previous studies of this phenomenon, as well as research on the *backlash* effect, we made several predictions starting with the expectation that counter-stereotypical adolescents would be judged more negatively than their stereotypical counterparts. We further predicted that counter-stereotypical boys would receive more severe sanctions than counter-stereotypical girls. Finally, we also wished to investigate the effect of participants' socioeconomic status on their reactions to their counter-stereotypical peers. In our study we found no evidence of *backlash* against counter-stereotypical targets. In fact, contrary to previous research, we found some evidence of adolescents evaluating counter-stereotypical targets less negatively. In particular, we found that adolescents perceived the counter-stereotypical targets as less *arrogant* than stereotypical
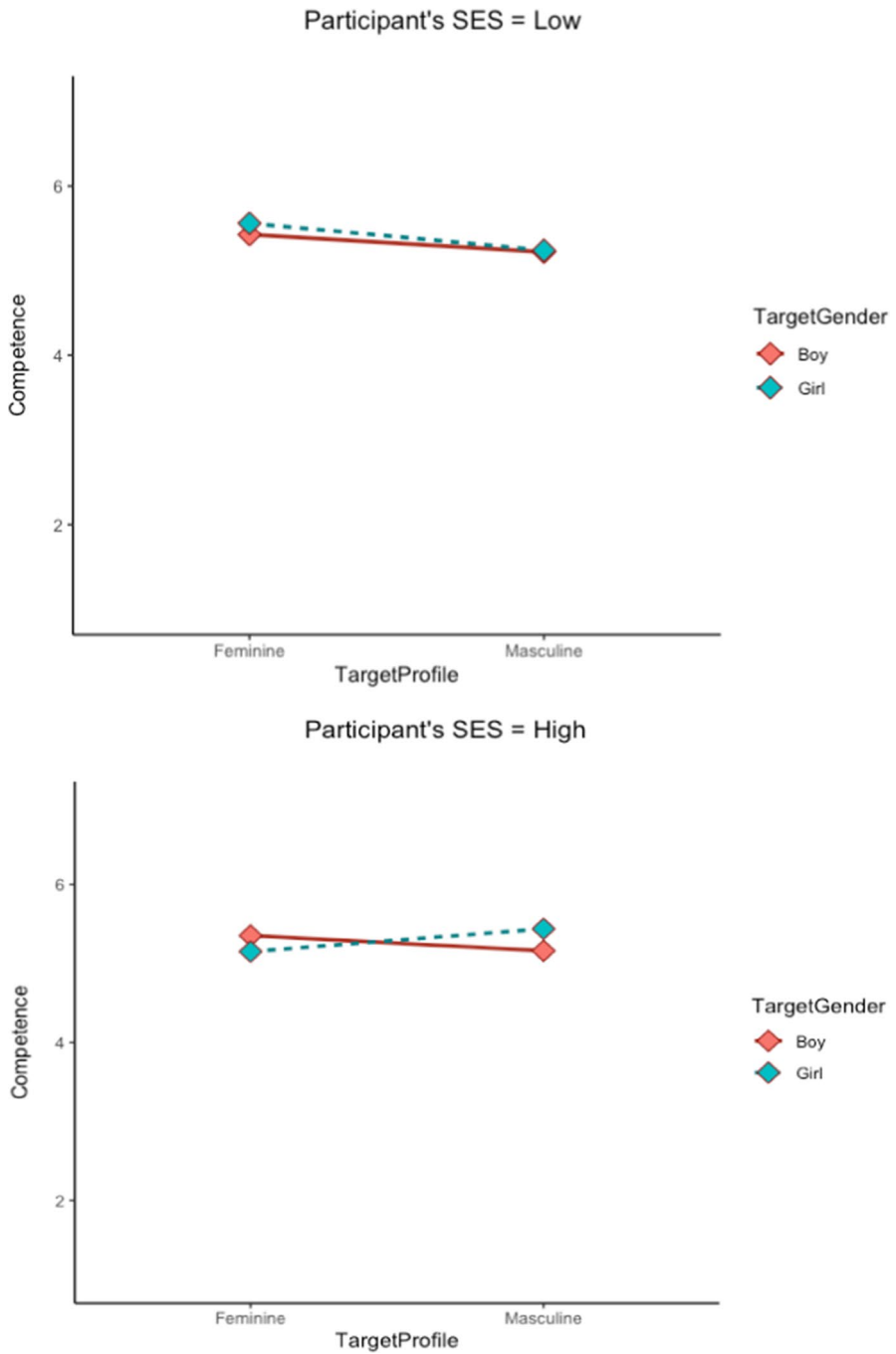
**Fig. 2** Perception of the target's competence by their Profile and Gender, for Low and High SES participants

targets. We also found evidence of the moderating role of adolescents' socio-economic status such that the higher perceived competence of counter-stereotypical targets was mostly driven by high SES adolescents in our data. Other results linked with adolescents' SES were identified; however, these effects were of little interest for our research and no clear patterns emerged.

## 4.1 Stereotypes and *system justification theory*

First, not yet mentioning our findings regarding perceptions of counter-stereotypical targets, what we have found overall are results which are congruent with gender stereotypes and their role in *system justification theory*. According to the system justification perspective, stereotypes legitimize status inequalities between men and women by assigning stereotypical characteristics to men which are linked to high social status. On the other hand, characteristics stereotypically assigned to women are linked to lower social status but more positive social perceptions (Jost & Banaji, 1994; Rudman et al., 2012b).This idea is also reflected in the results of research on benevolent sexism (Glick & Fiske, 2001; Sarlet & Dardenne, 2012), which shows that one way through which the gender hierarchy is preserved is by perceiving women (and feminine stereotypes) positively but with no real power and as ultimately inferior to men (and masculine stereotypes). This was precisely the case in our study, in which we found that in general, the feminine targets were viewed more positively than masculine targets, but at the same time were viewed as having lower power and influence. Participants rated feminine targets more positively on our *backlash* index, which included the *likeability* and *social inclusion* items. This *backlash* index also included the *election* measure, and though it is not clear whether the position of school representative reflects high social status and/or power, we interpret it as similar to the other two measures composing this index. Indeed, here the role of school representative might reflect a role of communicator/facilitator between students and teachers and might not involve any power or agency over others, which is instead reflected in the *popularity* measure.

On this measure of *popularity* and in line with the points just made, we have found that compared with feminine targets, adolescent boys and girls with a masculine profile were rated more popular, which we and previous authors (Brown, 2004; Kleiser Polk & Mayeux, 2022) have defined as encapsulating social status and power. Therefore, our findings on the *popularity* measure are consistent with studies showing a link between masculinity and popularity (Jewell & Brown, 2013; Lobel et al., 1993), rather than those showing a link between conformity to gender stereotypes and popularity (Kleiser & Mayeux, 2020, 2022). To summarize, as *system justification theory* posits, masculine stereotypes are more aligned with the power and high status of school popularity, while feminine stereotypes are more aligned with positive social perceptions but lower popularity. These findings are also consistent with past research, which has found that popularity and "social preference" are different constructs (van den Berg et al., 2020), and we have seen here how gender stereotypes are differently related to popularity than to social preference. This finding regarding popularity also lends credence to the idea that school contexts might not be as "feminine" as some literature suggests (see also

Verniers et al., 2016 for a similar position on this issue). Indeed, here we show that masculine characteristics might be favored in perceptions of popularity.

Contrary to our hypotheses, in our study we did not find social sanctions or negative social perceptions of adolescents whose traits and behaviors were incongruent with their gender. In fact, we found the opposite, that those counter-stereotypical adolescents were judged as less arrogant by our participants. However, as expected, we also found a moderating effect of adolescents' SES on their perception of counter-stereotypical peers. Next, we discuss possible explanations and implications of these findings.

## 4.2 Possible explanations

### 4.2.1 Positive reactions to counter-stereotypicality

To explain our findings, which run contrary to what has previously been reported, we point to the general social/societal context our study took place in. Our study was carried out in French middle and high-schools in 2022. The current French government has on several occasions expressed that gender equality and the reduction of gender discrimination would be a "Great national cause" of the presidential term. In fact, since 2018, the Ministry of National Education and Youth has defined several measures that are to be implemented in French schools in order to work on these questions of gender equality and gender discrimination (such as appointing in-school gender equality referents; see, for example, Avenel, 2021). Lending support to this view, after discussions with the participating schools, we learned that students from two schools—at least—received several interventions discussing questions of gender equality during the school year, before our study took place. Furthermore, it is possible that by contacting schools, specifically to participate in a study on gender stereotypes, we may have attracted mostly schools which are especially concerned with these questions.

In addition to the political propositions from the government, adolescents and younger generations in general seem to be concerned with questions of gender norms and gender fluidity (Bragg et al., 2018). Indeed, through traditional as well as social media, adolescents are increasingly exposed to content that deals with gender expressions or stereotypes or that questions the traditional binary vision of gender (Gravillon, 2022; for recent work on the role of media in transmission of gender stereotypes, see Lamer et al., 2022). Adolescents are now more aware than ever of discourse around questions of gender identity such as transgender identity, non-binary identification, and overall non-adherence to traditional gender norms (Bragg et al., 2018). As we were able to anecdotally observe during debriefing sessions after the study, adolescents were much more aware of concepts of gender identity, trans-identity, and sexual minorities than we anticipated. Thus, one possible explanation for our results is the changing of adolescents' norms towards gender expression due to governmental actions, broader discussions of these issues in society at large, or perhaps a combination of both.

However, since very little research in recent years has investigated the questions of gender norms in French samples of adolescents, it would be difficult to

definitively conclude that this explains our results. In fact, one of the few recent studies that has investigated the question of gender norms in French adolescents shows that boys still largely report feeling increased pressure to conform to gender norms as they move through adolescence, suggesting that changes in the perception of gender norms in adolescence might not be so straightforward (Hoffman et al., 2019). One possible explanation that could reconcile these contradictory points is that, as gender equality questions become more and more prevalent in schools and in society at large, discriminatory behaviors against counter-stereotypical individuals become less socially acceptable. Thus, it might not be that individuals' perception of counter-stereotypicality has become more positive, but rather that negative treatment of counter-stereotypical peers might now be extremely socially undesirable, as communicated by peers and or (social) media. Indeed, past research dealing with gender stereotypes and gendered behavior has shown the importance of considering the role of larger intersubjective norms in predicting individual behavior (Lamer et al., 2022). Thus, studying the current adolescent/school norms in France regarding perception and treatment of counter-stereotypical individuals—such as the acceptability of sanctioning these individuals—would be a worthwhile endeavor and might shed some light on current trends of gender discrimination in adolescence.

### 4.2.2 The role of socio-economic status

Regarding the inclusion of participants' SES as a moderator of *backlash*, we found evidence that a higher evaluation for counter-stereotypical targets on the *competence* measure was mostly accounted for by high SES adolescents. This evidence, although not at all definitive, shows that high SES adolescents perceive counter-stereotypical targets differently (in our study, it seems, more positively) than stereotypical targets, whereas we could not draw such conclusions for low SES adolescents. As mentioned in the introduction, it is perhaps the different types of upbringings of high and low SES individuals that could lead to different reactions to gender counter-stereotypicality. Indeed, work on the impact of social class on self-construal highlights that independent individuals, socialized in higher social class contexts, are more focused on individual self-expression and distinctiveness than others (Stephens et al., 2014). In this view they might place less emphasis on gender stereotypes to govern their behavior and by extension the behaviors of others as well. Thus, it could be that higher social class individuals show more positive reactions to counter-stereotypicality because they view individuality more positively (for discussion of this point, see also: Meijs et al., 2015). Future research is greatly needed not only to replicate our findings on this question but also to investigate what processes might lead to this effect. Indeed, though interesting, our findings are modest and further research is needed to replicate them. As we discuss next, future research on this specific question should take into account the distribution of high and low SES individuals in their samples, if the effect we have found here is to be further investigated.

## 4.3 Limitations and areas of improvement

While our pattern of results is interesting and contradicts past research, it is worth noting that several limitations exist in our study. A first important limitation of our work is the possibility that we did not successfully render salient the power associated with the school representative role. Indeed, theoretically, the *backlash* effect rests on the idea that, by behaving counter-stereotypically, individuals threaten the existing gender hierarchy (Rudman et al., 2012a). Due to the nature of the gender hierarchy, it might be especially threatened by counter-stereotypical women (girls) trying to obtain power or counter-stereotypical men (boys) at risk of losing power (Rudman et al., 2012b; Sirin et al., 2004). In our case, while we did intend for the school representative role to represent power, this might not have been the case. Indeed, in our study, we did not explain what a school representative could be or what their position entailed. It could be that students did not perceive that such a role would involve any power. Indeed, the correlation between the *election* measure and the *popularity* measure is very small and not significant ($r = 0.011$, $p = 0.75$). The lack of association between the *election* measure and *popularity* measure (which we know denotes power because of the definition of the concept included in our measure) does point to a lack of association with power. Future research on the *backlash* effect, specifically with adolescents, should take into account the contexts used and their link with power, as those contexts should be the ones most likely to elicit *backlash* (Mishra & Kray, 2022).

We point next to the unequal distribution of low and high SES students in our sample. We did find modest evidence showing that the higher perception of competence for counter-stereotypical targets was mostly driven by high SES adolescents. However, this effect was only observed on one measure. One possible explanation for this could be that our study was underpowered for detecting this moderating effect on other measures (Maxwell, 2004), specifically that the two groups of high and low SES were too unbalanced (Wilcox, 1992). Future research which examines the link between socio-economic status and *backlash* should put an emphasis on obtaining better distributed samples in terms of SES.

Finally, one last limitation of our research is our inability to interpret null results on our effects of interest. On most of our DVs, we did not observe significant interactions between our manipulated variables. In line with a proper use of null-hypothesis significance testing, we cannot interpret these non-significant effects as evidence of no effect. Instead, equivalence testing is required to determine a lack of effect (or rather a negligible effect), which requires selecting a smallest effect size of interest/that is meaningful for the research question (Lakens et al., 2018). Recommendations of best practice for using equivalence testing suggest that the smallest effect size of interest should be chosen before conducting hypotheses tests (Alter & Counsell, 2023). Thus, a future area of research for studies focused on the *backlash* effect should be to make use of equivalence testing and its ability to make sense of non-significant and/or negligible effects. This process would allow for a more complete understanding of when individuals display *backlash* and when they do not.

Bearing in mind these limitations, in the present research, we studied a population which has seldom been investigated in previous *backlash* research. We did

so by using a methodology well suited to investigating this phenomenon, as well as statistical analyses adapted to such contexts, which has not always been the case in previous research. Furthermore, our study benefits from having good statistical power for addressing this research question and being pre-registered. In this paper, we have raised important theoretical and methodological questions for future *backlash* research, specifically with adolescent participants, which we hope will inspire future research aimed at understanding the *backlash* effect.

**Data availability** Data available at the following link: https://tinyurl.com/48x3546a.

## Declarations

**Competing interests** The authors report there are no competing interests to declare.

## References

Alter, U., & Counsell, A. (2023). Determining negligible associations in regression. *The Quantitative Methods for Psychology, 19*(1), 59–83. https://doi.org/10.20982/tqmp.19.1.p059

An, B. P. (2015). The role of social desirability bias and racial/ethnic composition on the relation between education and attitude toward immigration restrictionism. *The Social Science Journal, 52*(4), 459–467. https://doi.org/10.1016/j.soscij.2014.09.005

Antonoplis, S. (2022). Studying socioeconomic status: Conceptual problems and an alternative path forward. *Perspectives on Psychological Science, 18*(2), 275–292. https://doi.org/10.1177/17456916221093615

Avenel, C. (2021, October). *IL, Elle, On : pour l'égalité filles-garçons, contre les violences sexistes et sexuelles [He, She, They: for gender equality, against sexist and sexual violence]*. Académie de Versailles. Retrieved October 27, 2022, from https://www.ac-versailles.fr/il-elle-pour-l-egalite-filles-garcons-contre-les-violences-sexistes-et-sexuelles-124217.

Berkovits, I., Hancock, G. R., & Nevitt, J. (2000). Bootstrap resampling approaches for repeated measure designs: Relative robustness to sphericity and normality violations. *Educational and Psychological Measurement, 60*(6), 877–892.

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles, 48*(7-8), 411–419. https://doi.org/10.1023/A:1023574427720

Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology, 65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Bragg, S., Renold, E., Ringrose, J., & Jackson, C. (2018). 'More than boy, girl, male, female': Exploring young people's views on gender diversity within and beyond school contexts. *Sex Education, 18*(4), 420–434. https://doi.org/10.1080/14681811.2018.1439373

Brown, B. B. (2004). Adolescents' relationships with peers. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 363–394). Wiley.

Brown, C. S., & Stone, E. A. (2016). Gender stereotypes and discrimination: How sexism impacts development. In S. S. Horn, M. D. Ruck, & L. S. Liben (Eds.) *Advances n Child Development and Behavior (Vol. 50): Equity and Justice in Developmental Science: Theoretical and Methodological Issues* (pp. 105–133). https://doi.org/10.1016/bs.acdb.2015.11.001

Charafeddine, R., Zambrana, I. M., Triniol, B., Mercier, H., Clément, F., Kaufmann, L., Reboul, A., Pons, F., & Van der Henst, J.-B. (2020). How preschoolers associate power with gender in male-female interactions: A cross-cultural investigation. *Sex Roles, 83*(7–8), 453–473. https://doi.org/10.1007/s11199-019-01116-x

Cohen, J. (1992). Things I have learned (so far). In *Annual Convention of the American Psychological Association*, 98th, Aug, 1990, Boston, MA, US. Presented at the aforementioned conference. American Psychological Association.

de Lemus, S., Moya, M., & Glick, P. (2010). When contact correlates with prejudice: Adolescents' romantic relationship experience predicts greater benevolent sexism in boys and hostile sexism in girls. *Sex Roles, 63*(3-4), 214–225.

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men?: An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly, 15*(2), 203–216. https://doi.org/10.1111/j.1471-6402.1991.tb00792.x

Giatti, L., do Camelo, L., Rodrigues, J. F., & Barreto, S. M. (2012). Reliability of the Macarthur scale of subjective social status -Brazilian longitudinal study of adult health (Elsa-Brasil). *BMC Public Health, 12,* 1096. https://doi.org/10.1186/1471-2458-12-1096

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*(2), 109–118. https://doi.org/10.1037/0003-066x.56.2.109

Gordon, A. R., Conron, K. J., Calzo, J. P., White, M. T., Reisner, S. L., & Austin, S. B. (2018). Gender expression, violence, and bullying victimization: Findings from probability samples of high school students in 4 US school districts. *Journal of School Health, 88*(4), 306–314. https://doi.org/10.1111/josh.12606

Goudeau, S., Autin, F., & Croizet, J.-C. (2017). Étudier, mesurer et manipuler la classe sociale en psychologie sociale: approches économiques, symboliques et culturelles [Studying, measuring and manipulating social class in social psychology: economic, symbolic and cultural approaches]. *International Review of Social Psychology, 30*(1), 1-19. https://doi.org/10.5334/irsp.52

Gravillon, I. (2022). Pourquoi la question du genre agite-t-elle les adolescents ? [Why questions around gender moves adolescents so much?] *L'école Des Parents*, *N° hors-série*(3), 8–13. https://doi.org/10.3917/epar.hs2.0008

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology, 106*(2), 375–389. https://doi.org/10.1037/a0034679

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Petersen (Eds.), *Girls at Puberty,* (pp. 201–228). https://doi.org/10.1007/978-1-4899-0354-9_10

Hoffman, A. J., Dumas, F., Loose, F., Smeding, A., Kurtz-Costes, B., & Régner, I. (2019). Development of gender typicality and felt pressure in European French and North African French adolescents. *Child Development, 90*(3), e306-e321. https://doi.org/10.1111/cdev.12959

Horn, S. S. (2007). Adolescents' acceptance of same-sex peers based on sexual orientation and gender expression. *Journal of Youth and Adolescence, 36*(3), 363–371. https://doi.org/10.1007/s10964-006-9111-0

Iacoviello, V., Valsecchi, G., Berent, J., Borinca, I., & Falomir-Pichastor, J. M. (2021). The impact of masculinity beliefs and political ideologies on men's backlash against non-traditional men: The moderating role of perceived men's feminization. *International Review of Social Psychology, 34*(1), Article 28. https://doi.org/10.5334/irsp.588

Institut National de la Statistique et des Études Economiques [National Institute of Statistical and Economic Studies]. (2012). *Immigrés et descendants d'immigrés en France.*

Jewell, J. A., & Brown, C. S. (2013). Relations among gender typicality, peer relations, and mental health during early adolescence. *Social Development, 23*(1), 137–156. https://doi.org/10.1111/sode.12042

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*(1), 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology, 88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Jury, M., Bruno, A., & Darnon, C. (2018). Doing better (or worse) than one's parents: Social status, mobility, and performance-avoidance goals. *British Journal of Educational Psychology, 88*(4), 659–674. https://doi.org/10.1111/bjep.12210

Klaczynski, P. A., Felmban, W. S., & Kole, J. (2020). Gender intensification and gender generalization biases in pre-adolescents, adolescents, and emerging adults. *British Journal of Developmental Psychology, 38*(3), 415–433. https://doi.org/10.1111/bjdp.12326

Kleiser, M., & Mayeux, L. (2020). Popularity and gender prototypicality: An experimental approach. *Journal of Youth and Adolescence, 50*(1), 144–158. https://doi.org/10.1007/s10964-020-01344-5

Kleiser Polk, M., & Mayeux, L. (2022). Associations between peer-perceived and self-perceived gender typicality and peer status in early adolescence. *The Journal of Early Adolescence, 43*(3), 320–341. https://doi.org/10.1177/02724316221105604

Koenig, A. M. (2018). Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.01086

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science, 1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Lamer, S. A., Dvorak, P., Biddle, A. M., Pauker, K., & Weisbuch, M. (2022). The transmission of gender stereotypes through televised patterns of nonverbal bias. *Journal of Personality and Social Psychology, 123*(6), 1315–1335. https://doi.org/10.1037/pspi0000390

Leaper, C., & Valin, D. (1996). Predictors of mexican american mothers' and fathers' attitudes toward gender equality. *Hispanic Journal of Behavioral Sciences, 18*(3), 343–355. https://doi.org/10.1177/07399863960183005

Lobel, T. E., Bempechat, J., Gewirtz, J. C., Shoken-Topaz, T., & Bashe, E. (1993). The role of gender-related information and self-endorsement of traits in preadolescents' inferences and judgments. *Child Development, 64*(4), 1285. https://doi.org/10.2307/1131340

Lobel, T. E., Nov-Krispin, N., Schiller, D., Lobel, O., & Feldman, A. (2004). Gender discriminatory behavior during adolescence and young adulthood: A developmental analysis. *Journal of Youth and Adolescence, 33*(6), 535–546. https://doi.org/10.1023/b:joyo.0000048067.83738.e6

Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods, 9*(2), 147–163. https://doi.org/10.1037/1082-989x.9.2.147

Mayeux, L., & Kleiser, M. (2019). A gender prototypicality theory of adolescent peer popularity. *Adolescent Research Review, 5*(3), 295–306. https://doi.org/10.1007/s40894-019-00123-z

McGuire, L., Palmer, S. B., & Rutland, A. (2021). Children's and adolescents' evaluations of peers who challenge their group: The role of gender norms and identity. *Social Development, 31*(2), 423–437. https://doi.org/10.1111/sode.12546

Meijs, M. H., Lammers, J., & Ratliff, K. A. (2015). Gender stereotype-inconsistent acts are seen as more acceptable than stereotype-consistent acts, if they are clever. *Social Psychology, 46*(5), 291–305. https://doi.org/10.1027/1864-9335/a000244

Mishra, S., & Kray, L. J. (2022). The mitigating effect of desiring status on social backlash against ambitious women. *Journal of Experimental Social Psychology, 102*, 104355. https://doi.org/10.1016/j.jesp.2022.104355

Mulvey, K. L., & Killen, M. (2014). Challenging gender stereotypes: Resistance and exclusion. *Child Development, 86*(3), 681–694. https://doi.org/10.1111/cdev.12317

Mulvey, K. L., Rizzo, M. T., & Killen, M. (2015). Challenging gender stereotypes: Theory of mind and Peer Group Dynamics. *Developmental Science, 19*(6), 999–1010. https://doi.org/10.1111/desc.12345

Navarro, R., Larrañaga, E., & Yubero, S. (2015). Gender identity, gender-typed personality traits and school bullying: Victims. *Bullies and Bully-Victims. Child Indicators Research, 9*(1), 1–20. https://doi.org/10.1007/s12187-015-9300-z

Owuamalam, C. K., Rubin, M., & Spears, R. (2018). Revisiting 25 years of system motivation explanation for system justification from the perspective of social identity model of system attitudes. *British Journal of Social Psychology, 58*(2), 362–381. https://doi.org/10.1111/bjso.12285

Plaza, M., Boiché, J., Brunel, L., & Ruchaud, F. (2016). Sport = male… but not all sports: Investigating the gender stereotypes of sport activities at the explicit and implicit levels. *Sex Roles, 76*(3–4), 202–217.

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereo-typical impression management. *Journal of Personality and Social Psychology, 74*(3), 629–645. https://doi.org/10.1037/0022-3514.74.3.629

Rudman, L. A., Moss-Racusin, C. A., Glick, P., & Phelan, J. E. (2012a). Reactions to vanguards: Advances in backlash theory. *Advances in Experimental Social Psychology, 45,* 167–227. https://doi.org/10.1016/b978-0-12-394286-9.00004-4

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012b). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology, 48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Sarlet, M., & Dardenne, B. (2012). Le sexisme bienveillant comme processus de maintien des inégalités sociales entre les genres [Benevolent sexism as a gender inequality maintaining process]. *L'année Psychologique, 112*(03), 435–463. https://doi.org/10.4074/s0003503312003053

Schroeder, K. M., & Liben, L. S. (2020). Felt pressure to conform to cultural gender roles: Correlates and consequences. *Sex Roles, 84*(3–4), 125–138. https://doi.org/10.1007/s11199-020-01155-9

Sirin, S., McCreary, D., & Mahalik, J. (2004). Differential reactions to men and women's gender role transgressions: Perceptions of social status, sexual orientation, and value dissimilarity. *The Journal of Men's Studies, 12*(2), 119–132. https://doi.org/10.3149/jms.1202.119

Sommet, N., & Morselli, D. (2021). Keep calm and learn multilevel linear modeling: A three-step procedure using SPSS, stata, R, and mplus. *International Review of Social Psychology, 34*(1), 555. https://doi.org/10.5334/irsp.555

Sommet, N., Quiamzade, A., Jury, M. A., & Mugny, G. (2015). The student-institution fit at university: Interactive effects of academic competition and social class on achievement goals. *Frontiers in Psychology, 6.* https://doi.org/10.3389/fpsyg.2015.00769

Stephens, N. M., Markus, H. R., & Phillips, L. T. (2014). Social class culture cycles: How three gateway contexts shape selves and fuel inequality. *Annual Review of Psychology, 65*(1), 611–634. https://doi.org/10.1146/annurev-psych-010213-115143

Stephens, N. M., Markus, H. R., & Townsend, S. S. M. (2007). Choice as an act of meaning: The case of social class. *Journal of Personality and Social Psychology, 93*(5), 814–830. https://doi.org/10.1037/0022-3514.93.5.814

Sullivan, J., Moss-Racusin, C., Lopez, M., & Williams, K. (2018). Backlash against gender stereotype-violating preschool children. *PLoS ONE, 13*(4), e0195503. https://doi.org/10.1371/journal.pone.0195503

van den Berg, Y. H. M., Lansu, T. A. M., & Cillessen, A. H. N. (2020). Preference and popularity as distinct forms of status: A meta-analytic review of 20 years of research. *Journal of Adolescence, 84*(1), 78–95. https://doi.org/10.1016/j.adolescence.2020.07.010

Verniers, C., Bonnot, V., Darnon, C., Dompnier, B., & Martinot, D. (2015). How gender stereotypes of academic abilities contribute to the maintenance of gender hierarchy in higher education. In K. Faniko, F. Lorenzi-Cioldi, O. Sarrasin, & E. Mayor (Eds.), *Men and women in social hierarchies* (pp. 26–38). Routledge.

Verniers, C., Martinot, D., & Dompnier, B. (2016). The feminization of school hypothesis called into question among junior and high school students. *British Journal of Educational Psychology, 86*(3), 369–381. https://doi.org/10.1111/bjep.12111

Wilcox, R. R. (1992). Why can methods for comparing means have relatively low power, and what can you do to correct the problem? *Current Directions in Psychological Science, 1*(3), 101–105. https://doi.org/10.1111/1467-8721.ep10768801

Xiao, S. X., Hoffer, A., Martin, C. L., & Jenkins, D. L. (2022). Early adolescents' gender typicality and depressive symptoms: The moderating role of parental acceptance. *The Journal of Early Adolescence, 42*(6), 822–840. https://doi.org/10.1177/02724316221078832

Yu, C., Zuo, X., Blum, R. W., Tolman, D. L., Kågesten, A., Mmari, K., De Meyer, S., Michielsen, K., Basu, S., Acharya, R., Lian, Q., & Lou, C. (2017). Marching to a different drummer: A cross-cultural comparison of young adolescents who challenge gender norms. *Journal of Adolescent Health, 61*(4), S48–S54. https://doi.org/10.1016/j.jadohealth.2017.07.005

**Ethan Meimoun**   is a Ph.D. student in social psychology at Université Paris Cité. His research focuses on reactions to gender norm violations and how gender stereotypes throughout development stages contribute to the maintenance of gender inequalities.

**Virginie Bonnot**   is a Professor in Social Psychology at Université Paris Cité and currently the Director of Cité du Genre, an interdisciplinary research institute in gender studies. Her research is aimed at understanding the role of ideologies (e.g., stereotypes, neoliberal ideology, intensive mothering ideology) in justifying and maintaining social hierarchies. Her work is mainly focused on gender issues.

**Julia Berenguer**   is a doctoral student in social psychology at Université Paris Cité. Her research focuses on the role of ideological factors (heteronormativity, homonormativity, neoliberal ideology) in maintaining inequalities and discrimination against LGBTQIA+ people.

**Cristina Aelenei**   is an Associate Professor in Social Psychology at Université Paris Cité. Her research focuses on understanding how ideological and structural dynamics (e.g., stereotypes, cultural mismatch, backlash effect, glass cliff) contribute to maintaining existing social hierarchies.