

RESEARCH ARTICLE

The weight of school grades: Evidence of biased teachers' evaluations against overweight students in Germany

Mona Dian¹, Moris Triventi^{2*}

1 Department of International Business Administration, University of Applied Sciences Worms, Worms, Rhineland-Palatinate, Germany, **2** Department of Sociology and Social Research, University of Trento, Trento, Trentino-Alto Adige, Italy

* moris.triventi@unitn.it



OPEN ACCESS

Citation: Dian M, Triventi M (2021) The weight of school grades: Evidence of biased teachers' evaluations against overweight students in Germany. PLoS ONE 16(2): e0245972. <https://doi.org/10.1371/journal.pone.0245972>

Editor: Sharon Mary Brownie, Waikato Institute of Technology, NEW ZEALAND

Received: September 5, 2020

Accepted: January 11, 2021

Published: February 8, 2021

Copyright: © 2021 Dian, Triventi. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort 3 – 5th Grade, doi:[10.5157/NEPS:SC3:8.0.0](https://doi.org/10.5157/NEPS:SC3:8.0.0). From 2008 to 2013, NEPS data were collected as part of the Framework Programme for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, the NEPS survey is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. Data and documents are released for academic research and are available for registered

Abstract

Discrimination and prejudice against overweight people is common in Western societies. In this article we aim to understand whether these attitudes reverberate into the school setting, by investigating whether teachers grade overweight students more severely than comparable normal weight students. By relying on the Attribution-Value Model of Prejudice (AVMP) and previous studies, we test a series of hypotheses using data from the German National Educational Panel Study (NEPS SC3) on a sample of students enrolled in the 7th grade (lower secondary education). We used hierarchical ordered logit regression to assess whether overweight and obese students receive systematically lower grades by their teachers in German and mathematics, adjusting for subject-specific competences measured with a standardized test, and a rich set of socio-demographic and socio-psychological students' characteristics (e.g. the "big five"). Results suggested that overweight and obese students were more severely graded in both subjects. The penalty for overweight students, and especially for obese students, was slightly larger in German and in the lowest part of the grade distribution. There was also indication of heterogeneous penalties by gender, with overweight male students being especially penalized in math. Possible ways to help teachers in assigning grades in a fairer way are discussed at the end.

Introduction

Discrimination against overweight people in the labor market and in social life is well documented and increased over time [1, 2]. In the labor market, overweight people are less likely to be hired [3] and are at risk to receive a lower starting salary in comparison to normal weight people [4]. In social life, overweight people are more often bullied [5] and receive less friendship requests in social networks on the internet [6].

In this article we aim to understand whether obese and overweight people constitute stigmatized and penalized social categories in a specific context that has attracted comparatively less attention, namely the school. In particular, we are interested in how teachers evaluate the

users who have signed a data use agreement with the LfBi (fdz@lifbi.de). This release is uniquely identified by one DOI. The sub-datasets used in this research are: SC3_CohortProfile_D_8-0-0, SC3_pTarget_D_8-0-0, SC3_xTargetCompetencies_D_8-0-0, and SC3_pParent_D_8-0-0. The third party data set "Percentile data files with LMS values: BMI-for-age charts, 2 to 20 years, LMS parameters and selected smoothed BMI (kilograms/meters squared) percentiles by sex and age", version of May 30, 2000, is provided by the National Center for Health Statistics (Hyattsville). It is publicly available from https://www.cdc.gov/growthcharts/percentile_data_files.htm. Interested researchers can replicate the study by directly obtaining and merging the data from the third parties and following the protocol in our methods section.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

scholastic performance of overweight students compared to normal weight students in German lower secondary school.

In most previous studies, school grades are mainly intended as indicators of academic proficiency, assuming that body mass diminishes performance in school via poor health and a resulting reduction in study effectiveness [7]. Nonetheless, this perspective overlooks the fact that overweight is a socially sanctioned feature of the physical body [8] and that lower grades may result from social pathways involving discrimination and stigma by teachers [9].

The literature on teachers' grading in school posits that grades are not always a fair measure of subject-specific competences because they can also reflect teachers' attitudes about the students [10, 11]. As for all common people, teachers' attitudes are likely to be influenced by attributes that are not relevant to cognitive outcomes such as student gender [12], migration background [13], and body size [9]. These attributes might elicit stereotypes about the students' competence and thus lead to discrimination in grading [12–14].

While gender and migration background are commonly considered to be unchangeable, overweight is perceived as a controllable attribute that is culturally undesirable at the same time [15]. The Attribution-Value Model of Prejudice (AVMP) argues that attributions of controllability and cultural value of the attribute the stereotype concerns are required to be simultaneously present to lead to prejudice. When weight is believed to be controlled volitionally, overweight people can be held responsible for their situation and be treated as if they deserved the blame [16]. But overweight is only considered a negative attribute if there is a cultural preference for thinness. Without the negative cultural value people would not be blamed for overweight [15, 17].

It fits to the controllability aspect of the AVMP that overweight people are perceived as lacking personal control, being lazy, less conscientious [18, 19], and less disciplined [20]. As control and discipline are important for competence development [21], stereotypes about overweight people also include stupidity [18], slowness in thinking [22], lower competence and sloppiness [19]. In comparison to prejudice against other groups, negative attitudes about overweight people are widespread and expressed more openly. Social desirability does not hold opinions back as much as in the case of racism or homophobia [23, 24]. Indeed, a large part of the population displays prejudice against overweight people referring to their intelligence and willingness to perform [18, 23, 25]. Some studies suggest that also school teachers might have prejudice against overweight people [26, 27].

The cultural preference for thinness in Western countries is also well documented. In all socioeconomically highly developed societies thin bodies have the highest ratings of attractiveness [28]. Weight loss up to thinness or underweight is the aim of many women who are uncomfortable with their body [29]. Therefore, there is ample evidence suggesting that attributions of controllability and negative cultural value are met in Western societies, especially regarding women's bodies. A likely result from prejudiced attitudes and beliefs is discriminatory behavior [4], and indeed, overweight students receive lower grades than their peers [30, 31].

Against this background, our research questions are: Are overweight and obese students graded more severely by their teachers than equally competent normal weight students? Moreover, does the potential grading bias vary by subject and gender? The answers to these questions are relevant for at least two reasons.

The first reason is fairness for the sake of the effect of grades on students' motivation to learn and on their well-being [32]. Students who felt treated unfairly by their teachers are more stressed and are more likely to have psychological problems [33]. The second reason is fairness for the sake of the validity of grades in signaling competence and mastery of knowledge in specific subjects [34]. Grades are relevant signals in educational-related decisions [35,

36] and thereafter in application processes, e.g. when future employers screen potential job applicants and when the university admits new students [10, 37]. Grades of overweight students can result from a mix of prejudice and objective ratings of competence, leading to the underestimation of their academic proficiency and potential for skill development. If this is the case, employers and institutions risk to unintentionally disregard qualified employees.

We investigate the potential grading bias by using data of the German National Educational Panel Study with students attending the seventh grade [38]. By this way, we aim to contribute to the literature by using recent high-quality nationally representative data on lower secondary education in Germany. The German context differs from the US context in two ways. First, German students are placed in educational tracks already in secondary education [see 35 for more information on the German tracking system]. Teachers' grades are particularly important for students' track placement, which in turn is crucial for later educational transitions. Second, while in the US obesity is comparatively widespread (37.3% in 2016), the incidence of this phenomenon in Germany is lower (25.7% in 2016) [see 39 for prevalence rates]. It could be that obesity is a less visible characteristic in the German context, leading to a stronger stigma as the society may be less aware of discrimination against overweight people [40].

The paper is organized in the following way: We set the basis for the research hypotheses by outlining results from the previous literature about grading bias towards overweight people. After presenting the German National Educational Panel Study data, variables and our methods, we report the empirical findings of the hierarchical ordinal logistic regression models. The article concludes with a discussion of the results and possible implications for teachers and schools.

Teachers, students' weight and grading

In the school context, the beliefs of teachers about students' competences and work ethics matter and overweight students face difficulties that can have consequences for their future careers. In a study of Neumark-Sztainer and colleagues [26], half of the school staff thought that overweight was self-caused. About one fifth believed that overweight people were less tidy and less successful at work. Fitting to the prejudicial attitudes, the teachers' perceptions of competences in math and reading decrease with increasing BMI of the students [41]. Teachers say that overweight students are burdensome to have in the classroom [27] and, for this reason, might judge them more severely than normal weight students.

Some studies reveal a relationship between student BMI and school grades: After controlling for sociodemographic factors, overweight students have a lower grade point average than normal weight students [42], a lower average grade in the main subjects [30], and are less likely to report having mostly higher grades [31]. However, these studies do not necessarily prove teacher bias as they did not control for student level of subject-specific competence in the statistical models. Therefore, the lower grades possibly express a lower competence of overweight compared to normal weight students, as suggested by a few studies about brain and memory functioning [43–45]. However, only a very low number of studies find a connection between overweight and lower competence measured by standardized test scores in a real-world setting [42, 46]. Mostly, competences do not differ between overweight and normal weight students after adjusting for sociodemographic characteristics [47–50].

Other studies, using American samples, go further and partly account for competences as an alternative cause of grade differences to discrimination. Sabia [51] reports an association between a higher BMI and a lower grade point average controlling for general intelligence in a sample of 14 to 17 years old white females. But holding general intelligence constant may not rule out differences in subject-specific competences as students also need to have enough discipline and motivation to study [52]. MacCann and Roberts [53], controlling for competence in

math and vocabulary, found a lower average grade in various school subjects among obese students compared to normal weight students in the US. Competence scores for science and social studies were, however, not included and could have still biased measures of grade differences. Research of this quality is not available for countries with a smaller share of overweight people in the population.

Despite some heterogeneity in the research findings, most works found that BMI is negatively related to academic achievement. Moreover, several studies report that teachers might have prejudice against overweight students. Following these insights, and regarding our research question if overweight and obese students are graded more severely, we expect that

overweight and obese students get lower grades than normal weight students, once adjusting for subject-specific competence and other individual characteristics

(Hyp. 1).

The second research question asked whether teachers' bias varies by subject. In this work, we focus on two key subjects in secondary education, namely German and mathematics. If discrimination against overweight students is just an inherent feature of the teachers in their role of graders, we could predict a similar degree of bias in both subjects. If instead the way in which students are commonly tested and the standards usually adopted by the teachers play a role as well, we could expect some differences in the grading bias across subjects. In mathematics, most exams are based on exercises with a precise solution, whereas in German teachers assess students more often with oral exams and open-ended questions, in which subjective elements have more room to affect the evaluation process. If these characteristics matter, we should expect that

overweight and obese students are graded less generously than comparable normal weight students especially in German and less in mathematics

(Hyp. 2).

Our third research question asks whether teachers' grading bias differs between males and females. It has been argued that obesity reduces the perception of femininity by others but not as much as it reduces the perception of masculinity [54]. In subjects related to reading and writing activities, for example English, femininity is a favorable attribute, whereas in mathematics it is often considered an unfavorable attribute. Obese girls are regarded less feminine and have a smaller femininity advantage in English classes than other girls. Consequently, in comparison to normal-weight girls, it is likely they face a specific penalization [9]. Indeed, previous studies in the US found that obese girls face higher grading penalization compared to normal-weight girls in English rather than in mathematics. For boys, no association between obesity and academic performance was found at all [9]. Following this logic, we hypothesize that

the weight penalty regarding grades is higher among girls than among boys, especially in German

(Hyp. 3).

Analytical design

Data. We make use of the German National Educational Panel Study (NEPS)–Starting Cohort 3, which provides information on students attending the grade 7 in 2012, who are

repeatedly interviewed in the subsequent years [38]. The longitudinal data set is used in a cross-sectional manner by analyzing data from wave 3, because the variables of interest are not measured repeatedly in every wave. Education- and school-related variables as well as competences and socio-demographics are assessed via paper-and-pencil interviews (PAPI) or provided by official student lists. Information on the parents is collected using computer-assisted telephone interviews (CATI) [55, 56].

We restricted the sample by excluding students without matches in the parent and institution data set. 47% of the cases from the original target data set were deleted because they had no matches (institution or parent data). After this operation, we explored the data and found that while the incidence of missing values is null or modest for several socio-demographic variables, the share of missing values is higher for the outcomes (around 7%) and especially for the main independent variable (25%) (see [S1 Table](#)). If we had performed a complete information case analysis, we would end up with 2,142 students. Since relying on listwise deletion would lead to underpowered and possibly biased estimates, we applied a multiple imputation (MI) procedure to impute the missing values [57], which allowed us to rely on an analytical sample of 3,754 cases. Under the assumption of random missing values conditional on the covariates, MI is able to minimize the bias, maximize the use of available information and obtain appropriate estimates of uncertainty around the point estimates [58, 59]. To impute the missing values, we followed the canonical three steps:

1. Imputation phase: We filled the missing data with estimated values and created a complete dataset using the chained equations/MICE approach, which uses a separate conditional distribution for each imputed variable [60]. We relied on multinomial logit to model student BMI, ordinal logit to model teacher grades and parental educational level (ISCED), a linear OLS to model students' test scores, parental occupational status (ISEI), psychological traits ("big five"), and the indexes measuring the attachment to school and homework duration. The other variables (age, gender, native language and school region, type of secondary school attended) do not present any missing value and therefore are used in the imputation models as additional predictors. Considering the amount of missing values and following recommendations from recent literature [61], we generated 50 imputations (completed datasets) under our chosen imputation model.
2. Analysis Phase: We applied the hierarchical ordinal logistic regression models (see below) to analyze teachers' grading in each of the 50 complete data sets.
3. Pooling Phase: The parameter estimates (coefficients and standard errors) obtained from each analyzed data set are then combined for inference using the so-called 'Rubin's rules', which take into account properly the variation between imputations [57, 58].

Variables. The variables come from several sub-data sets of the NEPS. The sub-data sets contain information from the student lists, the students and parents, and standardized competence tests. In the case of students, all time-variant variables were measured in the third wave while the time-invariant variables (gender, native language) were recorded in the first wave. The information on parents is measured once in the first wave and updated in subsequent waves in case of changes.

The grades in mathematics and German are self-reported by the students. Taking into consideration the highly skewed distribution of the original variable and to maintain enough statistical power for the analysis, we recoded the original grades into four categories: *low* (poor, failing, and passing), *medium-low* (satisfactory), *medium-high* (good) and *high* (very good).

Alternative but reasonable classifications lead to substantially similar results to the ones presented here.

Body mass index (BMI) categories were built using the BMI-for-age growth charts of the National Center for Health Statistics [62]. The charts provide BMI percentiles for children grouped by gender and age in months, measured in the middle of the months. We assigned the BMI percentiles to children with the respective age in months rounded down. The BMI is calculated using the formula weight divided by height in meters squared [63]. Height and weight were self-reported by the students. Biologically implausible values of weight, height, and BMI were excluded using an external reference table [64]. Following previous contributions [63], the variable BMI is classified in four qualitatively distinct weight types: 1) Underweight (lower than the 5th percentile); 2) Normal weight (BMI greater than or equal to the 5th percentile and lower than the 85th percentile); 3) Overweight (higher than or equal to the 85th percentile but lower than the 95th percentile); 4) Obese (higher than or equal to the 95th percentile). Even though the literature highlights that BMI is not a perfect measure of body fatness since it does not account for the ways in which body composition differs by individuals' characteristics such as muscularity, sex, and age [65], these problems are less marked in our context. Indeed, the growth charts we used are age- and gender-specific; moreover, muscularity should not be a big problem regarding children younger than 16 years old.

We control for subject-specific competence as an alternative cause of lower grades instead of teacher bias. Subject-specific competence is measured by weighted maximum likelihood estimates (WLE), which are corrected for the position of the test domain in the test book. The resulting score is the best estimate of each respondent's observed answers to several items belonging to a domain [66, 67]. The domains are mathematics and reading competence. A negative WLE score signals a below-average competence, a score of 0 an average competence, and a positive score an above-average competence.

Aside from competences, other variables need to be accounted for because they can influence both grades and weight. Parents' socioeconomic status (SES) is negatively associated to overweight [68] and positively related to grades [69]. Children with a higher conscientiousness show less risky health-related behavior [70] and have higher grades [71]. In general, personality traits are linked to over- and underweight [72] as well as to academic achievement [73]. Furthermore, migration background can be a confounder, since children with a migration origin are more often overweight than other children [74] and they also have a lower academic performance [13]. We control for native language as a proxy for migration background, which is not available in the data. Lastly, attitudes about learning and objective behaviors could be the reason why overweight students receive lower grades. To rule out that they put less effort in learning activities than normal-weight students, we control for attitudes using attachment to school and objective behaviors using homework duration. The complete list of control variables and their description is presented in [Table 1](#).

Methods

To identify teachers' grading bias against overweight and obese students with observational data we apply a so-called "grade-equation" approach [75] in which teachers' marks are expressed as a function of the categories of body mass index (BMI), adjusted for students' subject-specific competences measured with a standardized test (TESTSCORES), and an additional set of covariates capturing students' heterogeneity in socio-demographic characteristics (SOCIODEMO), attitudes (ATTITUDES) and behavior (BEHAV), which may be correlated with BMI and with teachers' grades as well. With this strategy it is possible to assess to what

Table 1. Characteristics of the control variables.

Variable	Item from NEPS	Type of Answer
Reading competence (test scores)	Booklet with 29/30 reading competence items	WLE score corrected for the test position, computed by the NEPS staff
Mathematics competence (test scores)	Booklet with 23 mathematics competence items	WLE score corrected for the test position, computed by the NEPS staff
Gender	School's list: gender child	0—female 1—male
School region	Please enter the postal code of your school.	1—former West Germany 2—former East Germany incl. Berlin
Age in years	When were you born?	Open question
Native language	Now let's talk about your mother tongue: which language did you learn as a child in your family?	0—German only 1—other native language
Parental SES	International Socio-Economic Index of occupational status (ISEI) (Highest most recent ISEI among pairs, most recent ISEI if single parent)	16 to 90
Parental ISCED	International Standard Classification of Education (ISCED) (Highest most recent ISCED among pairs, most recent ISCED if single parent)	1–2 or lower 2–3b 3–3a & 3c 4–4a & 5b 5–5a & 6
School type	Sample: type of school	0—Hauptschule 1—Realschule 2—Gymnasium 3—schools with different tracks
Openness	I do not care much about arts. [R] & I have an active imagination, I am an imaginative person.	9 step sum scale from 2 items: 1—does not apply at all to 5—applies completely
Conscientiousness	I am easy-going and tend to be a bit lazy. [R] & I am thorough.	
Extraversion	I tend to be cautious, reserved. [R] & I am out-going and sociable.	
Agreeableness	I tend to be critical of other people. [R] & I trust other people easily, I believe in the goodness in people.	
Neuroticism	I am relaxed and don't get easily stressed. [R] & I am considerate, sensitive.	
Attachment to school	Going to school for a long time is a waste of time.	1—completely disagree to 5—completely agree
Homework duration	How much time do you normally spend on your homework and learning for school?	1—less than half an hour per day 2—about half an hour to 1 hour per day 3—about 1 to 2 hours per day 4—about 2 to 3 hours per day 5—about 3 to 4 hours per day 6—more than 4 hours per day

<https://doi.org/10.1371/journal.pone.0245972.t001>

extent overweight and obese students receive on average different marks compared to their most similar normal weight peers, in terms of demonstrated performance in a specific subject and other relevant traits. This strategy assumes that standardized test scores are a less subjective measure of individual school-related competences than teachers' grades and could work as a yardstick to assess whether and how much factors other than student ability are taken into account by teachers when attributing grades to students [76]. For each school subject we estimated two model specifications, which can be succinctly described as follows:

$$\text{Model 1 : } \text{GRADES}_j = f(\text{BMI}, \text{TESTSCORE}, \text{SOCIODEMO})$$

$$\text{Model 2 : } \text{GRADES}_j = f(\text{BMI}, \text{TESTSCORE}, \text{SOCIODEMO}, \text{ATTITUDES}, \text{BEHAV})$$

Model 1 is intended to measure the total teacher grading bias, which we can interpret as the average deviation of teachers' evaluations from a more objective and standardized form of competence assessment: According to the models, if students with comparable competence scores receive different grades, the grade difference identifies teachers' bias. Model 2 includes possible mediators that can explain the total teacher grading bias. If the included mediators are sufficient to capture all the possible relevant mechanisms by which teachers' judgements are distorted, then we can interpret the coefficients corresponding to the BMI variable as the result of (taste or statistical) discrimination processes [76].

In the last part of the analysis, we investigate whether the grading bias due to overweight and obesity is heterogeneous across boys and girls, by introducing an interaction term between gender (SEX) and BMI. For the sake of simplicity, we omitted the main effects of the interacted variables in the notation, which are included in all models to avoid unbiased estimates, as suggested by Brambor et al. [77].

$$\text{Model 3 : } \text{GRADES}_j = f(\text{BMI} \times \text{SEX}, \text{TESTSCORE}, \text{SOCIODEMO}, \text{ATTITUDES}, \text{BEHAV})$$

Given the nature of our outcome variables and the data/sampling structure, we use three-level ordinal hierarchical (random-intercept) regression models in which individuals are nested in courses that are nested in schools. By this way we allow the standard errors to be dependent within classes and schools, taking into account that students within the clusters are more similar to each other than between the clusters, resulting from exposure to similar contexts. The statistical models are estimated separately for German and mathematics, and in each model, the subject-specific test score is included as a control variable. Accounting for potential non-linearities in the relationship between the competence variables and the outcomes does not alter the results in a significant way. As described above, each model was estimated on each of the 50 datasets generated from our imputation model and then combined in order to take into account the variability between the estimates across the imputed data.

Results

Descriptive statistics and bivariate analysis

The analytical sample is constituted by 51% of males and 20% of the students speak other native languages beneath German. Among the types of secondary schools, 61% were enrolled in the Gymnasium (high-level track), 20% in the Realschule (medium-level track), 5% in the Hauptschule (lowest-level track), and 14% in schools with different tracks (comprehensive). Regarding the BMI, 80% (n = 1719) of the students have a normal weight, 11% (n = 230) have underweight, 7% (n = 149) have overweight, and 2% (n = 44) are obese.

Looking at the socio-demographic characteristics, we see that overweight and obese students are more often males, come more often from socio-economically disadvantaged families and are more likely to have a non-native background. Moreover, the share of students attending the Gymnasium is much lower among overweight (44%) and obese (36%) students compared to normal weight students (63%) (see [S2 Table](#)). This could be related to the fact that, as we will show below, obese students have lower academic performance than normal weight ones. In terms of psychological traits, overweight and obese students display a lower level of extraversion, conscientiousness, agreeableness, openness, and higher levels of neuroticism than normal weight students. However, given the relatively small share of obese students and

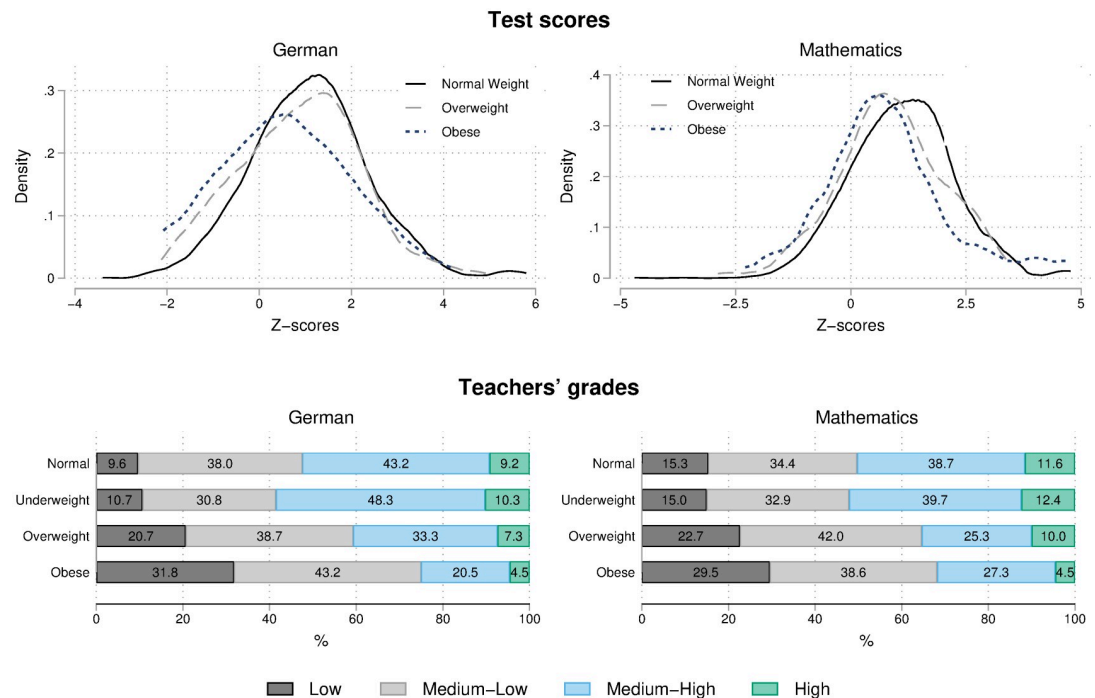


Fig 1. Distribution of test scores and teachers' grades by students' weight category. The upper-left graph shows Kernel density estimates of the distribution of test scores in German, the upper-right graph in mathematics. The lower-left graph shows the distribution of teachers' grades by students' weight category in German, and the lower-right graph in mathematics.

<https://doi.org/10.1371/journal.pone.0245972.g001>

the fact that some of these differences are pretty small, the only statistically significant differences are found on extraversion and conscientiousness. Instead, students with different BMI do not differ significantly in terms of attachment to school and time spent on homework.

We now proceed by describing the bivariate relationship between students' BMI and academic performance in lower secondary school in more detail. In the upper panel of Fig 1, the kernel density estimates of the distribution of students' test scores in German (left) and mathematics (right) indicate that, on average, normal weight students have a higher performance than overweight ones, who in turn perform better than obese students. In the lower panel of Fig 1 we report the distribution of teachers' grades in German (left) and mathematics (right) according to the categories of BMI.

The patterns are similar across the two subjects: The incidence of children who received a low grade is largest among obese students, decreases but stays rather high among overweight students and is lower among normal weight students. The opposite pattern is found when looking at high grades, but in this case the difference between normal weight and obese students is more pronounced as when comparing normal weight to overweight students. Underweight students usually perform very similarly compared to normal weight students, either when academic performance is measured by teachers' grades or standardized tests.

Analysis of the grading bias against overweight and obese students

To assess whether overweight and obese students are less generously graded by their teachers compared to equally performing normal-weight students, we estimated a series of hierarchical ordered logistic regression models in which teachers' grades are modelled as a function of students' BMI categories, subject-specific performance in standardized tests, and a set of relevant

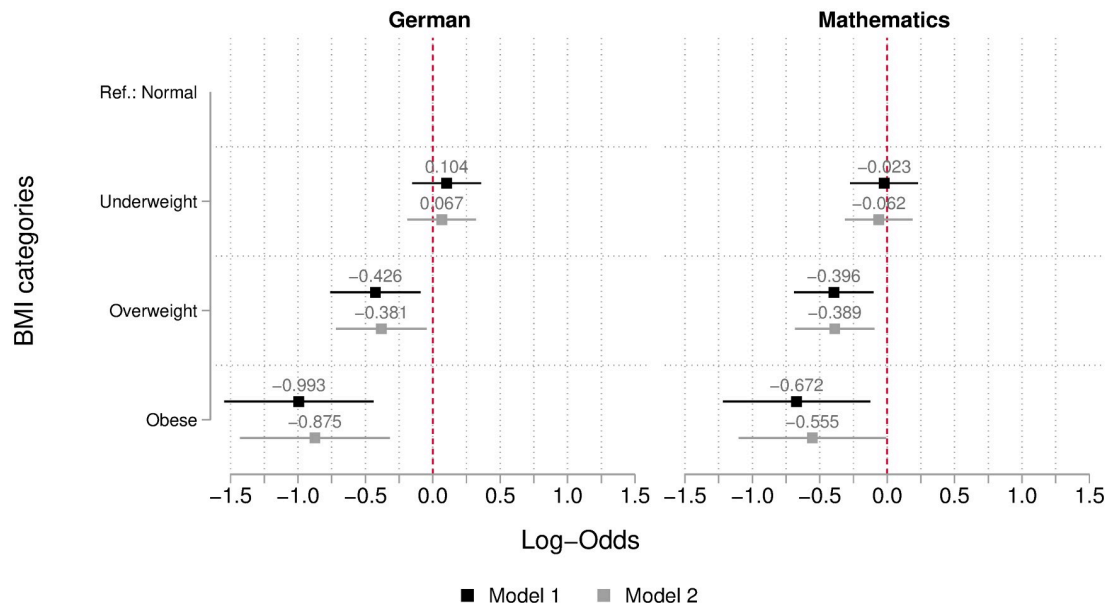


Fig 2. Teachers' grading bias (measured in log-odds) according to students' BMI. The graph shows log-odds ratios from the 3-level hierarchical ordered logit model on mathematics grades (left) and German (right). Bars represent 95% confidence intervals. Model 1 adjusts for students' competence level and socio-demographic variables, while model 2 also includes students' psychological traits, and school-related attitudes and behavior. $N = 3,754$. Analysis conducted on 50 multiply imputed datasets.

<https://doi.org/10.1371/journal.pone.0245972.g002>

individual characteristics. To correct for the missing values on relevant covariates, we conducted the analysis on 50 multiply imputed datasets. In Fig 2 we report the logit coefficients and 95% confidence intervals of the BMI categories of interest compared to normal weight students (the reference category). The complete models are reported in S3 Table. We report two model specifications: Model 1 adjusts for students' competence level and socio-demographic variables, while Model 2 also includes students' psychological traits, and school-related attitudes and behavior. However, since the two specifications lead to only minor differences in the estimated coefficients, in our comments we will focus on the results of the fully specified model.

The results indicate that underweight students are graded by their teachers very similarly to normal weight students: The estimates are close to zero and not statistically significant at the 95% confidence level. Conversely, overweight and obese students receive on average lower grades by their teachers than normal weight students with the same level of subject-specific competence. The disadvantages of obese students are larger than those of overweight students and are slightly more pronounced in German than in mathematics.

While relying on the log-odds gives us a synthetic picture of the overall relationship between BMI and teachers' grading, it prevents us to quantify the strength of this relationship in a meaningful way. For this purpose, we present average partial effects of being underweight, overweight, or obese on the four levels of teachers' grades in German (upper panel) and mathematics (lower panel) in Fig 3 (based on S5 Table). We see that, on average, overweight and obese students are more likely to receive low and medium-low grades, whereas they have lower chances to obtain high or medium-high grades compared to equally competent normal weight students with analogous socio-demographic characteristics, psychological traits, and school-related attitudes and behavior.

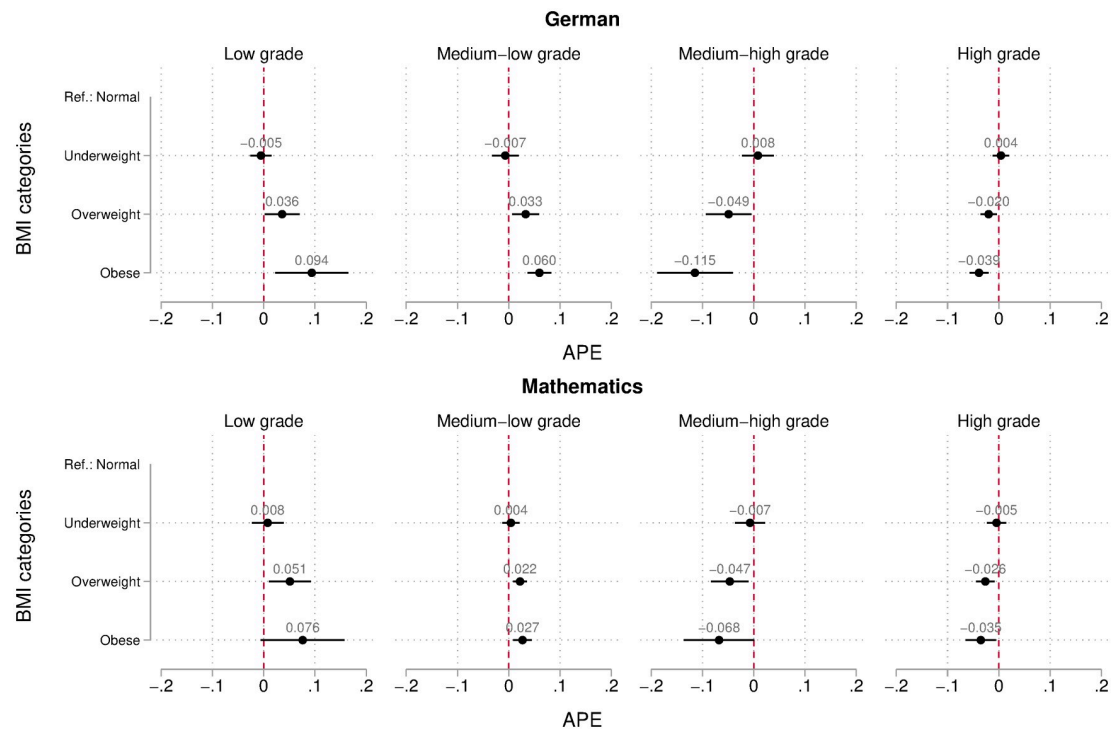


Fig 3. Teachers' grading bias (measured in average partial effects) according to students' BMI. The graph shows average partial effects from the 3-level hierarchical ordered logit model on mathematics grades (upper part) and German (lower part). Bars represent 95% confidence intervals. The models are adjusted for students' competence level, socio-demographic characteristics, psychological traits, and school-related attitudes and behavior. $N = 3,754$. Analysis conducted on 50 multiply imputed datasets.

<https://doi.org/10.1371/journal.pone.0245972.g003>

The largest differences among social groups are found at the lower end of the grade distribution and for medium-high grades. For instance, obese students have on average an 8 (mathematics) to 9 (German) percentage points larger probability of receiving a low grade by their teachers than equally competent normal weight students. They also have a lower probability of receiving medium-high (7–12 percentage points) and high (3–4 percentage points) grades than comparable normal weight students. Overweight is also associated to a penalization in grading, but its magnitude is smaller, ranging between 2 and 5 percentage points. From a qualitative point of view, the estimated differences appear to be larger in German than in mathematics, albeit the confidence intervals around the point estimates are to a large extent overlapped, making it difficult to state that the effect size across subjects actually differs in the reference population.

We estimated additional models in which we included an interaction term between the BMI categories and test scores, with the aim of assessing whether the grading bias differs among students with higher or lower levels of academic proficiency (see [S7 Table](#)). We found that this interaction is not statistically significant and estimated coefficients are substantially small, thus suggesting that the under-evaluation of overweight/obese students takes place at the same extent among low- and high-ability students.

Heterogeneity in the overweight penalty by gender

As a last step, we investigate whether the penalty in grading associated with being overweight differs by gender, as found by a previous study in the US [9]. Given the small number of obese

Table 2. Teachers' grading bias (measured in average partial effects) over the grades distribution, according to students' BMI and gender.

Effect of being overweight/ obese vs normal weight	German				Mathematics			
	Low	Medium-Low	Medium-High	High	Low	Medium-Low	Medium-High	High
Females	-0.006 (0.015)	-0.012 (0.029)	0.010 (0.025)	0.008 (0.019)	0.054 (0.032)	0.026 (0.012)	-0.051 (0.029)	-0.030 (0.015)
Males	0.105*** (0.025)	0.064*** (0.009)	-0.129*** (0.026)	-0.040*** (0.007)	0.046 (0.023)	0.029* (0.012)	-0.044* (0.022)	-0.030* (0.013)
$\Delta_{\text{Male-Female}}$	0.110*** (0.029)	0.075* (0.030)	-0.139*** (0.035)	-0.048* (0.020)	-0.009 (0.039)	0.003 (0.016)	0.007 (0.036)	-0.000 (0.019)

Note: The first two main rows present the average partial effect of being overweight/obese by gender and the last row reports a formal statistical test of the interaction BMI and gender. Standard errors in parentheses, statistical significance levels (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). $N = 2,194$, underweight students are excluded from this analysis. 50 multiply imputed datasets.

<https://doi.org/10.1371/journal.pone.0245972.t002>

students, for this analysis we were forced to merge the categories of overweight and obese students together, in order to gain statistical power. Table 2 presents the results of the models that include an interaction between students' BMI and gender (see S4 Table): More specifically, as suggested by Brambor et al. [77], we report the average partial effect of being overweight/obese versus normal weight on the probability of receiving different grades, among the group of females and males (adjusted for covariates included in the full models). In the last two rows we provide an estimate of the difference between these two effects, its corresponding standard error and information on whether it is statistically significant at different levels of confidence.

Fig 4 (based on S6 Table) reports average predicted probabilities of obtaining a low grade by BMI and gender, derived from the same models. The results are pretty straightforward and in contrast to previous evidence from the US. Being overweight or obese is not related to any penalization in grading among female students. Consequently, the penalty observed in the main analysis is mostly due to teachers' grading of male students. The pattern is particularly pronounced in German: Among males, indeed, overweight/obese students have a much higher probability of receiving a low grade (+11 percentage points) compared to normal weight students with the same subject-specific proficiency, psychological traits, and school-related attitudes and behavior. Overweight/obese males also have a substantially lower probability of getting medium-high (-13 percentage points) and high (-4 percentage points) grades in German.

The penalization in German due to being overweight/obese among males is substantially larger than among females, ranging from 5 to 14 percentage points, and statistically significant across all the four grades. In mathematics, the qualitative pattern differs to some extent: Albeit we detected statistically significant effects of being overweight only among males (ranging between 2 and 4 percentage points), the effect size among females is very similar. The differences in weight penalty by gender are not statistically significant in mathematics. Thus, given the empirical evidence, we cannot reject the null hypothesis of no difference in the grading penalization against overweight/obese students between boys and girls in the reference population.

Discussion and conclusions

From previous studies we knew that overweight students get, on average, lower grades than their peers in school [30, 31] but it was not clear if the differences in grades stem from discrimination or from actual differences in competences. One should expect that students with the same demonstrated competence in a given subject, and same school-related attitudes and

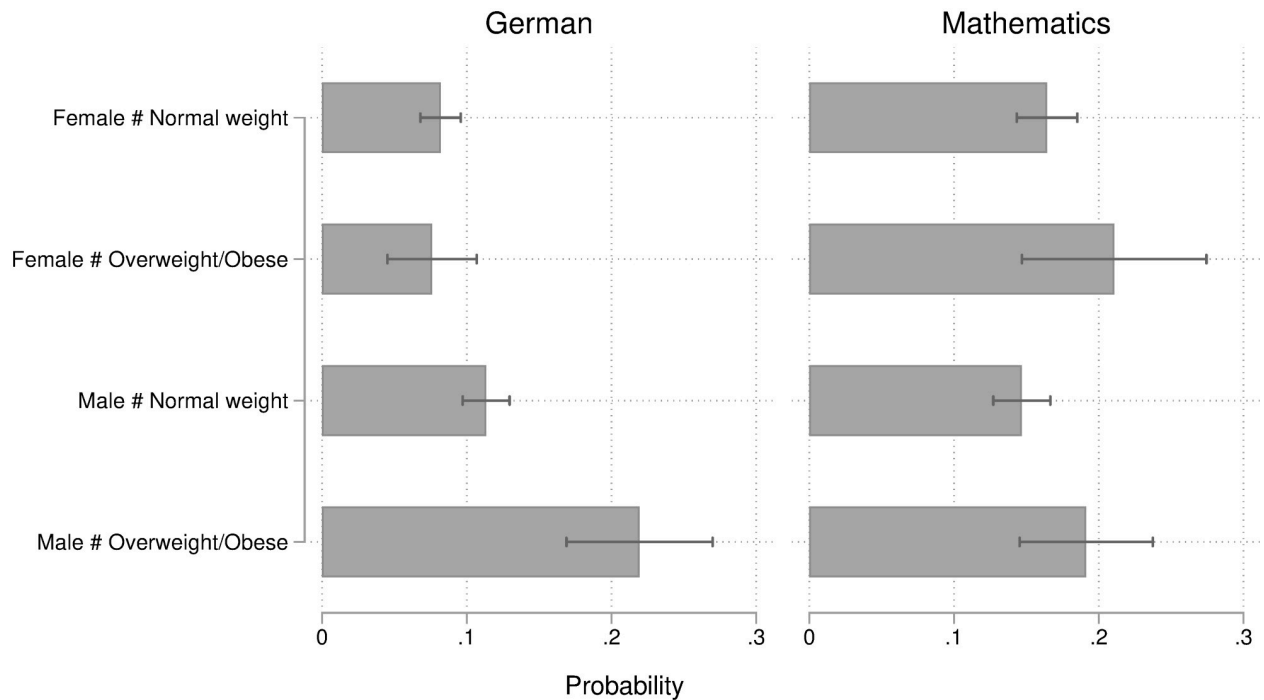


Fig 4. Predicted probabilities of obtaining a low grade by BMI and gender. Predicted probabilities (and 95% confidence intervals) derived from hierarchical ordinal logistic regression models of teachers' grading in German (left graph) and mathematics (right graph), adjusted for students' competence level, socio-demographic characteristics, psychological traits, and school-related attitudes and behavior. $N = 3,754$. Analysis conducted on 50 multiply imputed datasets.

<https://doi.org/10.1371/journal.pone.0245972.g004>

behavior will get similar grades from their teachers if fair grading is in place. We asked if lower mathematics and German grades are attributed to overweight compared to normal weight students, even when they have the same level of subject-specific competence, main psychological traits, and school-related attitudes and behavior. The focus on students in German lower secondary education allows us to concentrate on a key developmental stage within an underexplored country context, in which obesity could lead to stronger stigma due to the comparatively lower incidence of this phenomenon.

By means of hierarchical ordinal logistic regression models we have shown that both, overweight and obese students, receive on average lower grades than comparatively similar normal weight students, thus corroborating our first hypothesis. This pattern is found irrespective of whether we adjust for basic socio-demographic characteristics alone or for a more extensive set of individuals' features including students' personality traits, school-related attitudes and behavior.

While it is difficult to strictly prove "discrimination" within an observational setting, the rich set of control variables and the rather large size of the estimated differences seem to suggest that overweight, and especially obese students, are subject to some forms of penalty in the evaluations they receive by their teachers. The penalization is not only statistically significant, but also substantial. Indeed, the effect size of being obese rather than normal weight is larger than that of being male against being female, a comparison that has attracted much more attention among scholars [e.g., 78]. It is important to note two additional aspects. First, the consequences of being overweight and obese cannot be ascribed to a simple deviance-from-the-norm effect since underweight students do not experience any penalty in teachers' evaluations. Second, considering a wide array of students' traits, attitudes and behavior does not alter

substantially the estimating teacher grading bias, pointing to the possibility that a large part of the total teacher grading bias stems from discrimination and stigmatization processes.

Empirical evidence is instead less clear in supporting our second hypothesis, which states that the grading penalty should be larger in German than in mathematics. In the common wisdom, math teachers are expected to base decisions on grades on a certain score that is linked to pre-determined correct solutions. Consequently, one might argue that math exams leave less room for interpretation regarding which grades the students should receive. Apart from that, teachers may set a variety of exams in the subject German with a different scope for interpretation if a result is right or wrong. While dictation exercises are solved either correctly or incorrectly, there is not a unique correct answer in essays and text comprehension exercises, and therefore the grade could incorporate subjective sources of bias. We found that, from a qualitative point of view, the effect sizes associated to being obese or overweight across subjects follows the expected pattern. From a statistical inference point of view, the confidence intervals around the point estimates are instead rather large and overlapped to a large extent, thus not supporting the hypothesis of differential effects across subjects. Anyway, our findings suggest that—despite mathematics being usually considered a more objective subject—this discipline is also not exempted from teachers' grading bias linked to body appearance.

Additional insights on the heterogeneity in the penalization of overweight and obese students come from the second analysis, in which we investigate whether gender moderates the effect of BMI. Following theoretical arguments on the role of femininity in contemporary Western societies and previous findings from the US, we expected a larger detrimental effect of obesity on grades among females. Contrarily, we found that the more severe grading to which overweight/obese students are exposed in lower secondary school is mainly concentrated among male students, a pattern that is particularly pronounced in the German subject and less in mathematics. Differently from what was observed in the US, in Germany it seems that body appearance linked to negative stereotypes amplifies the stricter grading standards to which males are already exposed in comparison to females. While this might be due to processes of statistical discrimination by teachers, part of this result might be linked to individuals' behavior in classrooms that we are not able to capture with our variables. If the two stereotypes of boys being less diligent than girls [79] and overweight people being less disciplined than normal weight people are simultaneously present, this can lead to a double disadvantage for overweight boys. For overweight girls, in contrast, the gender stereotype of being more diligent can counteract the weight stereotype of being less disciplined, leading to lower bias. The latter might be especially relevant in a subject like German, where femininity is a favorable attribute.

Before concluding with policy recommendations, it is important to discuss some limitations. First, the grades are self-reported which means that the outcome variable can suffer from measurement error if the students intentionally or unintentionally reported wrong grades. Yet, we do not believe this issue should be a major concern, for two reasons. First, although especially students with unfavorable grades might have social desirability biases and report a better grade to increase their self-esteem, self-reported grades were found to be very similar to grades in the report cards in previous studies [80]. Second, if overweight/obese students on average obtain lower grades and are thus, on average, more prone to increase their self-reported grades, this could lead to an underestimation of their penalty. If this is the case, our findings would be based on conservative estimates.

The second limitation lies within the cross-sectional nature of the study, which does not allow us to disentangle whether overweight causes lower grades or if lower grades cause overweight by leading to stress eating [51]. Nevertheless, we believe that even if there are self-reciprocating effects, on theoretical ground it is much more plausible that the largest effect is of

BMI on grades and not *vice versa*. Furthermore, additional models in which we adjusted for life satisfaction to rule out extreme stress provide similar results.

The last limitation refers to the main assumption behind the grade-equation models, which posits that standardized test scores act as a good proxy for latent subject-specific competences. This assumption could be questioned if some unobserved factors associated with student BMI also affect the idiosyncratic performance in the standardized test. We are not aware of specific studies that try to tackle this aspect in-depth. However, results from previous studies on teachers' grading bias against immigrants or boys using a grade-equation model [76] report qualitatively similar findings than the few studies employing experimental designs [81].

Despite these limitations, we believe our study contributed to the literature by showing that overweight students have lower grades in German which can neither be attributed to lower subject-specific competence nor to psychological characteristics or key school-related attitudes and behavior. If institutions use grades as information on the competence of an applicant [82, 83], they should not contain anything else than academic competence in the given subject to be a valid source [10, 83]. But in reality, grades include other more subjective measures besides competences [10], and this can penalize students with the same competences but lower grades. To decrease the influence of discrimination and physical appearance on grades, we point towards a variety of educational policies at different levels of intervention. At the macro-level, our findings suggest the possibility of using competence test scores together with teachers' grades to provide an overall assessment of students' academic proficiency, especially at the end of educational cycles or in proximity of important educational transitions. At the meso- and micro-level, our findings point towards possible interventions on the grading policies and practices adopted by schools and teachers. First, in line with other scholars' advice [34], it appears to be important to grade students not in comparison to their classmates but to external criteria. These criteria should be stated in class and include what the students are expected to achieve and how the proof of achievement looks like [11]. Additionally, aside from competence-oriented practices, the censoring of the students' names on exams could decrease bias due to personal characteristics [81]. Another possible suggestion are training courses for raising the teachers' awareness of unconscious bias. This could enable them to detect and counteract discriminatory behavior. Although, awareness training in the context of grading is not researched by now, there are hints that raising awareness might lower bias in general: Indeed, it was found that people who believe to be probably biased against women actually judge them in a less biased way [84]. Not only for overweight students, but also for other discriminated minorities, external and fixed standards as well as increased awareness of possible discrimination might offer the opportunity to reduce inequalities of opportunities in the educational context.

Supporting information

S1 Table. Distribution of missing values across the variables (N = 3,814).
(DOCX)

S2 Table. Distribution of attributes in the analytical sample of students (N = 2,142).
(DOCX)

S3 Table. Three-level hierarchical ordinal logistic regression model: Log-odds ratio on teachers' grades.
(DOCX)

S4 Table. Three-level hierarchical ordinal logistic regression model with interaction: Log-odds ratio on teachers' grades.
(DOCX)

S5 Table. Average partial effect (APE) estimates reported in Fig 3.
(DOCX)

S6 Table. Average partial effect (APE) estimates reported in Fig 4.
(DOCX)

S7 Table. Interaction between BMI categories and students' test scores (N = 3,814).
(DOCX)

Author Contributions

Conceptualization: Mona Dian, Moris Triventi.

Data curation: Mona Dian.

Formal analysis: Mona Dian, Moris Triventi.

Methodology: Moris Triventi.

Supervision: Moris Triventi.

Visualization: Moris Triventi.

Writing – original draft: Mona Dian, Moris Triventi.

Writing – review & editing: Moris Triventi.

References

1. Andreyeva T, Puhl RM, Brownell KD. Changes in perceived weight discrimination among Americans: 1995–1996 through 2004–2006. *Obesity*. 2008; 16(5):1129–34. <https://doi.org/10.1038/oby.2008.35> PMID: 18356847
2. Latner JD, Stunkard AJ. Getting worse: The stigmatization of obese children. *Obes Res*. 2003; 11(3):452–6. <https://doi.org/10.1038/oby.2003.61> PMID: 12634444
3. Agerström J, Rooth D-O. The role of automatic obesity stereotypes in real hiring discrimination. *J Appl Psychol*. 2011; 96(4):790–805. <https://doi.org/10.1037/a0021594> PMID: 21280934
4. O'Brien KS, Latner JD, Ebneter D, Hunter JA. Obesity discrimination: The role of physical appearance, personal ideology, and anti-fat prejudice. *Int J Obesity*. 2012; 37(3):1–6. <https://doi.org/10.1038/ijo.2012.52> PMID: 22531085
5. Janssen I, Craig WM, Boyce WF, Pickett W. Associations between overweight and obesity with bullying behaviors in school-aged children. *Pediatrics*. 2004; 113(5):1187–94. <https://doi.org/10.1542/peds.113.5.1187> PMID: 15121928
6. Strauss RS, Pollack HA. Social marginalization of overweight children. *Arch Pediat Adol Med*. 2003; 157(8):746–52. <https://doi.org/10.1001/archpedi.157.8.746> PMID: 12912779
7. Ding W, Lehrer SF, Rosenquist JN, Audrain-McGovern J. The impact of poor health on academic performance: New evidence using genetic markers. *J Health Econ*. 2009; 28(3):578–97. <https://doi.org/10.1016/j.jhealeco.2008.11.006> PMID: 19217678
8. Saguy A. *What's wrong with fat?* Oxford: Oxford University Press; 2013.
9. Branigan AR. (How) Does obesity harm academic performance? Stratification at the intersection of race, sex, and body size in elementary and high school. *Sociol Educ*. 2017; 90(1):25–46. <https://doi.org/10.1177/0038040716680271> PMID: 29593365
10. Allen JD. Grades as valid measures of academic achievement of classroom learning. *Clearing House*. 2005; 78(5):218–23. <https://doi.org/10.3200/TCHS.78.5.218-223>
11. Close D. Fair grades. *Teach Philos*. 2009; 32(4):361–98. <https://doi.org/10.5840/teachphil200932439>
12. Salikutluk Z, Heyne S. Do gender roles and norms affect performance in maths? The impact of adolescents' and their peers' gender conceptions on maths grades. *Eur Sociol Rev*. 2017; 33(3):368–81. <https://doi.org/10.1093/esr/jcx049>

13. Bonefeld M, Dickhäuser O. (Biased) grading of students' performance: Students' names, performance level, and implicit attitudes. *Front Psychol*. 2018; 9(1):1–13. <https://doi.org/10.3389/fpsyg.2018.00481> PMID: 29867618
14. Dunkake I, Kiechle T, Klein M, Rosar U. Schöne Schüler, schöne Noten? [Good looks, good grades?]. *Z Soziol*. 2012; 41(2):142–61. <https://doi.org/10.1515/zfsoz-2012-0206>
15. Crandall CS, D'Anello S, Sakalli N, Lazarus E, Nejtardt GW, Feather NT. An attribution-value model of prejudice: Anti-fat attitudes in six nations. *Pers Soc Psychol Bull*. 2001; 27(1):30–7. <https://doi.org/10.1177/0146167201271003>
16. Crosnoe R, Muller C. Body mass index, academic achievement, and school context: examining the educational experiences of adolescents at risk of obesity. *J Health Soc Behav*. 2004; 45(4):393–407. <https://doi.org/10.1177/002214650404500403> PMID: 15869112
17. Crandall CS. Prejudice against fat people: Ideology and self-interest. *J Pers Soc Psychol*. 1994; 66(5):882–94. <https://doi.org/10.1037//0022-3514.66.5.882> PMID: 8014833
18. Puhl RM, Heuer CA. The stigma of obesity: A review and update. *Obesity*. 2009; 17(5):941–64. <https://doi.org/10.1038/oby.2008.636> PMID: 19165161
19. Roehling MV. Weight-based discrimination in employment: Psychological and legal aspects. *Pers Psychol*. 1999; 52(4):969–1016.
20. Larkin JC, Pines HA. No fat persons need apply: Experimental studies of the overweight stereotype and hiring preference. *Sociol Work Occup*. 1979; 6(3):312–27. <https://doi.org/10.1177/073088847900600303>
21. Duckworth AL, Seligman MEP. Self-discipline outdoes IQ in predicting academic performance of adolescents. *Am Psychol Soc*. 2005; 16(12):939–44. <https://doi.org/10.1111/j.1467-9280.2005.01641.x> PMID: 16313657
22. Paul RJ, Townsend JB. Shape up or ship out? Employment discrimination against the overweight. *Empl Responsib Rights J*. 1995; 8(2):133–45.
23. Wang SS, Brownell KD, Wadden TA. The influence of the stigma of obesity on overweight individuals. *Int J Obesity*. 2004; 28(10):1333–7. <https://doi.org/10.1038/sj.ijo.0802730> PMID: 15278101
24. Latner JD, O'Brien KS, Durso LE, Brinkman LA, MacDonald T. Weighing obesity stigma: The relative strength of different forms of bias. *Int J Obesity*. 2008; 32(7): 1145–52. <https://doi.org/10.1038/ijo.2008.53> PMID: 18414421
25. Puhl RM, Brownell KD. Bias, discrimination, and obesity. *Obes Res*. 2011; 9(12):788–805. <https://doi.org/10.1038/oby.2001.108> PMID: 11743063
26. Neumark-Sztainer D, Story M, Harris T. Beliefs and attitudes about obesity among teachers and school health care providers working with adolescents. *J Nutr Educ*. 1999; 31(1):3–9. [https://doi.org/10.1016/S0022-3182\(99\)70378-X](https://doi.org/10.1016/S0022-3182(99)70378-X)
27. Wilson SM, Smith AW, Wildman BG. Teachers' perceptions of youth with obesity in the classroom. *Adv Sch Ment Health Promot*. 2015; 8(4):231–43. <https://doi.org/10.1080/1754730X.2015.1074054>
28. Swami V, Frederick DA, Aavik T, Alcalay L, Allik J, Anderson D, et al. The attractive female body weight and female body dissatisfaction in 26 countries across 10 world regions: Results of the International Body Project I. *Pers Soc Psychol Bull*. 2010; 36(3):309–25. <https://doi.org/10.1177/0146167209359702> PMID: 20179313
29. Calogero RM, Boroughs M, Thompson JK. The impact of Western beauty ideals on the lives of women and men: A sociocultural perspective. In: Swami V, Furnham A, editors. *Body beautiful: Evolutionary and sociocultural perspectives*. London: Palgrave Macmillan; 2007. p. 259–98.
30. Sigfúsdóttir ID, Kristjánsson ÁL, Allegrante JP. Health behaviour and academic achievement in Icelandic school children. *Health Educ Res*. 2007; 22(1):70–80. <https://doi.org/10.1093/her/cyl044> PMID: 16766605
31. Florin TA, Shults J, Stettler N. Perception of overweight is associated with poor academic performance in US adolescents. *J Sch Health*. 2011; 81(11):663–70. <https://doi.org/10.1111/j.1746-1561.2011.00642.x> PMID: 21972986
32. Guskey TR. Making the grade: What benefits students? *Educ Leadersh*. 1994; 52(2):14–20.
33. Hjern A, Alfven G, Östberg V. School stressors, psychological complaints and psychosomatic pain. *Acta Paediatr*. 2008; 97(1):112–7. <https://doi.org/10.1111/j.1651-2227.2007.00585.x> PMID: 18076714
34. Organisation for Economic Co-operation and Development. *Grade expectations: How marks and education policies shape students' ambitions*. Paris: OECD Publishing; 2012.
35. Blossfeld H-P, Buchholz S, Skopek J, Triventi M, editors. *Models of secondary education and social inequality: An international comparison*. Cheltenham, Northampton: Edward Elgar Publishing Inc.; 2016.

36. Jackson M., editor. *Determined to succeed? Performance versus choice in educational attainment*. 1st ed. Stanford: Stanford University Press; 2013.
37. Johnes G. Standards and grade inflation. In: Johnes G, Johnes J, editors. *International handbook on the economics of education*. Cheltenham: Edward Elgar Publishing; 2004. p. 462–83.
38. National Educational Panel Study (NEPS): Starting cohort grade 5 (SC3) version 8-0-0 [dataset]. Bamberg: Leibniz Institute for Educational Trajectories, German Federal Ministry of Education and Research. 2019 Jan 11 [cited 2019 Apr 27].
39. Prevalence of obesity among adults, BMI \geq 30 (crude estimate) (%) [dataset on the Internet]. Geneva: World Health Organisation. 2017 Sep 22 [cited 2020 Aug 20]. [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-adults-bmi-30-\(crude-estimate\)-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-adults-bmi-30-(crude-estimate)-(-))
40. Carr D, Friedman MA. Is obesity stigmatizing? Body weight, perceived discrimination, and psychological well-being in the United States. *J Health Soc Behav*. 2005; 46(3):244–59. <https://doi.org/10.1177/002214650504600303> PMID: 16259147
41. Kenney EL, Gortmaker SL, Davison KK, Austin SB. The academic penalty for gaining weight: A longitudinal, change-in-change analysis of BMI and perceived academic ability in middle school students. *Int J Obesity*. 2015; 39(9):1408–13. <https://doi.org/10.1038/ijo.2015.88> PMID: 25982793
42. Shore SM, Sachs ML, Lidicker JR, Brett SN, Wright AR, Libonati JR. Decreased scholastic achievement in overweight middle school students. *Obesity*. 2008; 16(7):1535–8. <https://doi.org/10.1038/oby.2008.254> PMID: 18451772
43. Coppin G, Nolan-Poupart S, Jones-Gotman M, Small DM. Working memory and reward association learning impairments in obesity. *Neuropsychologia*. 2014; 65(1):146–55. <https://doi.org/10.1016/j.neuropsychologia.2014.10.004> PMID: 25447070
44. Gunstad J, Paul RH, Cohen RA, Tate DF, Gordon E. Obesity is associated with memory deficits in young and middle-aged adults. *Eat Weight Disord*. 2006; 11(1):e15–9. <https://doi.org/10.1007/BF03327747> PMID: 16801734
45. Reinert KRS, Po'e EK, Barkin SL. The relationship between executive function and obesity in children and adolescents: A systematic literature review. *J Obesity* [Internet]. 2013 Feb [cited 2019 Jan 22]; Article ID 820956:[10 p.]. <https://doi.org/10.1155/2013/820956> PMID: 23533726
46. Datar A, Sturm R. Childhood overweight and elementary school outcomes. *Int J Obesity*. 2006; 30(9):1449–60. <https://doi.org/10.1038/sj.ijo.0803311> PMID: 16534518
47. Judge S, Jahns L. Association of overweight with academic performance and social and behavioral problems: An update from the Early Childhood Longitudinal Study. *J Sch Health*. 2007; 77(10):672–8. <https://doi.org/10.1111/j.1746-1561.2007.00250.x> PMID: 18076412
48. Kaestner R, Grossman M. *Effects of weight on children's educational achievement*. Cambridge (MA): National Bureau of Economic Research; 2008 Jan. Working Paper No. 13764.
49. Li Y, Dai Q, Jackson J, Zhang J. Overweight is associated with decreased cognitive functioning among school-age children and adolescents. *Obesity*. 2008; 16(8):1809–15. <https://doi.org/10.1038/oby.2008.296> PMID: 18551126
50. Veldwijk J, Fries MC, Bemelmans WJ, Haveman-Nies A, Smit HA, Koppelman GH, et al. Overweight and school performance among primary school children: The PIAMA birth cohort study. *Obesity*. 2012; 20(3):590–6. <https://doi.org/10.1038/oby.2011.327> PMID: 22030985
51. Sabia J. The effect of body weight on adolescent academic performance. *South Econ J*. 2007; 73(4):871–900.
52. Duckworth AL, Quinn PD, Tsukayama E. What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *J Educ Psychol*. 2012; 104(2):439–51. <https://doi.org/10.1037/a0026280> PMID: 24072936
53. MacCann C, Roberts RD. Just as smart but not as successful: Obese students obtain lower school grades but equivalent test scores to nonobese students. *Int J Obesity*. 2013; 37(1):40–6. <https://doi.org/10.1038/ijo.2012.47> PMID: 22531092
54. Pine KJ. Children's perceptions of body shape: A thinness bias in pre-adolescent girls and associations with femininity. *Clin Child Psychol Psychiatry*. 2001; 6(4):519–36. <https://doi.org/10.1177/1359104501006004006>
55. Blossfeld H-P, Roßbach H-G, von Maurice J, editors. *Education as a lifelong process—the German National Educational Panel Study (NEPS)*. [Special Issue] *Z Erzieh Wiss* No. 14. Wiesbaden: Springer VS; 2011.
56. Leibniz Institute for Educational Trajectories. Starting Cohort 3: Grade 5 (SC3) study overview waves 1 to 8 [Internet]. Bamberg: Leibniz Institute for Educational Trajectories; 2018 [cited 2019 Jan 22].
57. Rubin DB. *Multiple imputation for nonresponse in surveys*. Hoboken: John Wiley & Sons; 1987.

58. Little RJA, Rubin DB. *Statistical analysis with missing data*. 2nd ed. Hoboken: John Wiley & Sons; 2002.
59. Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behav Res*. 1998; 33(4):545–71. https://doi.org/10.1207/s15327906mbr3304_5 PMID: 26753828
60. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011; 30(4):377–99. <https://doi.org/10.1002/sim.4067> PMID: 21225900
61. Horton NJ, Kleinman KP. Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *Am Stat*. 2007; 61(1):79–90. <https://doi.org/10.1198/000313007X172556> PMID: 17401454
62. Percentile data files with LMS values: BMI-for-age charts, 2 to 20 years, LMS parameters and selected smoothed BMI (kilograms/meters squared) percentiles by sex and age [dataset on the Internet]. Hyattsville: National Center for Health Statistics. 2000 May 30 [cited 2019 Jan 22]. https://www.cdc.gov/growthcharts/percentile_data_files.htm
63. Centers for Disease Control and Prevention [Internet]. Atlanta: Division of Nutrition, Physical Activity, and Obesity, National Center for Chronic Disease Prevention and Health Promotion; [last reviewed 2020 Apr 10]. Overweight and obesity: Defining childhood obesity, BMI for children and teens; [last reviewed 2018 Jul 3; cited 2019 Jan 22]; [about 1 screen]. <https://www.cdc.gov/obesity/childhood/defining.html>
64. National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. 2013 YRBS data user's guide. Atlanta: CDC Stacks; 2014 [cited 2019 Jan 22]. <https://stacks.cdc.gov/view/cdc/50245>
65. Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? *Am J Epidemiol*. 1996; 143(3):228–39. <https://doi.org/10.1093/oxfordjournals.aje.a008733> PMID: 8561156
66. Pohl S, Carstensen CH. NEPS technical report: Scaling the data of the competence tests. Bamberg: University of Bamberg; 2012 Oct. Working Paper No. 14. <https://www.neps-data.de/Portals/0/Working%20Papers/WPXIV.pdf>
67. Leibniz Institute for Educational Trajectories. Information on competence testing. NEPS starting cohort 3–grade 5. Paths through lower secondary school—educational pathways of students in grade 5 and higher. Wave 3: grade 7 [Internet]. Bamberg: Leibniz Institute for Educational Trajectories; 2018 [cited 2020 Aug 19]. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/3-1-0/NEPS_SC3_Compences_W3_en.pdf
68. Shrewsbury V, Wardle J. Socioeconomic status and adiposity in childhood: A systematic review of cross-sectional studies 1990–2005. *Obesity*. 2008; 16(2):275–84. <https://doi.org/10.1038/oby.2007.35> PMID: 18239633
69. Sirin SR. Socioeconomic status and academic achievement: A meta-analytic review of research. *Rev Educ Res*. 2005; 75(3):417–53. <https://doi.org/10.3102/00346543075003417>
70. Bogg T, Roberts BW. Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychol Bull*. 2004; 130(6):887–919. <https://doi.org/10.1037/0033-2909.130.6.887> PMID: 15535742
71. Wagerman SA, Funder DC. Acquaintance reports of personality and academic achievement: A case for conscientiousness. *J Res Pers*. 2007; 41(1):221–9. <https://doi.org/10.1016/j.jrp.2006.03.001>
72. Terraciano A, Sutin AR, McCrae RR, Deiana B, Ferrucci L, Schlessinger D, et al. Facets of personality linked to underweight and overweight. *Psychosom Med*. 2009; 71(6): 682–9. <https://doi.org/10.1097/PSY.0b013e3181a2925b> PMID: 19414622
73. Chamorro-Premuzic T, Furnham A. Personality predicts academic performance: Evidence from two longitudinal university samples. *J Res Pers*. 2003; 37(4):319–38. [https://doi.org/10.1016/S0092-6566\(02\)00578-0](https://doi.org/10.1016/S0092-6566(02)00578-0)
74. Lange D, Plachta-Danielzik S, Landsberg B, Müller MJ. Soziale Ungleichheit, Migrationshintergrund, Lebenswelten und Übergewicht bei Kindern und Jugendlichen [Social inequality, migration background, life environments, and overweight of children and adolescents]. *Bundesgesundheitsblatt—Gesundheitsforschung—Gesundheitsschutz*. 2010; 53(7):707–15. <https://doi.org/10.1007/s00103-010-1081-4> PMID: 20631975
75. Dardanoni V, Modica S, Pennisi A. Grading across schools. *B. E. J Econ Anal Policy*. 2009; 9(1):16. <https://doi.org/10.2202/1935-1682.2065>
76. Kiss D. Are immigrants and girls graded worse? Results of a matching approach. *Educ Econ*. 2013; 21(5):447–63. <https://doi.org/10.1080/09645292.2011.585019>

77. Brambor T, Clark WR, Golder M. Understanding interaction models: Improving empirical analyses. *Polit Anal.* 2006; 14(1):63–82. Lavy V. Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *J Public Econ.* 2008;92(10–11):2083–105.
78. Cassady JC. Self-reported GPA and SAT: A methodological note. *Pract Assess Res Evaluation.* 2001; 7(12):1–6. <https://doi.org/10.7275/5hym-y754>
79. Heyder A, Kessels U. Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Soc Psychol Educ.* 2015; 18(3):467–85. <https://doi.org/10.1007/s11218-015-9303-0>
80. Nitko AJ. *Educational assessment of students.* 3rd ed. Upper Saddle River: Merrill; 2001.
81. Sprietsma M. Discrimination in grading: Experimental evidence from primary school teachers. *Empirical economics.* 2013; 45(1), 523–538.
82. Thoms P, McMasters R, Roberts MR, Dombkowski DA. Resume characteristics as predictors of an invitation to interview. *J Bus Psychol.* 1999; 13(3):339–56. <https://doi.org/10.1023/A:1022974232557>
83. Guskey TR. Helping standards make the grade. *Educational Leadership.* 2001; 59(1):20–7.
84. Uhlmann EL, Cohen GL. Constructed criteria: Redefining merit to justify discrimination. *Psychol Sci.* 2005; 16(6):474–80. <https://doi.org/10.1111/j.0956-7976.2005.01559.x> PMID: 15943674