I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15811

Daily Use of Social Media Is Associated with More Body Dissatisfaction of Teenage Girls in a Large Cross-Cultural Survey

Clotilde Napp
Thomas Breda

DECEMBER 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15811

Daily Use of Social Media Is Associated with More Body Dissatisfaction of Teenage Girls in a Large Cross-Cultural Survey

Clotilde Napp
Université Paris Dauphine, CNRS and IZA

Thomas Breda
Paris School of Economics, CNRS and IZA

DECEMBER 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Daily Use of Social Media Is Associated with More Body Dissatisfaction of Teenage Girls in a Large Cross-Cultural Survey

Most teenagers spend several hours per day on social media. We provide a large-scale investigation of the relationship between social media daily usage and body dissatisfaction among a sample of more than 50,000 15 y.o. students. This relation is positive and large for girls—higher use of social networks is associated with higher dissatisfaction about their body—and negative for boys. The positive relation for girls is observed in all eight countries included in the study, covering very different cultural contexts (e.g., Georgia, Ireland, Spain, Mexico, Panama or Hong Kong). It is observed for all girls, no matter their body mass index (BMI), their academic performance, and their socioeconomic background. Instrumenting social networks consumption by students' or students' peers' internet access at home while controlling finely for other students' or students' peers' household characteristics suggests that the relationship between social media consumption and girls' body dissatisfaction could be causal.

JEL Classification: I12, L82
Keywords: social media, body dissatisfaction

Corresponding author:
Thomas Breda
Paris School of Economics
48 Bd Jourdan
75014 Paris
France
E-mail: thomas.breda@ens.fr

Introduction

Social media are extensively used by adolescents (1,2) and have received a lot of research attention as a possible risk factor for body dissatisfaction (3). Sociocultural theories of body dissatisfaction, such as the Tripartite influence model (4), propose that social media consumption shapes a person's body dissatisfaction through the processes of social comparison and internalization of appearance ideals. Popular social media platforms used by adolescents such as Instagram, Facebook or Snapchat (1) are highly visual and appearance focused platforms that extend opportunities to engage in behaviors that induce body dissatisfaction (3). Indeed the platforms above do not only feature models and celebrities but also the users themselves who usually post personal images and seek out comparisons with their peers and reactions from them (5,6). According to objectification theory (7), girls could be particularly subject to the detrimental impact of social media, since they are constantly looked at and objectified, which makes them come to see themselves as an object for others to evaluate based on their appearance.

However, it is also argued that social media can have positive effects on body image. Viewing body-positive content on social media platforms can help individuals become more accepting and appreciative of their bodies and social media have also gained traction in promoting diverse body types, shapes, and sizes (8–10). Additionally, social media enable to connect with others to build a body-accepting community and share resources, messages, and images that are healing or affirming (11).

The effect of social media on body dissatisfaction is therefore debated. This debate has important health implications as body dissatisfaction, defined as “a person’s negative thoughts and feelings about his/her body” (12), has been shown repeatedly to be a major risk factor for physical and psychological health issues, like eating-related pathologies, low self-esteem and depression symptomatology (13–16). The sharp increase in the past twenty years in depression or eating-related pathologies among teenage girls or women that accompanied the birth of social media—for example, according to a recent meta-analysis (17), prevalence rates of eating-related pathologies for women have more than doubled from 3.5% in the 2000–2006 period to 7.8% in the 2013–2018 period—is also a real public-health and societal concern that deserves urgent attention.

We provide a large-scale multi-country study of the relationship between social media consumption on a regular basis (daily consumption) and body dissatisfaction. Doing so, we contribute to a large body of empirical studies whose conclusions are so far inconsistent (18), with some studies finding a positive association between social media use and body dissatisfaction (*e.g.*, (19,20)), and others finding the opposite (8) or no relationship (21,22). We however improve over former observational studies—see (11,23) for meta-analyses—in three directions. First, based on large representative samples of 15 y.o. students, we provide evidence that is fully consistent across eight different countries capturing very distinct cultural contexts. This contrasts with existing studies that usually focus on samples that are (i) not representative of the general population they cover, (ii) small in size (typically around 150 individuals, and always less than 1,000 individuals in the 20 studies

covered in (23)), and (iii) focused on a single country (the only exceptions being (24,25) that both compare Korea and the U.S.).

Second, using a rich multi-country survey, we are able to control finely for teenager characteristics, including measures of their physical appearance (size, weight and BMI), their academic performance, their social background, or their family wealth, ensuring that we do not simply capture the fact that social media consumption varies depending on physical appearance (*e.g.*, if underweight individuals do more sport and go less on social media) or other obvious factors.

Third, we try to establish a causal link by exploiting variation across students in the availability of internet at home as an instrument generating plausibly exogenous variation in students' social media consumption. This strategy is made possible by the fact that we focus on countries where internet is not yet available to all students. When looking at a specific teenager, we exploit both the availability of internet to her, and to the other students in the same school. Indeed, a student's incentive to go on social media increases when her peers also do so, hence when peers have an internet connection, they go more on social media, which in turn increases the student propensity to use social media. We exploit these reasonings to obtain plausibly exogenous variation across students in social media consumption and study its likely causal impact on body dissatisfaction.

Results

Data

Analyses are based on data from PISA2018, an every-3-year international assessment of the knowledge and skills of 15-year-old students in mathematics, reading, and science in more than 70 countries. PISA2018 focuses on academic performance but it also includes measures of students' socioeconomic background and an optional well-being questionnaire that was distributed by 9 countries about students' satisfaction with different aspects of their lives, like their health, their life at school and, of particular interest for our analysis, their body and look. Moreover, to better understand students' use of the internet, an optional Information and Communication Technology (ICT) familiarity questionnaire was distributed by 53 countries and included questions about how teenagers use digital devices. Details about PISA survey (e.g. sampling design, sampling weights, representativeness) and the data we analyse are provided in the Methods section and in the Supplementary Materials (SM).

The sample of teenagers responding to both the well-being and the ICT familiarity questionnaire is reduced to 51,539 observations in 8 countries with complete information on both body satisfaction and social media consumption: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818).

The measure of Body Dissatisfaction and First Observations

The PISA survey includes a well-being questionnaire with six different items dealing with body, look and appearance: 1. 'I am not concerned about my weight', 2. 'I like my body', 3. 'I am satisfied with the way I look', 4. 'I consider myself to be attractive', 5. 'I like my look just the way it is', and 6. 'I like the way my clothes fit me'. The first three items reflect body dissatisfaction along three margins: one's weight, the dislike of one's body in general, and the dissatisfaction about one's look or appearance. These are arguably the main aspects of body dissatisfaction. On the contrary, item 4 is about attractivity, and one can consider oneself attractive but at the cost of a lot of time and effort and feeling attractive is not inconsistent with body concerns and dissatisfaction. Reciprocally, one can consider oneself not attractive without really caring about it or paying attention to it, and without it being associated with body dissatisfaction. Item 5 is about one's look and is redundant with item 3. Finally, item 6 is about one's clothes, and is not directly related to one's body. As for item 4, I can like my clothes but be dissatisfied with my body, or have no problem with my body but pay no attention to my clothes. For these reasons, our main measure of Body Dissatisfaction (BD) only keeps the first three items. BD is constructed as the equally weighted average of the three (reverse-coded) items, that we standardize so that it has a mean of zero and a Standard Deviation (SD) of 1 in each country—see details in the SM.

Consistent with previous studies based on smaller samples (e.g. (26–32)), we find that body dissatisfaction is larger for girls than for boys in all countries in our sample (Table S1). The gender gap in body dissatisfaction is equal to 0.33SD on the whole sample. We observe for example that 45.9% of girls but "only" 38.7% of boys express concerns about their weight, and that 32.4% of girls but "only" 22.5% of boys declare not liking their body. The gender gap in our index of body dissatisfaction and the differences above are robust to controlling for Body Mass Index (BMI), socioeconomic background, academic performance, as well as other individual characteristics like students' life satisfaction that could have an impact on students' dissatisfaction about their bodies (Table S1)—see the SM for the description of these variables.

Analysing the role of these individual characteristics, we find, as expected, that higher BMI (as informed by the students) is associated with higher body dissatisfaction for both boys and girls (Figure S1). This is more the case for girls than for boys and the gender gap in body dissatisfaction is higher when BMI is higher (0.29SD for above average BMI and 0.22SD for below average BMI). Moreover, for girls, the increase is almost linear whereas for boys, body dissatisfaction is almost constant as a function of BMI for students whose BMI is below median and begins to increase linearly only for BMI above median.

High academic performance (as proxied by the average of science and math performance) is associated with more body dissatisfaction for girls, whereas the level of body dissatisfaction does not vary much with performance for boys (Figure S2). As a result, the gender gap in body dissatisfaction is more than twice larger among highest performers than among lowest ones.

Concerning socioeconomic status, girls from wealthier and more educated families suffer from higher body dissatisfaction relative to boys (Figure S3). This pattern is also observed at the macro level across the eight countries in our sample: the gender gap in body

dissatisfaction is highly heterogenous across countries, ranging from 0.06SD in Panama to 0.56SD in Ireland (Table S1, column 1), and it is actually higher in more economically developed, wealthy or gender equal countries, with a correlation coefficient r=0.87 with the Human Development Index (HDI), r=0.85 with the Gross Domestic Product (GDP) and r=0.66 with the Gender Gap Index which captures gender equality in rights (see SM).

In sum, body dissatisfaction for boys and for girls is related to a wide range of individual and cultural characteristics like academic performance, socio economic background, BMI or the country level of development. In line with objectification theory, girls' body dissatisfaction is higher than that of boys and the relation with individual and cultural characteristics also differs by gender. Note that PISA2018 provides data on students' life satisfaction and self-esteem—see the SM—and we find on our whole sample that body dissatisfaction is more "central" for girls than for boys in the sense that it is a stronger predictor of life satisfaction, of negative feelings, of self-efficacy or of finding a meaning in life for girls than for boys (Table S2).

Social media consumption and body dissatisfaction

In the PISA2018 questionnaire on ICT familiarity, 15-year-old students are asked in a series of distinct items to report how frequently ("never or hardly ever", "once or twice a month", "once or twice a week", “almost every day”, "every day") they use the Internet for different activities, such as playing games, chatting online, browsing for fun, obtaining practical information, uploading or downloading content, or using email. There is a specific item for students’ participation in social media: “how often do you use digital devices outside of school for participating in social networks (e.g., Facebook, MySpace)”. Figure S4 shows that when all countries are considered together, more intensive usage of social media is associated with higher body dissatisfaction for girls and lower body dissatisfaction for boys. These relationships are mostly driven by the most intensive users of social media: body dissatisfaction is 0.13SD larger among girls using social media every day than among other girls, and 0.16SD smaller among boys using social media every day than among other boys (Table 1, col. 1). This suggests that a mild usage of social media may not have the same effects as a daily usage, and we therefore consider in what follows as our main measure of social media consumption the fact of being a daily user. This approach splits the sample in two halves of roughly equal size as 51% of the students in our sample use social media every day. It is also motivated by the fact that daily users of social media can potentially use it several hours per day, implying that the intensity of their usage (which we do not measure perfectly) is likely to contrast sharply from that of milder users. For example, in 2019 in the U.S., female teenagers (ages 13-18) using social media (not necessarily daily) spend *on average* 2:17 hours per day on these media (against 1:31 for male teenagers, see (33)).

There is an excess representation of girls relative to boys among daily users of social media in all the countries in our sample (Table S3). The same question about participation in social networks was asked in PISA2012 and both the percentage of daily users and the gender gap have increased in the interval between the two surveys.

Daily usage of social media (DSM) is also positively related to academic performance and socioeconomic background and slightly negatively related to BMI, especially for girls (Figure S5). At the country level, DSM is higher in more developed countries, with a coefficient of correlation r=0.88 between the percentage of daily users of social media and the HDI (N=8 countries).

Turning to the main results, DSM is associated with higher body dissatisfaction for girls and lower body dissatisfaction for boys in all countries in our sample (Figure 1). In all these countries, the gender gap in body dissatisfaction is more than twice larger on average among daily users of social media than among lower users. It is striking to observe such a consistent pattern across such distinct cultures and parts of the world as Asia (Hong Kong), southern Europe (Spain), Eastern Europe (Bulgaria and Serbia), former USSR (Georgia), Central America (Mexico and Panama) and an Anglo-Saxon culture (Ireland).

Based on linear regression models, we investigate if the associations between social media consumption and body dissatisfaction for girls and boys are robust to controlling for students' characteristics. Table 1, col. 1 first retrieves from a regression with no additional controls the gender gap in body dissatisfaction among low users of social media (0.11SD), as well as the partial effect of daily social media consumption on girls' body dissatisfaction (0.13SD) and boys' body dissatisfaction (-0.16SD). The second column shows that these partial effects are robust to controlling for country fixed effects and quintiles of the BMI. The third column adds a rich set of control variables for (i) students' characteristics (grade, grade repetition, and math and science performance), (ii) for their mothers' and fathers' detailed education and occupational status, and (iii) for characteristics of their households (*e.g.*, availability of books, cars, TVs, personal bedroom and cultural goods at home, see SM for details). In the fourth column, we also interact these controls with student's sex to allow them to have a different effect for girls and boys. In the fifth one, school fixed effects are also included (there are about 2,600 schools in our sample, with an average of 20 students surveyed per school). The relationship between DSM and BD decreases for boys when more controls are added (it is divided by about two) but remains very stable for girls.

The link between social media consumption and body dissatisfaction may also be confounded by other usages of the internet, such as playing online games, chatting online, browsing for fun, obtaining practical information, uploading or downloading content, or using email. Indeed, for both girls and boys, DSM is positively correlated with all these possible alternative usages (Table S4). We therefore include them (also interacted with students' gender) as additional control variables in Table 1 (col. 6). For girls, the link between DSM and BD remains almost unchanged. For boys, however, the relationship becomes weaker and statistically insignificant, suggesting that social media consumption was partially capturing other internet usages.

To get a better understanding of how teenagers in our sample use the internet, we have performed a principal component analysis of the various possible usages for girls and boys separately (Table S5). For both girls and boys, the first component loads positively and almost equally on all possible usages, showing that students who use the internet tend to use it for everything. The second component, however, loads positively on playing games,

obtaining practical information, using emails, and uploading or downloading content, but it loads somewhat negatively on browsing the Internet for fun and very negatively on social media consumption, chatting online. Hence, we can identify with these two main components two profiles of students (no matter their gender): those who use the internet for everything, and those who only play video games, upload or download content, and use it for emails and practical matters. The last column of Table 1 shows that these two typical usages of the internet have opposite effects on body satisfaction for girls: using the internet for everything is associated with an increase in body dissatisfaction, while using it only to play games, uploading or downloading content, or practical things is associated with a non-statistically significant decrease. This highlights that not all internet activities are associated with a lower female body satisfaction.

The associations between DSM and BD are finally observed no matter girls' and boys' BMI (Figure 2), academic performance (Figure S6) and socioeconomic background (Figure S7). Focusing on girls, we observe that the link between DSM and BD is stronger for girls with high BMI, high academic performance, and low socioeconomic background. These heterogeneous effects for girls according to academic performance or socioeconomic background are statistically significant, robust to testing jointly for the three sources of heterogeneity and robust to the inclusion of control variables (Table S6). This is however not the case for BMI.

Causal analysis

To what extent can we interpret the associations above as a causal effect of daily social media consumption on body dissatisfaction? To answer this question, we try to obtain exogenous variation in students' social media consumption stemming from their ability to access the internet. This strategy is made possible by the fact that our sample includes several countries in which a significant fraction of households still does not have access to the internet. Students from such households are likely to be prevented from using social media simply because they are unable to connect to the internet. Our approach will therefore consist in comparing the body satisfaction of students who can and cannot use social media rather than that of students who choose to use and not to use social media. We argue that the former comparison is more likely to reflect a causal impact of social media on body satisfaction.

We implement this idea by using the availability of an internet connection at home as an instrumental variable for DSM. This technique (34) amounts to extracting variations in DSM that solely arise from the availability of an internet connection at home and to using only these latter variations to explain variations in body dissatisfaction. It identifies the causal impact of DSM on BD under the assumption that having an internet connection at home affects body satisfaction only through its effect on social media consumption. Table 2, col. 1, shows that when DSM is instrumented by internet access at home, it has a much larger effect on girls' BD, amounting to 0.40SD.

The assumption above is however likely to be violated because households with an internet connection differ substantially from those without: they are for example richer and from

higher social background. These differences may in their own lead to differences in students' body satisfaction as, for example, high and low SES students receive a different education. We therefore take advantage of the detailed information available in PISA2018 to control finely for home possessions, and households and students' characteristics. The idea is to compare students from very similar households, except that one has internet access while the other does not. When we do so (Table 2, col. 2), the effect of DSM on BD remains significant and large (0.42SD).

We then exploit a second idea to improve the credibility of our instrumental variable approach: as networks, social media become more attractive for someone when her relatives or peers use them as well. However, the fact that a student's peers use social media also depend on the fact that the student does it, and cannot be directly used as an instrument due to this reflection problem (35). We therefore directly use as an additional or alternative instrument the availability of internet at home for the peers. The idea being that when a student's peers have access to the internet, these peers are more likely to go on social media, which in turn increases the student's own likelihood to use these media. In practice, our instrument is for each student the share of other students in her school who have access to internet at home. This alternative instrument allows us to identify the causal effect of DSM on BD providing that the availability of internet at home for a student affects the body satisfaction of another student in the same school only through the fact that it will increase the probability that the latter student uses social media.

This additional instrument is used jointly with the first in the third column of Table 2. The fourth column also includes controls for the characteristics of the peers and their households. This is to ensure that a student peers' access to the internet does not capture other observable peer characteristics that could potentially be correlated with the student's body dissatisfaction. In both columns, the link between DSM and BD for girls is large, around 0.7 SD. Column 5 then splits the peers according to their sex and uses three instruments instead of two: the student's internet access at home, the share of other girls in the school having internet at home, and the share of boys in the school having internet at home. Finally, the last column drops from the instruments the student's own internet connection as it may be more endogenous and could have driven our results so far. The large effect of DSM and BD is maintained across all these specifications.

To validate our instrumental variable approach, we provide a number of additional tests in the SM (p. 9-11 and Table S7). First, students' access to the internet at home indeed (strongly) increases daily social media usage (from 15 to 27 percentage points, depending on control variables included, see Table S7). The share of other students' (or other girls') in the school with internet access at home also increases a student's probability to use social media daily (columns 3 to 6 of Table S7). In all models estimated in Table 2, Fisher statistics additionally show that instruments are not weak. Finally, in all specifications that include at least two instruments except col. 2, we cannot reject the validity of our over-identification restrictions. We conclude from all these tests that our instrumental variables are valid from a statistical point of view. We also discuss additional robustness checks in the SM. In total, the evidence allows us to conclude that a daily consumption of social media is likely to have a large negative impact on girls' body satisfaction.

Interestingly, the opposite relationship found for boys becomes statistically insignificant when we use the same instrumental variable approach (Table S8), suggesting that it could be non-causal. It could be the case for example that mostly boys with a high body satisfaction use social media, so that there is a positive association between the two even though social media does not directly impact boys' body satisfaction.

Robustness to other measures of body dissatisfaction and of use of social networks

For the sake of robustness and reliability, we have considered as alternative measures of body dissatisfaction the three (reverse-coded) individual items involved in the computation of our measure BD, taken separately—i.e., the body weight concern, the general body dislike and the body look dissatisfaction. We have checked that our main results in both Tables 1 and 2 are overall robust to the consideration of these three alternative measures. We have also checked that they hold when considering social media consumption linearly rather than an indicator of daily usage. We have finally checked that results in Table 1 remain when we do not use sampling weights.

Discussion

In our data, the relationship between daily social media consumption and body dissatisfaction is systematically positive for teenage girls, no matter their country, BMI, performance at school or socioeconomic background, and it is negative or inexistent for boys. These results established on a large scale and for representative samples of students, complement former observational studies (18–25). They highlight the importance of analysing girls and boys separately, potentially explaining why studies that considered them together obtained contrasting results. The magnitude of the relationship between social media consumption and body dissatisfaction also varies according to teenagers' BMI, academic performance or socioeconomic background, which could further explain why studies based on non-representative samples may have reached different conclusions.

The present study however suffers from two main empirical limitations. First, we are not able to measure social media consumption very precisely, and cannot isolate among daily users those who spend several hours a day on these media, and may be the most at risk. This limitation is perhaps less acute in our context as we focus on countries where the penetration of social media remains somewhat moderate in 2018 (as compared to the U.S. for example) but further work is nevertheless necessary to better understand the effects of using social media several hours a day. Second, body dissatisfaction is self-reported and only observed through three direct questions. This contrasts with studies that could use detailed questionnaires or a series of depictions with which subjects can rate their current and ideal body shape. Such studies are arguably better able to distinguish different aspects of body dissatisfaction.

Going beyond these limitations, we try to take advantage of the large amount of information available in PISA2018 to make progress toward interpreting the positive association between girls' social media usage and their body dissatisfaction as causal. To

do so, we exploit plausibly exogenous variation in social media consumption arising from differences in students and their peers' internet access at home. Results established using such variation are very large in magnitude, and larger than the simple association between girls' social media usage and their body dissatisfaction, suggesting that this association may actually underestimate the causal impact of social media. This would be the case for example if the girls with a high body satisfaction tend to go more on social media. It could also be that the associations presented in Table 1 are subject to attenuation bias due to measurement errors, a problem that the use of instrument variables can solve (see details in SM).

Of course, in the absence of experimental variation in social media consumption, we cannot make causal claims with certainty. Our approach would be invalid for example if students having an internet access at home differed in some characteristics that we cannot control for and that directly affect their body satisfaction. Using the internet access of peers as an alternative instrument allows us to avoid this problem. However, if students who have internet access at home physically interact with their peers in ways that can directly affect their body satisfaction, our second approach will be invalid as well. Such threats to our strategy to obtain causality remain reasonably unlikely, implying that the present study makes substantial progress toward assessing the causal impact of social media consumption on body dissatisfaction using observational data. Doing so, it complements existing experimental studies (30,36–40) that establish a causal effect of social media on body image with certainty, but are only able to evaluate the effect of specific, short-term and limited variations in social media consumption. In contrast, our study captures the average effect on body satisfaction of social media consumption on a regular basis. This effect can arise from various channels induced by the dynamic, interactive, and personalized nature of social media, and further research is necessary to better understand these channels.

Regarding exact mechanisms, we can therefore only make a couple of general comments. First, the much larger estimates obtained when social media usage is instrumented by peers' access to the internet is consistent with the idea that peers' usage of social media increases the detrimental effect of these media on girls' body satisfaction (see the technical discussion of this point in the SM). This is in line with predictions from social comparison theory, which posits that the need for self-evaluation pushes people to seek out comparisons with similar rather than dissimilar others (41). As a consequence, social media could be detrimental to girls' body satisfaction not because it exposes them to external models, but because it induces upward comparisons with peers that do not occur in standard offline interactions. Hence, the detrimental effect of social media may arise not because it connects teenage girls to a broader community, but because it enhances new types of interactions with their already established network.

Second, our main results are fully consistent with the tripartite influence model and objectification theory. They highlight the specificity of body image concerns for girls who appear particularly vulnerable to the implicit body appearance comparisons occurring on social media. Doing so, they contribute to the understanding of the possible detrimental effects of the recent surge in social media use among teenagers (42–46), in particular females.

Regarding the quantitative importance of social media for teenagers—who can spend several hours a day using them—understanding their effects on health and well-being should arguably represent a major social and public health objective. We contribute to this objective by providing large-scale evidence on their likely negative effects on female teenagers' body image. It seems necessary to better inform parents and the general public about such possible detrimental effects and to start thinking about public health or social policies that could help mitigating these effects. Information campaigns would probably be good start. Further research is also required to better understand the consequences of what is a huge societal change.

Methods

Details on Data

This study uses exclusively data from the 2018 Programme for International Student Assessment (PISA 2018). It focuses on 8 countries with complete information on both body satisfaction and social media consumption: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818). Those are the countries where both the ICT and well-being questionnaire were administered (see (47), chapter 17 for details). The questionnaires are filled on a computer in all countries.

The desired base PISA target population in each country consisted of 15-year-old students attending educational institutions in grades 7 and higher. The sampling design used for the PISA assessment to reach these students is a two-stage stratified sample design. The first-stage sampling units consisted of individual schools having 15-year-old students, or the possibility of having such students at the time of assessment. Schools were sampled systematically from a comprehensive national list of all PISA-eligible schools, known as the school sampling frame, with probabilities that were proportional to a measure of size. The measure of size was a function of the estimated number of PISA-eligible 15-year-old students enrolled in the school.

The second-stage sampling units were students within sampled schools. Once schools were selected to be in the sample, a complete list of each sampled school's 15-year-old students was prepared. The target cluster size in each school is 42 students who were drawn with equal probability among pre-selected students within the school. For additional details on population coverage, exclusions, and student participation rates, see (47), chapter 4.

Survey weights are finally provided to make the students observed in PISA 2018 representative of the target population (see (47), chapter 9). We provide a general description of PISA2018 and its purpose in the SM.

Statistical Methods

We use the PISA 2018 ICT questionnaire to construct a measure of daily social media consumption (DSM) equal to one for students going on social media every day. We use the wellbeing questionnaire to measure body dissatisfaction (BD) from three questions. Details on these data constructions can be found in the Results section and in the SM. We systematically use sampling weights when computing summary statistics (i.e. when computing means of BD or DSM by subgroups).

To analyse the link between DSM and BD, we first use linear regression models estimated using Weighted Least Squares. The weights are the sampling weights provided by PISA. The main model used to obtain the results in Table 1 is as follows:

$$BD_{is} = \alpha(DSM_i * Boy_i) + \beta(DSM_i * Girl_i) + \gamma X_i + \delta(X_i * Girl_i) + \eta_s + (\eta_s * Girl_i) + \epsilon_i$$

where the subscript *i* denotes students, the subscript *s* denotes their school, X_i are a set of characteristics of student *i* and her/his household, and η_s is a series of school fixed effects that are also interacted with students' gender. This specification captures the partial effect of DSM on BD separately for boys (coefficient α) and girls (coefficient β), conditional on various students' characteristics. Controls are added progressively in the subsequent columns of Table 1, and the school fixed effects (also interacted with gender) are only included in column 5. These controls are described in the Table notes and details on their construction are provided in the SM (Appendix A). The last column replaces DSM_i by the first two principal components of a PCA of the various usages students make of the Internet (see details in the Results section).

The Instrumental Variable (IV) results presented in Table 2 use (unweighted) two-stage least squares. The analysis is restricted to girls. The first stage equation is:

$$DSM_{is} = \alpha IC_i + \beta\overline{IC_{-is}} + \gamma X_i + \delta\overline{X_{-is}} + \epsilon_i$$

IC_i is the first instrument, i.e. an indicator variable equal to one if girl *i* has access to the Internet at home. $\overline{IC_{-is}}$ is the second instrument, i.e. the average of IC_i for all students in school *s* except girl *i*. In the first two columns of Table 2, $\overline{IC_{-is}}$ is not included. In columns 4 to 6 of Table 2, $\overline{IC_{-is}}$ is split in two variables: the share of boys in the school with internet access at home, and the share of girls (except *i*) with internet access at home. $\overline{X_{-is}}$ are the average of the students' characteristics in school *s*, excluding student *i*. In columns 5 and 6 of Table 2, the characteristics of girls and boys are considered separately. From the first stage equation, we get $\widehat{DSM_i}$, the daily social media consumption predicted by the instrument(s) and control variables. We then plug it in the second stage equation:

$$BD_{is} = \alpha\widehat{DSM_i} + \beta X_i + \gamma\overline{X_{-is}} + \epsilon_i$$

α captures the (causal) effect of DSM on BD according to our instrumental variable approach.

In all regression Tables (Tables 1 and 2 as well as SM Tables), we report robust standard errors. Additional technical details and relevant statistical tests are provided in the SM.

Data availability statement: All data used in the analysis are publicly available and deposited at [TO BE COMPLETED]

Code availability statement: All codes used in the analysis are publicly available and deposited at [TO BE COMPLETED]

References and Notes

1. Anderson M, Jiang J. Teens, social media & technology 2018. Pew Res Cent. 2018;31(2018):1673-89.
2. Rodgers RF, Lombardo C, Cerolini S, Franko DL, Omori M, Fuller-Tyszkiewicz M, et al. The impact of the COVID-19 pandemic on eating disorder risk and symptoms. Int J Eat Disord. 2020;53(7):1166-70.
3. Rodgers RF, Melioli T. The relationship between body image concerns, eating disorders and internet use, part I: A review of empirical support. Adolesc Res Rev. 2016;1(2):95-119.
4. Thompson JK, Coovert MD, Stormer SM. Body image, social comparison, and eating disturbance: A covariance structure modeling investigation. Int J Eat Disord. 1999;26(1):43-51.
5. Fardouly J, Vartanian LR. Social media and body image concerns: Current research and future directions. Curr Opin Psychol. 2016;9:1-5.
6. Heinberg LJ, Thompson JK. Body Image and Televised Images of Thinness and Attractiveness: A Controlled Laboratory Investigation. J Soc Clin Psychol. 1 déc 1995;14(4):325-38.
7. Fredrickson BL, Roberts TA. Objectification theory: Toward understanding women's lived experiences and mental health risks. Psychol Women Q. 1997;21(2):173-206.
8. Cohen R, Fardouly J, Newton-John T, Slater A. # BoPo on Instagram: An experimental investigation of the effects of viewing body positive content on young women's mood and body image. New Media Soc. 2019;21(7):1546-64.
9. Slater A, Varsani N, Diedrichs PC. # fitspo or# loveyourself? The impact of fitspiration and self-compassion Instagram images on women's body image, self-compassion, and mood. Body Image. 2017;22:87-96.
10. Verduyn P, Ybarra O, Résibois M, Jonides J, Kross E. Do social network sites enhance or undermine subjective well-being? A critical review. 2017;
11. Saiphoo AN, Vahedi Z. A meta-analytic review of the relationship between social media use and body image disturbance. Comput Hum Behav. 2019;101:259-75.
12. Grogan S. Body image: Understanding body dissatisfaction in men, women and children. Routledge; 2016.
13. Jacobi C, Hayward C, de Zwaan M, Kraemer HC, Agras WS. Coming to Terms With Risk Factors for Eating Disorders: Application of Risk Terminology and Suggestions for a General Taxonomy. Psychol Bull. 2004;130(1):19-65.
14. Stice E, Shaw HE. Role of body dissatisfaction in the onset and maintenance of eating pathology: A synthesis of research findings. J Psychosom Res. 2002;53(5):985-93.

15. Paxton SJ, Neumark-Sztainer D, Hannan PJ, Eisenberg ME. Body Dissatisfaction Prospectively Predicts Depressive Mood and Low Self-Esteem in Adolescent Girls and Boys. J Clin Child Adolesc Psychol. 1 déc 2006;35(4):539-49.
16. Cruz-Sáez S, Pascual A, Wlodarczyk A, Echeburúa E. The effect of body dissatisfaction on disordered eating: The mediating role of self-esteem and negative affect in male and female adolescents. J Health Psychol. 1 juill 2020;25(8):1098-108.
17. Galmiche M, Déchelotte P, Lambert G, Tavolacci MP. Prevalence of eating disorders over the 2000–2018 period: a systematic literature review. Am J Clin Nutr. 1 mai 2019;109(5):1402-13.
18. Mahon C, Hevey D. Processing Body Image on Social Media: Gender Differences in Adolescent Boys' and Girls' Agency and Active Coping. Front Psychol. 2021;12:1703.
19. Fardouly J, Pinkus RT, Vartanian LR. The impact of appearance comparisons made through social media, traditional media, and in person in women's everyday lives. Body Image. 2017;20:31-9.
20. Scully M, Swords L, Nixon E. Social comparisons on social media: online appearance-related activity and body dissatisfaction in adolescent girls. Ir J Psychol Med. 2020;1-12.
21. Ferguson CJ, Muñoz ME, Garza A, Galindo M. Concurrent and prospective analyses of peer, television and social media influences on body dissatisfaction, eating disorder symptoms and life satisfaction in adolescent girls. J Youth Adolesc. 2014;43(1):1-14.
22. Cohen R, Newton-John T, Slater A. The relationship between Facebook and Instagram appearance-focused activities and body image concerns in young women. Body Image. 2017;23:183-7.
23. Holland G, Tiggemann M. A systematic review of the impact of the use of social networking sites on body image and disordered eating outcomes. Body Image. 2016;17:100-10.
24. Lee HE, Taniguchi E, Modica A, Park H. Effects of witnessing fat talk on body satisfaction and psychological well-being: A cross-cultural comparison of Korea and the United States. Soc Behav Personal Int J. 1 sept 2013;41(8):1279-95.
25. Lee HR, Lee HE, Choi J, Kim JH, Han HL. Social Media Use, Body Image, and Psychological Well-Being: A Cross-Cultural Comparison of Korea and the United States. J Health Commun. 2 déc 2014;19(12):1343-58.
26. Fardouly J, Vartanian LR. Negative comparisons about one's appearance mediate the relationship between Facebook usage and body image concerns. Body Image. 1 janv 2015;12:82-8.
27. Tiggemann M, Miller J. The Internet and adolescent girls' weight satisfaction and drive for thinness. Sex Roles. 2010;63(1-2):79-90.
28. Tiggemann M, Slater A. NetGirls: The Internet, Facebook, and body image concern in adolescent girls. Int J Eat Disord. 2013;46(6):630-3.
29. Tiggemann M, Slater A. NetTweens: The Internet and Body Image Concerns in Preteenage Girls. J Early Adolesc. 1 juin 2014;34(5):606-20.
30. Mabe AG, Forney KJ, Keel PK. Do you "like" my photo? Facebook use maintains eating disorder risk. Int J Eat Disord. 2014;47(5):516-23.

31. Vandenbosch L, Eggermont S. Understanding Sexual Objectification: A Comprehensive Approach Toward Media Exposure and Girls' Internalization of Beauty Ideals, Self-Objectification, and Body Surveillance. J Commun. 1 oct 2012;62(5):869-87.
32. de Vries DA, Peter J, Nikken P, de Graaf H. The Effect of Social Network Site Use on Appearance Investment and Desire for Cosmetic Surgery Among Adolescent Boys and Girls. Sex Roles. 1 nov 2014;71(9):283-95.
33. Rideout V, Robb, Michael. The common sense census: media use by tweens and teens [Internet]. Common Sense Media; 2019 [cité 11 janv 2022]. Disponible sur: https://apo.org.au/node/58360
34. Wooldridge JM. Econometric Analysis of Cross Section and Panel Data, second edition. MIT Press; 2010. 1095 p.
35. Manski CF. Identification of Endogenous Social Effects: The Reflection Problem. Rev Econ Stud. 1 juill 1993;60(3):531-42.
36. Tiggemann M, Zaccardo M. "Exercise to be fit, not skinny": The effect of fitspiration imagery on women's body image. Body Image. 1 sept 2015;15:61-7.
37. Haferkamp N, Krämer NC. Social Comparison 2.0: Examining the Effects of Online Profiles on Social-Networking Sites. Cyberpsychology Behav Soc Netw. 1 mai 2011;14(5):309-14.
38. Fardouly J, Diedrichs PC, Vartanian LR, Halliwell E. Social comparisons on social media: The impact of Facebook on young women's body image concerns and mood. Body Image. 1 mars 2015;13:38-45.
39. Mills JS, Musto S, Williams L, Tiggemann M. "Selfie" harm: Effects on mood and body image in young women. Body Image. 1 déc 2018;27:86-92.
40. Cohen R, Blaszczynski A. Comparative effects of Facebook and conventional media on body image dissatisfaction. J Eat Disord. 2 juill 2015;3(1):23.
41. Festinger L. A theory of social comparison processes. Hum Relat. 1954;7(2):117-40.
42. Twenge JM, Haidt J, Joiner TE, Campbell WK. Underestimating digital media harm. Nat Hum Behav. avr 2020;4(4):346-8.
43. Allcott H, Braghieri L, Eichmeyer S, Gentzkow M. The Welfare Effects of Social Media. Am Econ Rev. mars 2020;110(3):629-76.
44. Haidt J, Allen N. Scrutinizing the effects of digital technology on mental health. Nature. févr 2020;578(7794):226-7.
45. Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. Nat Hum Behav. févr 2019;3(2):173-82.
46. Twenge JM, Martin GN. Gender differences in associations between digital media use and psychological well-being: Evidence from three large datasets. J Adolesc. 1 févr 2020;79:91-102.
47. PISA 2018 Technical Report. Paris: OECD.

Acknowledgments: To be completed

Author contributions: Both authors contributed equally to each aspect of this work.

Funding: The authors have no funding to declare.

Competing interests: Authors declare that they have no competing interests.

Supplementary Materials

Data

Methods

Figs. S1 to S7

Tables S1 to S8

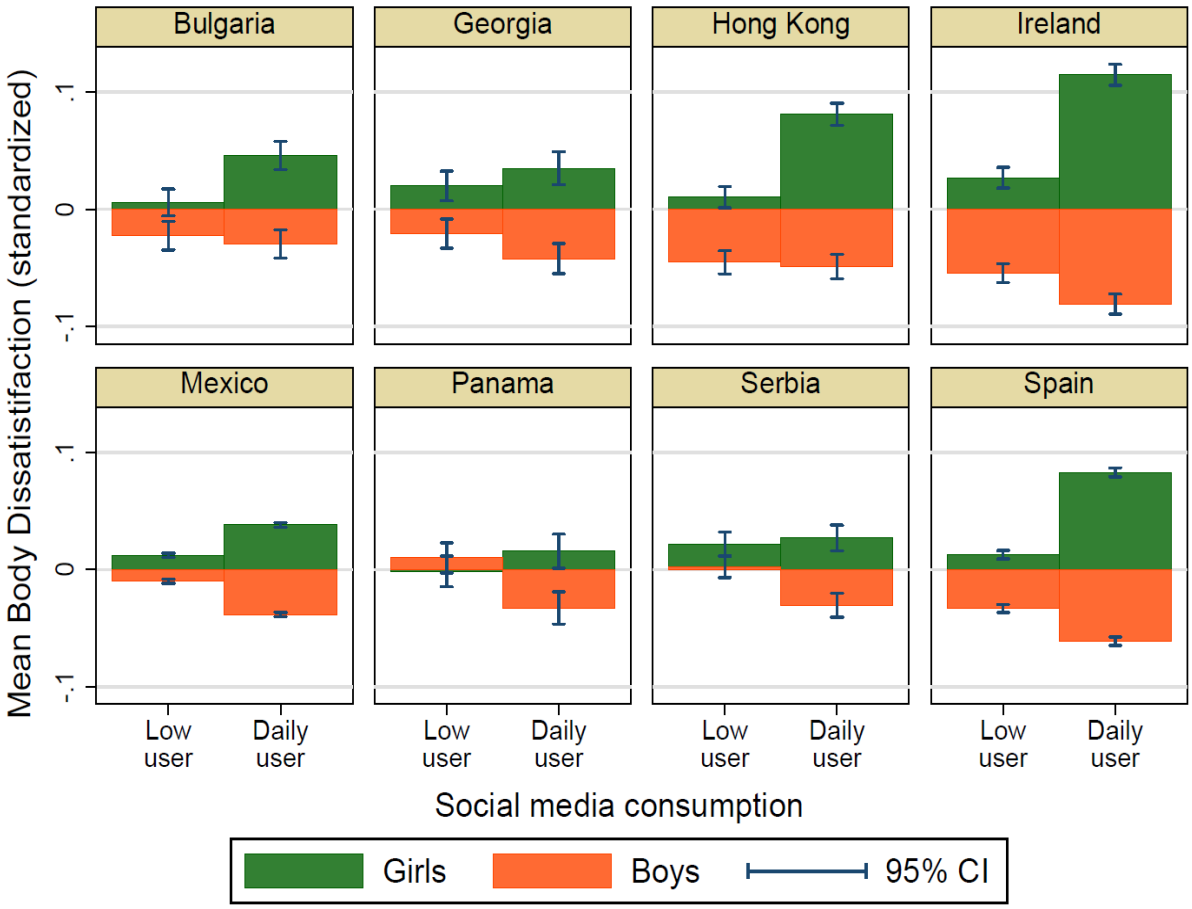


Figure 1: Girls' and boys' body dissatisfaction for low and daily users of social media. The Figure provides in each country the average level of body dissatisfaction (standardized index) for girls and boys (weighted using sample weights) depending on their usage of social media. The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818).

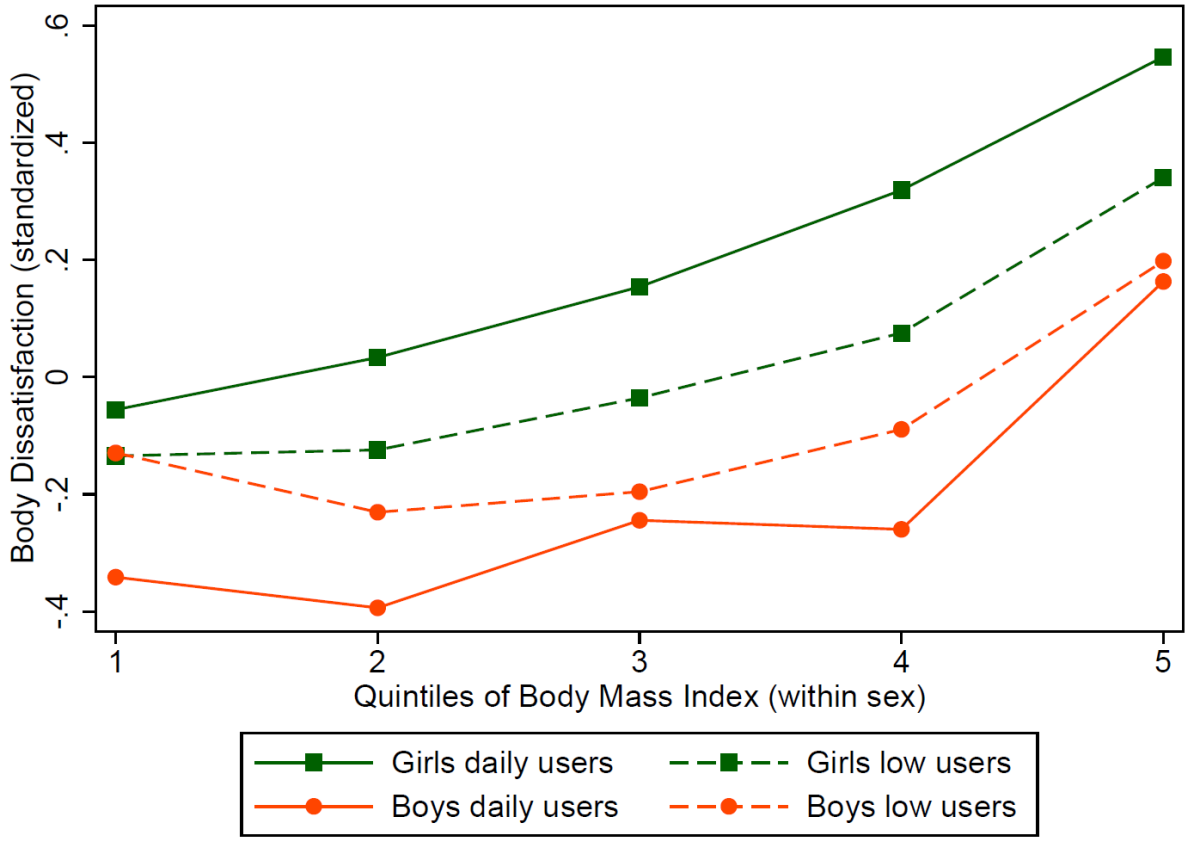


Figure 2: Link between social media consumption and body dissatisfaction as a function of the body mass index. The Figure provides the average level of body dissatisfaction (standardized index) for girls and boys depending on their usage of social media and their BMI (5 quintiles defined separately for each sex). The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818).

Table 1: Link between social media consumption (daily usage) and body dissatisfaction for girls and boys. Estimates from linear regression models with various sets of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent variable: Index of body dissatisfaction						
Girl	0.111***	0.115***	0.100***				
	(0.0316)	(0.0324)	-0.0377				
Girl*Daily user	0.131***	0.164***	0.162***	0.121***	0.118***	0.104**	
	(0.0304)	(0.0335)	-0.038	(0.0393)	(0.0399)	(0.0491)	
Boy*Daily user	-0.157***	-0.132***	-0.126***	-0.0879**	-0.0924***	-0.0450	
	(0.0298)	(0.0306)	(0.0346)	(0.0346)	(0.0348)	(0.0483)	
Girl*First comp. of internet usage							0.0267**
							(0.0124)
Girl*Second comp. of internet usage							-0.0283
							(0.0192)
Boy*First comp. of internet usage							-0.0276***
							(0.00974)
Boy*Second comp. of internet usage							-0.00919
							(0.0169)
Control variables:							
Quintiles of BMI and country fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Students and households' controls	No	No	Yes	Yes	Yes	Yes	Yes
The above interacted with gender	No	No	No	Yes	Yes	Yes	Yes
School Fixed effects (also interacted with gender)	No	No	No	No	Yes	No	No
Other possible internet usages (also interacted with gender)	No	No	No	No	No	Yes	No
Observations	51,539	43,688	33,673	33,673	33,673	30,818	30,818
R-squared	0.020	0.054	0.073	0.086	0.114	0.089	0.086

Notes: Results from linear regression models. Controls for students and households' characteristics include controls for the student grade, performance in math and science, repeater status, possessions and resources available at home as well as controls for her or his parents' education and occupation and her or his household characteristics. The complete list is given in Appendix A. In columns 4 to 7, the estimate for girls is unavailable as it is absorbed by the other controls. The first and second components of the pca of internet usages are described in Table S5. Observations are weighted using students' sampling weights. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Girls only. Link between social media consumption (daily usage) and body dissatisfaction. Estimates from instrument variable models.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Index of Body Dissatisfaction					
Daily usage of social media (instrumented)	**0.404*** (0.123)**	**0.421** (0.198)**	**0.704*** (0.179)**	**0.751*** (0.224)**	**0.619** (0.285)**	**1.005** (0.506)**
Instruments:						
Own internet access at home	Yes	Yes	Yes	Yes	Yes	No
Other students in school average internet access	No	No	Yes	No	No	No
Other girls in school average internet access	No	No	No	Yes	Yes	Yes
Boys in school average internet access	No	No	No	Yes	Yes	Yes
Control variables:						
Country fixed effects and BMI	Yes	Yes	Yes	Yes	Yes	Yes
Students and households' characteristics	No	Yes	Yes	Yes	Yes	Yes
Peers' characteristics	No	No	No	No	Yes	Yes
Observations	21,739	17,078	16,400	10,330	10,330	10,330

Notes: Results from two-stage least squares regression models. The sample only includes girls. When internet access of peers is used as instrument, only schools with at least 9 peers are included in the sample (to avoid noisy estimates). This amounts to keeping schools with 10 students or more in column (3) and with at least 10 boys and 10 girls in subsequent columns. This selection rule has little impact on the final results. Averages for peers are computed taking the mean of the variables considered across all relevant peers (*e.g.*, the share of other girls in the school having access to the internet). Controls are similar to those in Table 1 (see Appendix A for details) with the exception that for peers, we replace the set of dummies mother education, father education and number of books at home by one single continuous variable (in order to avoid increasing dimensionality too much). This again has little impact on the final results. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Supplementary Material for

Daily use of social media is associated with more body dissatisfaction among teenage girls in a large cross-cultural survey

Authors: Clotilde Napp[1,2] Thomas Breda[3,4]

Affiliations:

[1] Université Paris Dauphine, Paris Sciences et Lettres Research University, 75016 Paris, France

[2] CNRS, UMR7088, 75016 Paris, France

[3] Paris School of Economics, 75014 Paris, France

[4] CNRS, UMR8545, 75014 Paris, France

*Corresponding author: Thomas Breda
Email: thomas.breda@ens.fr

This PDF file includes:

Supplementary text about materials and methods
Figures S1 to S7
Tables S1 to S8

Other supplementary materials for this manuscript include the following:
Databases and codes allowing the replication of the results will be deposited on a third-party server.

Appendix A: Data

General Description of PISA 2018

The Programme for International Student Assessment (PISA) is an every-three-year international survey of 15-year-old students aimed at determining their knowledge and skills in different domains. Students' abilities are assessed in the three curricular domains: mathematics, reading, and science. Students also answer a background questionnaire, seeking information about the students themselves, their homes, and their school and learning experiences.

The PISA target population is made up of all students in any educational institution between the ages of 15 years and 3 months and 16 years and 2 months at the time of the assessment. This specific age has been chosen because it is close to the end of compulsory education in most countries. Efforts have been made to ensure the absence of cultural or national biases in the test items and in the evaluation of performance.

We analyse data from the PISA 2018 survey.

PISA surveys systematically assess students' performance and knowledge in three core subjects: mathematics, reading and science. However, one of the three core subjects is chosen to be covered in greater depth in each survey. In 2018, reading literacy is the major subject area, as it was in 2009.

For the first time, PISA2018 distributed a well-being questionnaire about students' satisfaction with different aspects of their lives, like their health, their life at school, the way they look. Questions about students' body dissatisfaction are included in the well-being questionnaire. PISA2018 includes more than 70 countries, but only 8 countries distributed the well-being questionnaire.

PISA2018 also distributed a questionnaire on ICT (Information and Communications Technology) familiarity. The *ICT familiarity questionnaire* of *PISA 2018* assesses students' interest in *ICT,* use of *ICT*, perceived competence, and autonomy in using *ICT,* and the use of social media.

Our sample consists of more than 51,539 students responding to both the well-being and the ICT familiarity questionnaire, with OECD countries like Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235) and non-OECD countries like Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818). The national samples are representative of all 15-year-old students in the country.

Variables of interest in PISA 2018

Body dissatisfaction and other body-related items. Several questions in the well-being questionnaire deal with students' body, look and appearance. Questions WB153 are about a student's body image. Students are asked, thinking about themselves, how much they agree with each of the following statements: 'I like my look just the way it is' (WB153q01ha), 'I consider myself to be attractive' (WB153q02ha), 'I am not concerned about my weight' (WB153q03ha), 'I like my body' (WB153q04ha), 'I like the way my clothes fit me' (WB153q05ha).
Answers are given on a four-point Likert scale with response categories ranging from "Strongly disagree", "Disagree", "Agree", to "Strongly agree". An additional response option "I don't have an opinion" was treated as missing, unless otherwise specified.
An additional question WB155q02ha asks students how satisfied they are with the way they look. Answers range from 1 (not at all satisfied) to 4 (totally satisfied).
For all these body and appearance items and index, we reverse the scale so that higher scores correspond to higher dissatisfaction with one's body.

We take as our main measure of body dissatisfaction (BD) the equally weighted average of the (reverse-coded) items 'I am not concerned about my weight' (WB153q03ha), 'I like my body' (WB153q04) and 'how satisfied are you about the way you look' (WB155q02). We think that the measure BD captures "body dissatisfaction", considering the concerns about one's weight, the dislike of one's body in general, and the dissatisfaction about one's look or appearance. We standardize the measure BD at the country level (see below).

For the sake of robustness and reliability, we shall consider as alternative measures of body dissatisfaction the three (reverse-coded) individual items involved in the computation of our measure BD, i.e.,

- the *body weight concern or dissatisfaction*, measured by the reverse-coded item 'I am not concerned about my weight' (WB153q03ha),
- the *general body dislike or dissatisfaction*, measured by the reverse-coded item 'I like my body' (WB153q04ha).
- the *body look dissatisfaction* measured by the reverse coded item 'how satisfied are you about the way you look' (WB155q02ha)

In order to measure, at the country level, the percentage of boys and girls who are concerned about their body weight, or don't like their body and the gender gap, we introduce the binary variable equal to one when the student answers either "Strongly disagree" or "Disagree" to question WB153q03ha or WB153q04ha.

Students' use of social networks, Internet access and other ICT-related items
In the questionnaire on ICT familiarity, students are asked (in questions IC001) whether given devices are available for them to use at home. The list of devices includes a desktop computer (IC001Q01), a portable laptop (IC001Q02), Internet connection (IC001Q04), a

cellular phone without Internet access (IC001Q06), a cellular phone with Internet access (IC001Q07), a printer (IC001Q09).
Possible answers are "Yes and I use it", "Yes but I don't use it" and "No".
We consider that the student has an Internet access at home if she/he answers positively ("Yes and I use it" or "Yes but I don't use it") to question IC001Q04. We do not consider cellular phones with Internet access (IC001Q07) because access to such a device is arguably more endogenous to students' desires to go online: a student who is very much willing to go online (e.g. to use social media) might obtain a connected cellular phone from her parents. This is less the case for the general Internet connection available in the household.

PISA2018 questionnaire on ICT familiarity includes questions (IC008) about students' use of Internet for given activities, and more precisely how often students use digital devices for playing one-player games (IC008Q01), playing collaborative games online (IC008Q02), using email (IC008Q03), chatting online (e.g., MSN) (IC008Q04), participating in social networks (e.g., Facebook, MySpace) (IC008Q05), playing online games via social networks (*e.g.*, the Sims Social) (IC008Q07), browsing the Internet for fun (IC008Q08), reading news on the Internet (IC008Q09), obtaining practical information from the Internet (IC008Q10), downloading music, films, games or software from the Internet (IC008Q11), uploading created contents for sharing (IC008Q12), downloading new apps on a mobile device (IC008Q13).
Answers are given on a five-point scale with response categories "Never or hardly ever", "Once or twice a month", "Once or twice a week", "Almost every day" and "Every Day".
Of particular interest to us is the question on the use of social media. We denote by DSM the daily use of social media, *i.e.*, students answering "Every day" to question IC008Q05. This splits the sample in two halves of roughly equal size as 51% of the students in our sample use social media every day.

Students' performance and socioeconomic background

- We use individual-level PISA scores in math and science to measure *students' performance* (we exclude the score in reading for the measure because it is not available in Spain). These scores are on a 0-1000 scale. They have been scaled during the first PISA survey in 2000 to have a mean of 500 and a standard deviation of 100. Performances in math and science are highly correlated and we use the average of science and math performance to capture students' academic level. Note that PISA adopts the Item Response Theory model and does not provide for each student actual science or math scores but plausible values for their ability in these fields. These plausible values (10 for PISA 2018) are random numbers drawn from the distribution of scores that could be reasonably assigned to each individual, given his or her answers - that is, the marginal posterior distribution. When students' performance is used as a control variable, we simply control for the first plausible values for math and science performances (with no incidence on our results). When we instead examine heterogeneity in outcomes depending on students' performance (Figures S2 and S4), we compute estimates obtained for each plausible value for

the average of math and science performance, and take the average of these estimates as our final estimate.[1]

- In PISA, a student's socio-economic background is estimated by the PISA index of *economic, social and cultural status* (ESCS), which is based on information about parental education, highest parental occupation, and home possessions including books in the home. We use it in heterogeneity analyses.

Students' life satisfaction, negative feelings and sense of self-efficacy. The standard students' questionnaire in PISA2018 includes questions about students' life satisfaction, meaning in life and self-efficacy.

- In question ST016q01na, students are asked to rate their *life satisfaction* ('Overall, how satisfied are you with your life as a whole these days?') on a scale from 0 (not at all satisfied) to 10 (completely satisfied).
- In questions ST188 students are asked to report the extent to which they agree ("strongly disagree", "disagree, "agree", "strongly agree") with the following statements about themselves: "I usually manage one way or another"; "I feel proud that I have accomplished things"; "I feel that I can handle many things at a time"; "My belief in myself gets me through hard times"; and "When I'm in a difficult situation, I can usually find my way out of it". These statements were combined to create the index of *self-efficacy* (RESILIENCE). Positive values in this index mean that the student reported higher self-efficacy than did the average student across OECD countries.
- In questions ST186, students are asked to report about their *positive and negative feelings* and in particular how frequently ("never", "rarely", "sometimes", "always") they feel miserable.
- In questions ST185, students are asked to report the extent to which they agree ("strongly agree", "agree", "disagree", "strongly disagree") with the following statements: "My life has clear meaning or purpose"; "I have discovered a satisfactory meaning in life"; and "I have a clear sense of what gives meaning to my life". These statements were combined to form the index of *meaning in life* (EUDMO). Positive values in the index indicate greater meaning in life than the average student across OECD countries.

Body Mass Index

PISA includes in the well-being questionnaire questions about the student's height and weight (WB151q01ha and WB152q01ha), which permits to determine individual *Body Mass Index* (BMI), as informed/declared by the students. We recall that Body mass index (BMI) is a person's weight in kilograms divided by the square of height in meters. We divide the samples of girls and boys in 5 quintiles of BMI and control in some specifications by students' quintile of BMI, the quintile being determined separately for boys and girls.

Other control variables

The students' and household characteristics used as control variables in Tables 1 and 2 are the following:

[1] This is done using the Stata command Repest.

- Individual controls:
    - Student grade relative to modal grade in country (6 dummies)
    - Student is a repeater (one dummy equal to one if the student is currently repeating a grade or did repeat a grade in the past)
    - Math performance (first Plausible Value)
    - Science performance (first Plausible Value)
- Household controls:
    - Highest parental education in years of schooling (one variable)
    - Mother education. 7 dummies, one for each of the following possible educational attainments according to the ISCED classification
        - None,
        - ISCED 1,
        - ISCED 2,
        - ISCED 3B, C
        - ISCED 3A, ISCED 4,
        - ISCED 5B,
        - ISCED 5A, ISCED 6
    - Father education (7 dummies, constructed similarly to mother education)
    - Mother occupational rank (one variable between 0 and 100)
    - Father occupational rank (one variable between 0 and 100)
    - Index of cultural possessions at home built by PISA (e.g. books of classical literature, poetry books, pieces of art)
    - Home possessions and school resources:
        - At least one car at home
        - At least a TV at home
        - A desk to work of her/his
        - A room of her/his own
        - A quiet place to study
        - A dictionary at home
        - Books at home to help with school work
        - Technical reference books at home
        - Total number of books at home (6 dummies)

These controls are also averaged among students' peers (other students of the same sex in the school and students of the other sex in the school) and used as controls for peers' characteristics in Table 2. To reduce the total number of controls (with no impact on the final results), we however consider one single control for peers' mother education, father education, average grade, and number of books at home. This means that we do not create a set of dummies but simply average the initial variable across the relevant peers, ignoring its discrete nature.

Data on country-level measures of development, wealth and individualism

Human Development Index (HDI)
The Human Development Index (HDI) is a composite statistic of life expectancy, education, and per capita income indicators. A country scores a higher HDI when life expectancy, education level and per capita income is higher. Values have been taken for year 2018.
Source: http://hdr.undp.org/en/content/2019-human-development-index-ranking

Gross Domestic Product (GDP)
GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.
Source:https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD?end=2019&most_recent_year_desc=true&start=2018

Gender Gap Index (GGI)
The Gender Gap Index, from the World Economic Forum, synthesizes the position of women in any given country by taking into account economic opportunities, economic participation, educational attainment, political achievements, and health and well-being. Larger values point to a better position of women in society.
Source: GGI2018, World economic forum
http://www3.weforum.org/docs/WEF_GGGR_2018.pdf

Appendix B: Methods

Weights and representativeness

PISA provides weights to make surveyed students representative of the 15-year-old students of the surveyed countries. We use these weights in all country-specific and whole sample descriptive statistics, so that all the results we provide are representative of the whole student population in the eight countries we focus on. We also use the weights in the regression analyses in Table 1. As there are pros and cons to using weights in regression analysis, we have checked that the results are robust to not doing it. For the two-stage least square regressions in Table 2 and Table S7, we do not use weights as results from IVs with weights have no clear interpretation.

Standardized variables

To get results that can be interpreted both on our whole sample and country-by-country, we normalize the index of body dissatisfaction so that its weighted mean is zero in each country of the sample, while its weighted standard deviation is one in each country of the sample. Gender gaps in body dissatisfaction are then directly expressed as a fraction of its standard deviation and, as such, directly comparable across countries. This standardization implies that our final index of body dissatisfaction has the same mean and standard

deviation in each country, so that our main results are not driven by cross-country differences in the variable. For the same reason, we standardize similarly ("at the country level") our measures of academic performance and of socioeconomic background. This allows us to control for (or to examine the effect of) the relative academic performance or relative academic performance of a student in her or his country. Other variables (such as the BMI) are not standardized, unless otherwise specified.

Regressions

Empirical analyses presented in the paper rely primarily on student-level regressions that control for some observable students' characteristics as well as on country-level correlations.

Principal Components Analysis of internet usages

To perform the principal component analysis of the various usages of Internet, we first group the two items corresponding to "playing games" (IC008Q07NA and IC008Q02TA), considering that a student plays daily if she/he answers "daily" to one of these items. We perform a similar grouping for the two items corresponding to "downloading content" (IC008Q13NA and IC008Q11TA). We consider other items independently in the PCA (see the list in Table S5). The same list of variables for internet usages is used as controls in Table 1, column (6).

Instrumental variables analysis

First stage results

Results for the first stage are provided in Table S7. Students' access to the internet at home has a significant effect on daily social media usage in all specifications. Note that the survey also provides information on access to the internet on a mobile device. The rationale for only instrumenting by internet access at home is that access on a mobile device might be more endogenous. For example, students who strongly desire to use social media may strongly bargain with their parents to get a connected mobile phone, while we may think that internet access at home is something over which students have little direct control. We observe that students' access to the internet at homes increases their likelihood to consume social media by about 27 percentage points (Table S7, col. 1). When other instruments and controls are added, the effect remains very large (larger than 15 percentage points in all columns of Table S7).

Regarding the other instruments, only boys' average internet access in the school has no significant direct impact on a girl likelihood to use social media, even though the point estimates go in the expected direction. By contrast, the fact that other girls have access to the internet have a large and significant effect.

Statistical tests

We start by examining the Kleibegen-Paap Wald rk F statistic to detect possible weak identification. The F stats are always above 20, except in the last column, suggesting there is no problem of weak identification, except perhaps in the last specification (where it is still above 10, the rule of thumb level to consider that instruments are not weak).

We then propose three weak-instrument-robust inference tests based on the Anderson-Rubin Wald F statistic, the Anderson-Rubin Wald chi-squared statistic and the Stock-Wright LM S statistic. These tests are of primary importance when instruments are weak, but even when they are not, they are valid. They allow us to confirm that social media consumption has a significantly negative impact on girls' body dissatisfaction, even when allowing the instruments to be weak. Indeed, for all three tests, we strongly reject the null that the effect of daily social media consumption on girls' body dissatisfaction is equal to zero in all specifications except the two last ones. For those two last specifications, we only reject the null at the 10% level.

To wrap up, there is no clear sign of weak instruments, except perhaps when a very large number of control variables are included (col. 5 and 6 of Table S7). In these latter cases, the instruments' strength is still acceptable, but on the edge of accepted magnitudes for the Fisher statistics. However, the effect of social media consumption on body dissatisfaction remains significant at the 10% level in these specifications even when we use significance tests that allow for weak instruments. As our results are consistent across specifications with and without this large set of control variables used in columns 5 and 6, we are confident that the main conclusions are not driven by slightly weak instruments in these last specifications.

Regarding overidentification, the p-values associated to Hansen J statistics for specifications with at least two instruments estimated in Table 2 are the following: p=0.025 in col. 3, p=0.41 in col. 4, p=0.49 in col. 5 and p=0.49 in col. 6. This shows, that except in col. 3, we cannot reject the validity of our over-identification restrictions. To put it simply, this means that instruments are coherent with each other, or that, assuming that one is truly exogenous, the others are exogenous as well.

Additional robustness checks

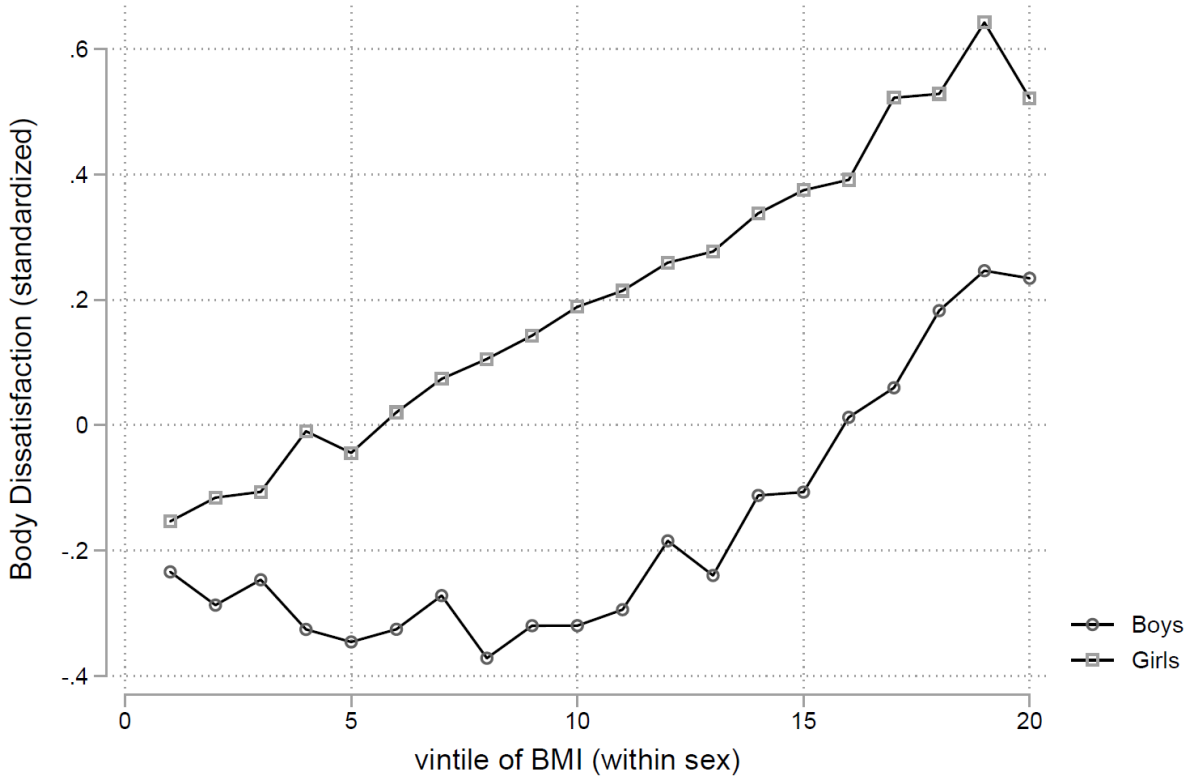
We finally tested several additional specifications. We first tested the stability of the estimates for each set of instruments to the addition or removal of certain control variables (e.g. adding controls for peers' characteristics in col. 3 of Table 2, removing these controls in col. 6). We then tested the effects of other combinations of our four possible instruments, such as using other students', other girls' or boys' in the school average internet access at home as a single instrument. Point estimates remain similar and statistically significant across most such specifications. It is only when we use boys' in the school average internet access at home as a single instrument that the estimate gets statistically not significant (while still large), likely because this instrument is a weak predictor of girls' daily usage of social media. Together, our additional checks show the robustness of our IV estimates to alternative specifications.

A note on the local nature of IV estimates and their interpretation

Note that the IV approach identifies the effect of social media consumption on body satisfaction on *compliers*, that is the students that do use social media because they have access to the internet at home or because their peers do. Technically, we estimate in Table 2 a Local Average Treatment Effect (LATE) of social media consumption on body dissatisfaction. This effect is for students complying with the instruments and may differ from the average effect of social media consumption on body dissatisfaction. In particular, compliers in columns (3) to (6) are students that use social media because their peers do so as well. The effect of social media on body image among these students may be different. This will be the case for example if peers' usage of social media exacerbates their detrimental effect on girls' body dissatisfaction (see related discussion and theoretical arguments provided for this claim in the main text). In that case, the effect of social media estimated in columns (3) to (6) of Table 2 should be larger, something we indeed observe. This is why we explain in the main text that our IV results are consistent with the idea that peers' usage of social media may exacerbate their negative effect on girls' body dissatisfaction.

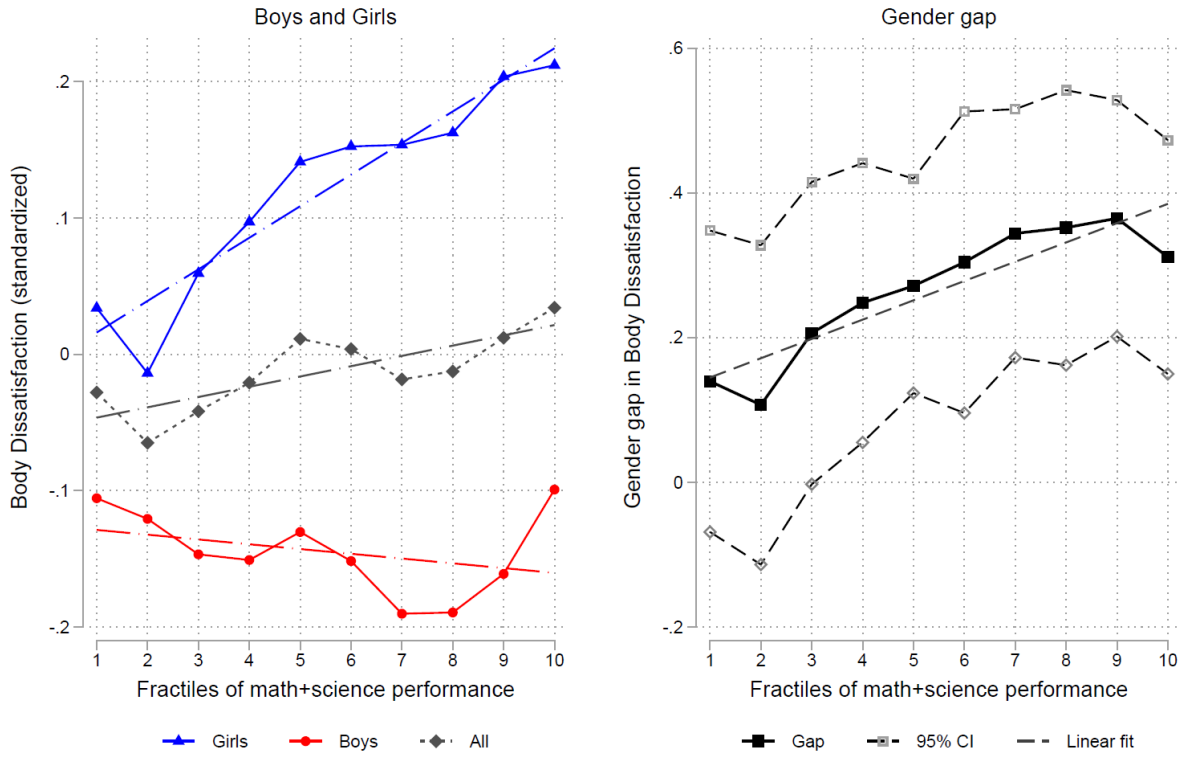
Appendix C: Additional Figures: Figures S1 to S7

Figure S1: Girls' and boys' body dissatisfaction as a function of their Body Mass Index (BMI)



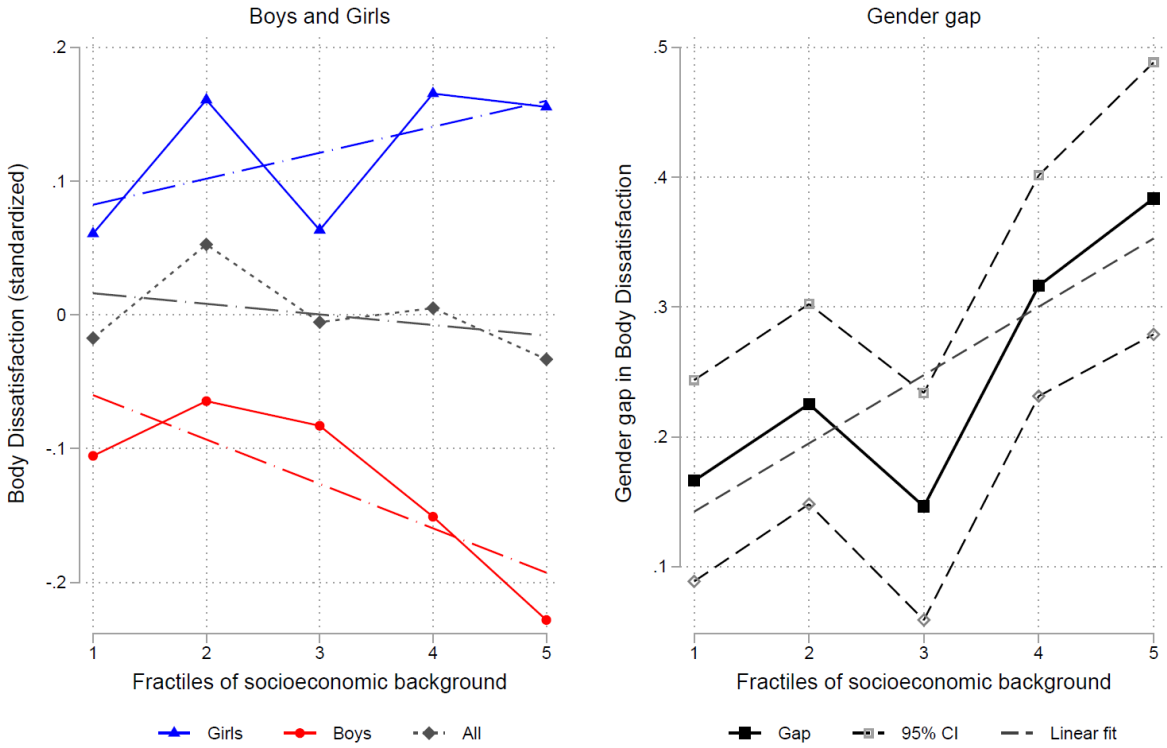
Note: The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818). The samples of girls and of boys are split in 20 vintiles of BMI containing the same number of observations. Body dissatisfaction is standardized to have a mean of 0 and a standard deviation of 1 in each country in the sample.

Figure S2. Body Dissatisfaction (BD) for boys and girls by decile of academic performance



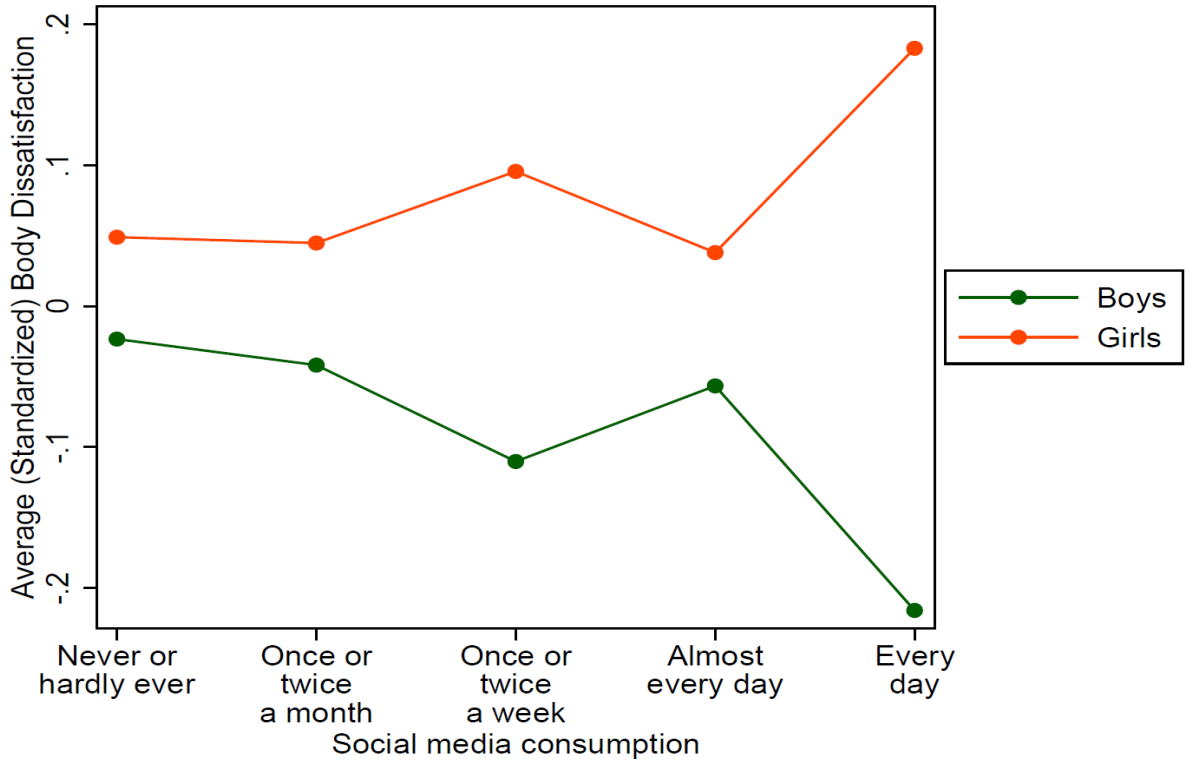
Notes: Analyses based on a sample of 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818). The left figure shows the mean levels of Body Dissatisfaction BD for boys (red), for girls (blue), for all (black) by deciles of math+science performance. The right figure shows the difference between girls' and boys' level of BD (girls-boys) in each decile of math+science performance, as well as the associated 95% confidence intervals. Science and math performance are standardized to have a weighted mean equal to zero and a weighted standard deviation equal to one in each country in the sample and averaged afterwards. The average is then split in deciles. Body dissatisfaction is also standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. The variables are described in detail in Appendix A. Estimates and standard errors involving measures of ability are based on plausible values and account for measurement error in these abilities on top of standard sampling error (see details in Appendix B).

Figure S3. Body Dissatisfaction (BD) for boys and girls by quintiles of socioeconomic background



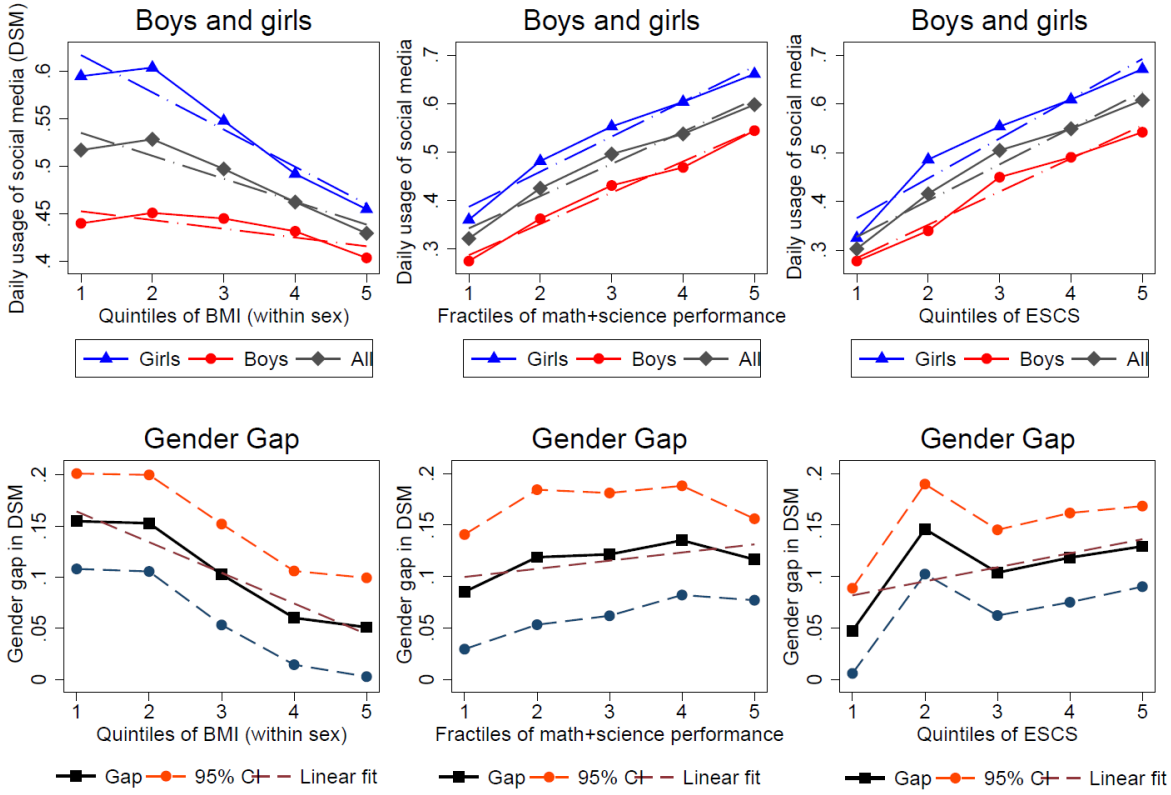
Notes: Analyses based on a sample of 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818). The left figure shows the mean levels of Body Dissatisfaction BD for boys (red), for girls (blue), for all (black) by quintiles of students' socioeconomic background. The right figure shows the difference between girls' and boys' level of BD (girls-boys) in each quintile of students' socioeconomic background, as well as the associated 95% confidence intervals. Socioeconomic background is an index provided by PISA (ESCS) that is standardized to have a weighted mean equal to zero and a weighted standard deviation equal to one in each country in the sample. Body dissatisfaction is also standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. The variables are described in detail in Appendix A.

Figure S4: Girls' and boys' body dissatisfaction as a function of their social media consumption (whole sample)



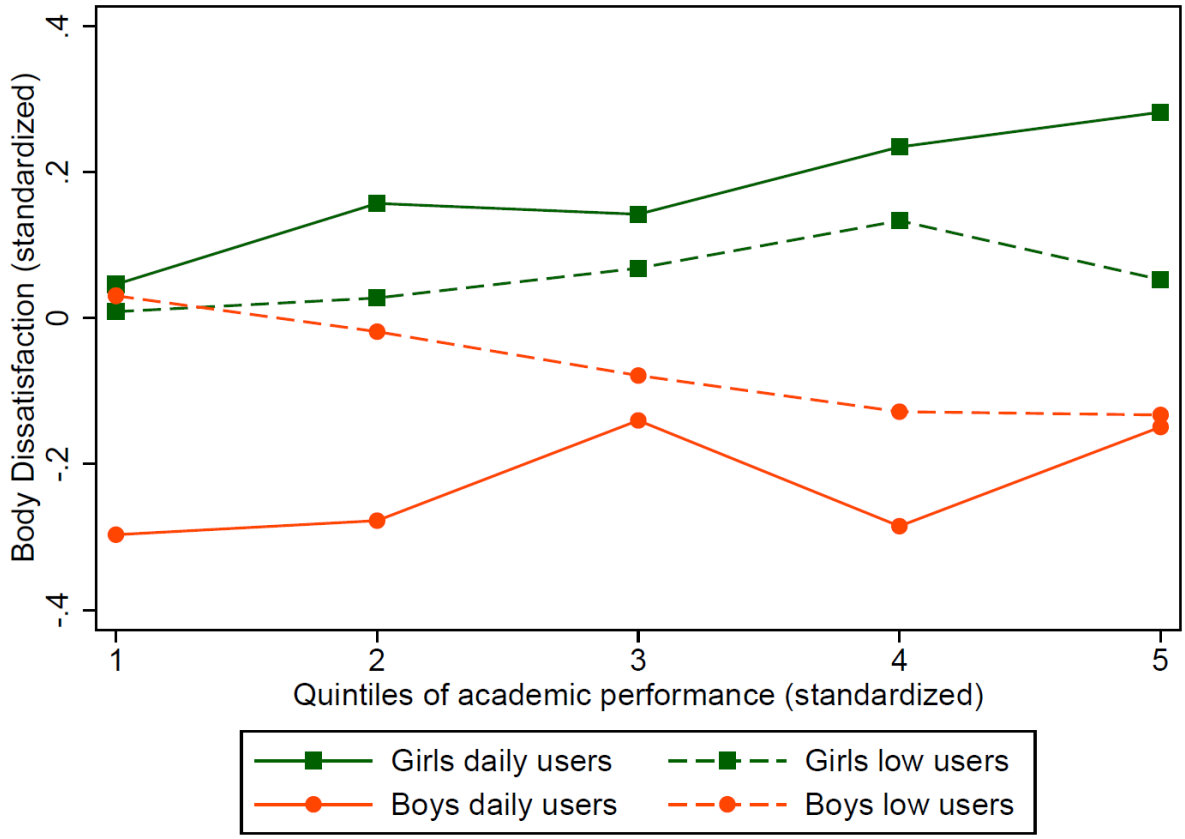
Note: The Figure provides the average level of body dissatisfaction (index standardized by country) for girls and boys (weighted using sample weights) depending on their usage of social media. The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818).

Figure S5: Daily usage of social media for girls and boys depending on their BMI, academic performance, or socioeconomic background



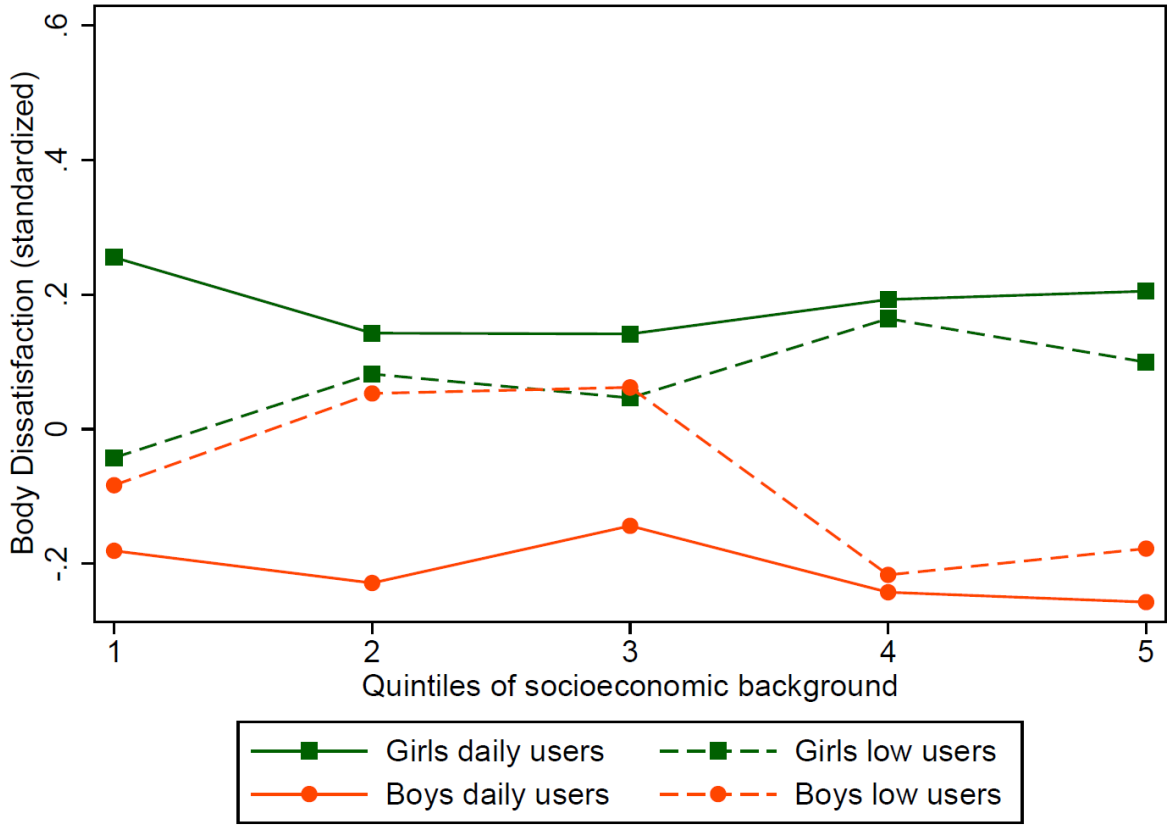
Notes: The top panel of the figure shows the percentage of boys (in red), of girls (in blue) and of all students (in black) who are daily users of social media as a function of quintiles of BMI (defined separately for each sex), academic performance (average of math and science performance standardized by country), or socioeconomic background. The bottom panel shows the corresponding gender gaps.

Figure S6: Link between social media consumption and body dissatisfaction as a function of academic performance



Note: The Figure provides the average level of body dissatisfaction (standardized index) for girls and boys depending on their usage of social media and their academic performance (5 quintiles). Academic performance is the average of the first plausible values of math and science performance which have previously been standardized to have a mean of 0 and standard deviation of 1 in each country in the sample. The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818).

Figure S7: Link between social media consumption and body dissatisfaction as a function of socioeconomic background



Note: The Figure provides the average level of body dissatisfaction (standardized index) for girls and boys depending on their usage of social media and their socioeconomic and cultural background (5 quintiles of the index escs provided in PISA). The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818).

Appendix D: Additional Tables: Tables S1 to S8

Table S1. Gender gaps in Body dissatisfaction by country: average gaps and gaps conditional on various sets of control variables

	Gender Gap (girls minus boys) in Body Dissatisfaction				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
All PISA2018 Countries	0.25	0.24	0.25	0.24	0.22
By country:					
Spain	0.38	0.38	0.39	0.39	0.33
Ireland	0.55	0.58	0.58	0.64	0.52
Mexico	0.20	0.18	0.21	0.19	0.18
Bulgaria	0.21	0.24	0.28	0.30	0.26
Georgia	0.24	0.21	0.27	0.28	0.27
Hong Kong	0.37	0.37	0.37	0.34	0.30
Panama	0.07	0.08	0.11	0.10	0.02
Serbia	0.15	0.16	0.21	0.23	0.18
CONTROLS	**NONE**	Individual controls	Body Mass Index (BMI)	Indiv controls+BMI	Indiv controls +BMI+ well-being

Notes: The table presents the gender gap in Body Dissatisfaction, with various controls. The measure of Body Dissatisfaction (BD) relies on questions about the students' look, weight and body and is standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. Individual controls are detailed in Appendix A. They include the level of education of the student's parents, measured both in years and kind of diploma obtained, grade repetition, measures of economic, social and cultural status of the household, measures of home educational resources and home possessions, and measures of performance in science and math (standardized by country). BMI is also described in Appendix A. Controls for well-being in the last column include students' levels of self-efficacy, of life satisfaction, of meaning in life, of negative feelings and an index of individual subjective well-being. These items are also described in Appendix A.

Table S2: Relation between Body Dissatisfaction and Life satisfaction, self-efficacy, meaning in life, or feeling miserable. Linear models for girls, boys and both sexes

Dep. Var.	Life satisfaction			Feeling miserable			Self-Efficacy			Life meaning		
	Girls	Boys	All	Girls	Boys	All	Girls	Boys	All	Girls	Boys	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
BD	-0.321***	-0.297***	-0.297***	0.235***	0.158***	0.158***	-0.294***	-0.226***	-0.226***	-0.296***	-0.279***	-0.279***
	(0.00566)	(0.00623)	(0.00630)	(0.00570)	(0.00665)	(0.00652)	(0.00579)	(0.00659)	(0.00652)	(0.00574)	(0.00659)	(0.00650)
Girl			-0.0625***			0.331***			0.0735***			0.0992***
			(0.00838)			(0.00874)			(0.00878)			(0.00872)
BD*Girl			-0.0242***			0.0771***			-0.0683***			-0.0166*
			(0.00843)			(0.00873)			(0.00876)			(0.00872)
Constant	-0.0162***	0.0462***	0.0462***	0.155***	-0.176***	-0.176***	0.0458***	-0.0277***	-0.0277***	0.0464***	-0.0527***	-0.0527***
	(0.00585)	(0.00599)	(0.00606)	(0.00593)	(0.00644)	(0.00631)	(0.00602)	(0.00640)	(0.00633)	(0.00596)	(0.00639)	(0.00630)
Observations	25,598	23,991	49,589	24,936	23,342	48,278	24,874	23,235	48,109	24,994	23,334	48,328
R-squared	0.112	0.087	0.105	0.064	0.023	0.079	0.094	0.048	0.072	0.096	0.072	0.084

Notes: The table presents estimates of linear regressions of Life satisfaction, self-efficacy, meaning in life and feeling miserable on Body Dissatisfaction (BD), for girls and boys separately, and for all. Regressions are run for girls, for boys, and for all students with a dummy for girls and the dummy interacted with the variable BD. The variable BD relies on students' answers to questions about their look, their weight and their body. The variable 'Life satisfaction' relies on students' answers to the question 'Overall, how satisfied are you with your life as a whole these days? on a scale from 0 to 10. The variable 'Feeling miserable' relies on students' answers to the question: 'how frequently ("never", "rarely", "sometimes", "always") do you feel miserable?' Self-efficacy and meaning in life are measured by indices provided by PISA, relying on a series of items. All variables are standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. See Appendix A for more details on these variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table S3. Percentage of boys and girls who are daily users of social networks

	ALL	Girls	Boys
All Countries	47.57%	52.68%	42.14%
Spain	69.61%	78.65%	60.34%
Ireland	60.07%	66.99%	52.93%
Mexico	40.53%	44.61%	36.15%
Bulgaria	41.66%	47.10%	36.20%
Georgia	44.04%	47.37%	40.44%
Hong Kong	63.71%	70.01%	57.08%
Panama	43.16%	46.37%	40.09%
Serbia	48.30%	55.06%	40.57%

Notes: The sample includes 8 countries: Spain (N=23,553), Ireland (N=4,664), Mexico (N=5,235), Bulgaria (N=2,944), Georgia (N=3,252), Hong-Kong (N=4,602), Panama (N=3,471) and Serbia (N=3,818). Students are described as intensive users of social networks if they declare participating in social networks at least once a day.

Table S4: Correlation between various daily internet usages

	Social media	Browse for fun	Browse for practical things	Play online games	Chat online	Upload crea-tive content	Emails	Down-load con-tent
Panel A: All students								
Social media	1.00	0.46	0.24	0.21	0.68	0.20	0.19	0.29
Browse for fun	0.46	1.00	0.35	0.35	0.45	0.25	0.22	0.41
Browse for practical things	0.24	0.35	1.00	0.31	0.22	0.46	0.35	0.49
Play online games	0.21	0.35	0.31	1.00	0.22	0.32	0.33	0.36
Chat online	0.68	0.45	0.22	0.22	1.00	0.16	0.19	0.26
Upload creative content	0.20	0.25	0.46	0.32	0.16	1.00	0.33	0.47
Emails	0.19	0.22	0.35	0.33	0.19	0.33	1.00	0.29
Download content	0.29	0.41	0.49	0.36	0.26	0.47	0.29	1.00
Panel B: Girls only								
Social media	1.00	0.41	0.19	0.16	0.66	0.16	0.13	0.24
Browse for fun	0.41	1.00	0.33	0.25	0.38	0.24	0.17	0.40
Browse for practical things	0.19	0.33	1.00	0.26	0.17	0.40	0.26	0.46
Play online games	0.16	0.25	0.26	1.00	0.16	0.28	0.26	0.30
Chat online	0.66	0.38	0.17	0.16	1.00	0.13	0.14	0.21
Upload creative content	0.16	0.24	0.40	0.28	0.13	1.00	0.24	0.44
Emails	0.13	0.17	0.26	0.26	0.14	0.24	1.00	0.21
Download content	0.24	0.40	0.46	0.30	0.21	0.44	0.21	1.00
Panel C: Boys only								
Social media	1.00	0.52	0.31	0.32	0.68	0.25	0.26	0.34
Browse for fun	0.52	1.00	0.37	0.43	0.53	0.26	0.27	0.42
Browse for practical things	0.31	0.37	1.00	0.35	0.27	0.51	0.41	0.52
Play online games	0.32	0.43	0.35	1.00	0.32	0.36	0.36	0.42
Chat online	0.68	0.53	0.27	0.32	1.00	0.20	0.25	0.32
Upload creative content	0.25	0.26	0.51	0.36	0.20	1.00	0.41	0.49
Emails	0.26	0.27	0.41	0.36	0.25	0.41	1.00	0.36
Download content	0.34	0.42	0.52	0.42	0.32	0.49	0.36	1.00

Notes: The Table shows the correlation between the various usages students make of the Internet. Each usage is measured with a dummy variable equal to one of the student declares resorting to this usage every day. Going on social media is most highly correlated with chatting online. The sample includes 8 countries (see other Tables' notes).

Table S5: Loadings of internet usages on the first two components of their Principal Component Analysis (PCA)

	All students		Girls only		Boys only	
	First Component of PCA	Second Component of PCA	First Comp.	Second Comp.	First Comp.	Second Comp.
Social media	0.35	-0.53	0.36	-0.54	0.36	-0.47
Browse for fun	0.38	-0.23	0.40	-0.17	0.37	-0.30
Browse for practical things	0.37	0.29	0.38	0.30	0.37	0.31
Play online games	0.33	0.18	0.31	0.22	0.34	0.07
Chat online	0.34	-0.55	0.34	-0.56	0.35	-0.52
Upload creative content	0.35	0.37	0.35	0.35	0.34	0.43
Emails	0.30	0.27	0.26	0.23	0.32	0.30
Download content	0.40	0.21	0.41	0.23	0.38	0.22

Notes: The Table shows how the first two components of a principal component analysis of possible internet usages load on these usages (factor loadings). The sample includes 8 countries (see other Tables' notes).

Table S6: Heterogeneity of the link between social media consumption and body dissatisfaction for girls according to BMI, academic performance and socioeconomic background

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: Index of Body Dissatisfaction				
Daily user of social media (DSM)	-0.0777	0.114***	0.0733**	-0.171	-0.161
	(0.195)	(0.0307)	(0.0345)	(0.194)	(0.215)
BMI	0.0379***			0.0396***	
	(0.00635)			(0.00632)	
DSM*BMI	0.0112			0.0104	0.00815
	(0.00896)			(0.00890)	(0.00982)
Academic performance		0.0111		0.00619	
		(0.0276)		(0.0272)	
DSM*Academic performance		0.0673*		0.0924**	0.129***
		(0.0347)		(0.0370)	(0.0414)
Socioeconomic background			0.0441**	0.0665***	
			(0.0184)	(0.0199)	
DSM*Socioeconomic background			-0.0513**	-0.0762***	-0.104***
			(0.0258)	(0.0289)	(0.0330)
Additional controls	No	No	No	No	Yes
Observations	21,966	26,510	26,440	21,932	17,232
R-squared	0.033	0.007	0.006	0.040	0.081

Notes: Results from linear regression models. The sample only includes girls. BMI is winsorized at the 1st and 99th percentile of the distribution of the variable. Academic performance is the average of students' performance in math and science (first plausible values) which have previously been standardized to have a mean of 0 and a standard deviation of 1 in each country. Controls in the last columns include country fixed effects, quintiles of BMI and a rich set of other students' characteristics (including their academic performance in math and science and measures of their socioeconomic background). See Appendix A for details on these controls. Observations are weighted using students' sampling weights. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table S7: Girls only. First stage results for IV (two-stage least square) estimates provided in Table 2

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: daily usage of social media					
Own internet access at home	0.268***	0.201***	0.156***	0.164***	0.166***	
	(0.0151)	(0.0186)	(0.0208)	(0.0273)	(0.0274)	
Other students in school average internet access			0.277***			
			(0.0423)			
Other girls in school average internet access				0.211***	0.143*	0.188**
				(0.0686)	(0.0735)	(0.0738)
Boys in school average internet access				0.0838	0.0258	0.117
				(0.0754)	(0.0827)	(0.0816)
Controls:						
Country fixed effects and BMI	Yes	Yes	Yes	Yes	Yes	Yes
Students and households characteristics	No	Yes	Yes	Yes	Yes	Yes
Peers' characteristics	No	No	No	No	Yes	Yes
Observations	21,739	17,078	16,400	10,330	10,330	10,330
R-squared	0.101	0.111	0.106	0.092	0.097	0.093
p-val	0	0	0	0.00415	0.0935	0.0868
Kleibergen-Paap Wald rk F statistic	313.6	116.8	77.97	33.60	19.50	10.78
Weak-instrument-robust inference:						
Anderson-Rubin Wald F statistic	10.82	4.524	10.24	4.391	2.115	2.422
p-val	0.00101	0.0332	0	0.00430	0.0960	0.0888
Anderson-Rubin Wald chi-sq stat.	10.82	4.537	20.55	13.24	6.405	4.889
p-val	0.00100	0.0324	0	0	0	0
Stock-Wright LM S statistic	10.84	4.575	20.34	13.16	6.530	5.012
p-val	0.000992	0.0334	0	0.00429	0.0885	0.0816
Over-identification:						
Hansen J (over-identification)	NA	NA	5.017	1.778	1.436	0.467
p-val Hansen J	NA	NA	0.0251	0.411	0.488	0.494

Notes: The Table provides for girls the results from a regression of daily usage of social media on internet access at home and the average internet access at home of peers. Robust standard errors in parentheses. Controls are identical to those in Table 2 in each corresponding column. See Table 2 notes for details. *** p<0.01, ** p<0.05, * p<0.1

Table S8: Boys only. Link between social media consumption (daily usage) and body dissatisfaction. Estimates from instrument variable models.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Index of Body Dissatisfaction					
Daily usage of social media	-0.440***	-0.197	-0.165	-0.263	-0.279	0.0503
(instrumented)	(0.136)	(0.215)	(0.178)	(0.214)	(0.275)	(0.388)
Instruments:						
Own internet access at home	Yes	Yes	Yes	Yes	Yes	No
Other students in school average internet access	No	No	Yes	No	No	No
Other boys in school average internet access	No	No	No	Yes	Yes	Yes
Girls in school average internet access	No	No	No	Yes	Yes	Yes
Controls:						
Country fixed effects and BMI	Yes	Yes	Yes	Yes	Yes	Yes
Students and households' characteristics	No	Yes	Yes	Yes	Yes	Yes
Peers' characteristics	No	No	No	No	Yes	Yes
Observations	21,411	16,234	15,495	12,771	12,771	12,771

Notes: Results from two-stage least squares regression models. The sample only includes boys. When internet access of peers is used as instrument, only schools with at least 9 peers are included in the sample (to avoid noisy estimates). This amounts to keep schools with 10 students or more in column (3) and with at least 10 boys and 10 girls in subsequent columns. This selection rule has little impact on the final results. Averages for peers are computed taking the mean of the variables considered across all relevant peers (*e.g.*, the share of other girls in the school having access to the internet). Controls are similar to those in Table 1 (see Appendix A for details) with the exception that for peers, we replace the set of dummies mother education, father education and number of books at home by one single continuous variable (in order to avoid increasing dimensionality too much). This again has little impact in the final results. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1