Zoom in: A Psycho-Institutional Intervention Improves Minority Student’s Grades in Two Field Experiments

Kinneret Endevelt  
The Hebrew University

Eran Halperin  
Hebrew University  https://orcid.org/0000-0002-3379-2935

Roni porat ( roni.porat@mail.huji.ac.il )  
The Hebrew University  https://orcid.org/0000-0002-0665-4131

Article

Keywords:

Posted Date: March 10th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2578351/v1

License: 🍀 This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Improving minority’s achievements in higher education is a crucial step towards closing the inequality gap. In this paper we introduce a new psycho-institutional approach targeting institutional changes that affect psychological processes that are necessary for minority’s success. We tested this approach in two field experiments in Israel (n > 20,000). Our intervention involved reconstructing institutional virtual spaces in a manner that signals to minority students that they belong. This was achieved by manipulating how lecturer’s name appeared on Zoom so that instead of the default option (i.e., Hebrew or English only) it appeared in three languages: Arabic, Hebrew, and English. This approach positively affected minority students’ sense of belonging, class participation, and grades (a 7.5-point class grade increase in experiment 1, and a 4-point semester grade increase in experiment 2). Taken together, these studies provide a new framework for designing interventions that can meaningfully impact minority’s outcomes.

Introduction

How can we improve minority achievements in higher education? For over two decades social scientists from various disciplines have been studying this key question. Scholar’s focus on higher education stems from findings suggesting that completing a bachelor’s degree is associated with a host of socio-economic benefits like one’s position in the labor market, productivity, career opportunities, and lifelong earnings. This implies that obtaining an academic degree can serve as an equalizer that can meaningfully reduce inequality and increase minority’s social mobility. However, studies repeatedly demonstrate that disparities in higher education loom large. In the U.S there is a 14 percentage difference in bachelor degree completion rates between non-Hispanic White and Black people. This is also true in Israel, where there is an 18 percentage difference in degree completion rates between Jewish and Palestinian citizens. These disparities imply that the very place that could potentially serve as the platform for social mobility, sometimes amplifies the gap and feeds back into the inequality cycle.

Over the years two approaches to improving minority achievements in higher education emerged. The first, termed here as a psychological approach, focuses on the individual needs and challenges minority students face. The second, termed here as an institutional approach, is largely situated within economics and political science, and focuses on modifying policies, rules, and programs within the academic institution. In what follows we review these approaches, highlighting findings, contributions, and limitations, and propose a new psycho-institutional approach that combines the best of both.

The psychological approach targets the individual needs and difficulties minority students face within higher education institutions (for a recent review and meta-analysis see Solanki et al., 2020). Academic settings where majority group members hold a disproportional number of prestigious positions like faculty positions present unique challenges for minority students. This includes a lack of support system, the absence of role models, and the feeling that they are not respected, and are not supported by
advantaged group members\textsuperscript{15}. Thus, minority students often experience a lingering feeling that they don't belong in these academic spaces\textsuperscript{16,17}. As a result, they are reluctant to actively participate in institutional activities like classes or social events in fear that their behavior will confirm a negative stereotype that is associated with their group\textsuperscript{18}. This sense of not belonging can undermine their performance\textsuperscript{17,19,20} and by that increase the chances of them dropping out\textsuperscript{21–23}.

To address the challenge of feeling that one doesn't belong psychologists developed and tested several interventions. These can be characterized as interventions that (1) target exclusively minority group members, that (2) need to make a meaningful effort to correct for their discrimination, and (3) this effort is typically a change in one's mindset or mental model rather than a change of the system in which discrimination occurs. These kinds of interventions include, for example, social belonging interventions\textsuperscript{24,25}, where senior students from similar backgrounds share their personal story about their experience transitioning into college, with an emphasis on the challenges they faced and how those got easier over time\textsuperscript{24}. These personal stories are intended to make minority students feel that the lack of belonging they are experiencing is normal, temporary, and widespread\textsuperscript{25}. Another psychological intervention is value-afﬁrmation that encourages students to reflect on enduring positive aspects of themselves\textsuperscript{26,27}, by identifying a core personal value, and then reﬂecting on experiences where the value had been important to them\textsuperscript{28}. Peer mentoring programs are also psychological interventions that typically involve senior minority students providing social, emotional and academic support to newcomer minority students\textsuperscript{29–32}. Peer mentoring has been shown to increase social support\textsuperscript{33}, social belonging and self-efficacy\textsuperscript{30,33}. Some studies ﬁnd that the most effective mentorship occurs when mentors come from minority populations\textsuperscript{30}. Finally, a recent intervention involved communicating to minority students that their peers hold pro-diversity attitudes and engage in inclusive behaviors. This intervention effected minorities sense of belonging as well as their grades\textsuperscript{34}.

While psychological interventions have shown some success\textsuperscript{35,36}, they are limited for two main reasons. First, a number of studies point out that these intervention require careful and delicate crafting to yield positive outcomes\textsuperscript{35,37–40}, making them implausible to scale-up by practitioners\textsuperscript{41}. Second, psychological interventions target minority group members rather than majority groups members. As such, they rely on meaningful efforts by minority students to “fix” their own discrimination. Recently, some concerns were raised\textsuperscript{42} regarding the implicit and normative message that stems from this approach. That is, intervening exclusively on minority students may suggest that failures in higher education should be attributed to students’ (lack of) persistence and performance rather than systemic failures and biases inherent to higher education institutions.

One alternative route to the psychological approach is an institutional approach, that encompasses interventions targeting policies, rules, and regulations within the institution changing how it operates\textsuperscript{43,44}. These changes aim to equalize the playing field by changing the rules of the game. One example are affirmative action programs that aim to provide equal opportunities to minorities to gain access and
succeed in higher education\textsuperscript{45}. These include admissions policies that consider race, ethnicity or gender to support the academic and social success of minority students\textsuperscript{44}. Another example are scholarship programs that aim to provide financial support to minority students in their pursuit of higher education. Such programs often offer financial assistance, mentorship, and other support services to help students overcome material barriers such as financial constrains or lack of access to resources\textsuperscript{6,46}. These institutional efforts have proved to be quite successful in addressing minorities material needs\textsuperscript{6}. However, because these programs require major funding, they are less feasible and accessible to many higher education institutions. By addressing minorities material needs but not their psychological needs, these interventions fall short in materializing these student’s full potential.

We propose a third route to addressing inequality in higher education – a psycho-institutional approach that directly targets the psychological needs and challenges of minority students, without placing the burden of addressing discrimination on their shoulders. We are not alone in this call to take an institutional perspective in psychology’s work on inequality. In recent years similar calls have been made by leading scholars in the field\textsuperscript{47–51}. At the same time the discipline is still silent on how in practice this could be done. Thus, our goal was to design a psycho-institutional intervention that is tailored to address the psychological needs and challenges of minority students. To do so, we take the best from both worlds described earlier. We propose designing interventions that (1) target institutions that are typically controlled by majority group members instead of targeting minority students; (2) The intervention should modify or reconstruct a component of the institution rather than a mental model or mindset of its students, faculty or administration members; (3) Reconstructing or modifying the institution should be done in a manner that directly addresses a unique psychological challenge or barrier; (4) The intervention should be easy to implement and scale up in terms of its design and cost.

In the current research we design a psych-institutional intervention that attempts to directly address minorities’ lack of belongingness. We do so by reconstructing institutional spaces in a manner that signals to minority students that they belong, and that the institution sees them as important and integral members. We hypothesized that this renewed perception can potentially alleviate concerns about not being valued, respected, and included, and as a result, increase their sense of belonging. Enhanced belonging may in turn lead these students to be more active learners which may result in better academic achievements\textsuperscript{53}.

We tested the intervention in two field experiments in Israel, looking at Palestinian minority students at the Hebrew University. While several social groups within Israel suffer from inequality, there is wide consensus that the Palestinian minority is the most marginalized\textsuperscript{54}. The Palestinian minority suffers from both social and institutional discrimination, resulting in deprivation in almost all life domains\textsuperscript{55}. One salient indicator of their marginalization is the absence of the Arabic language from the Israeli public sphere\textsuperscript{56,57}. Open-ended interviews conducted with Palestinian students revealed that the absence of the Arabic language from the Hebrew University campus creates feelings of alienation and distress. This qualitative evidence is supported by scholarly work suggesting that language is important for one’s sense
of belonging and feeling of acceptance. We hypothesized that an institutional change that increases the salience of Arabic in the public sphere would send a signal of inclusion that could enhance Palestinians student’s sense of belonging and by that increase their participation in class and improve their academic achievements.

The experiments were carried out during the COVID-19 pandemic, a time where minority students were disproportionally harmed in higher education as they experienced more technical and learning challenges, financial insecurity, and higher rates of mental health disorders. Given that classes at the Hebrew University were held remotely during this period, the institutional change to the public sphere involved changing the way the lecturer’s name appeared on Zoom so that instead of the default (i.e., English or Hebrew only) which served as the control condition, the name would appear in three languages: Arabic, Hebrew, and English (experimental condition) (see Fig. 1–2 for the experiments timeline and a depiction of the intervention).

Study 1 was conducted during the spring semester 2020, which was taught fully remote. We randomly assigned 34 classes (n = 834) from two faculties (Social Science and Humanities) to either receive an institutional signal of inclusion (17 classes) or not (17 classes). Lecturers in the experimental condition were asked to change how their names appeared four weeks before the end of the semester. We evaluated students’ achievements by examining their final grade for the course. Furthermore, we assessed student participation in class by coding their attendance to class and camera usage in the Zoom recordings before and after the intervention was administered. Finally, we evaluated students’ sense of belonging, using a questionnaire administered at the end of the semester (for a detailed timeline, see Fig. 1).

Study 2 was conducted during the spring semester 2021 that was remote for the first two weeks. We randomly assigned the 54 departments at the Hebrew University (out of 90 departments in total) that had both Jewish and Palestinian students (n = 20,733 students) to either experimental (n = 27) or control (n = 27) conditions. Lecturers in the experimental condition were asked by the university rector to change how their names appeared on Zoom at the start of the semester. Four weeks after the start of the intervention, we administered a questionnaire measuring the same outcomes as in Study 1.

In mid-semester (May 10th -21th ) there was a violent outbreak between the Jewish and Palestinian citizens of Israel that included street lynching and nightly riots in shared cities where Jews and Palestinians live side by side. These included riots in Jerusalem and at the Hebrew University’s main gate. In addition, these events were accompanied by daily Israeli air strikes on the Gaza Strip and the firing of rockets from the Gaza Strip to cities across Israel. Four weeks after these events, we administered another questionnaire to measure our primary outcomes. Finally, at the end of the semester, we obtained student’s grades at the departmental level. Details about the recruitment, design, representativeness of the samples, and robustness checks, are provided in the method section and in the supplementary information.
Results

Our pre-registered estimation strategy included multilevel modeling, controlling for covariates (gender, Palestinian proportion, class/department size and others specified in the method section), accounting for clusters (classes in Study 1 and departments in Study 2) with intercept random effect. While some variables (i.e., achievements) were measured for the entire population, others (i.e., belonging) were measured with a sample of the population where differential survey attrition was corrected with weights\(^6^3\) (see Table 1). We replicated our results with two alternative models. The first is a linear model controlling for covariates with robust standard errors. The second is the same linear model without controlling for covariates (see supplementary materials for results of both models). Both models replicate the main findings.

We begin by reporting findings on students’ achievements, the primary outcome for both Studies 1 and 2 (see Fig. 3). We found that the grades of Palestinian students in the experimental condition were higher in both studies than the grades of those in the control conditions. In Study 1, we found a significant effect, where Palestinian students grades in the experimental condition were 7.5 points higher (on a 0-100 scale) than their counterparts in the control condition \((b = 11.35, 95\% \text{ CI} = [5.41, 17.30], SE = 2.99, t(81) = 3.80, p < .001)\). In study 2, we found a marginally significant effect, where their grades in the experimental condition were on average 4.10 points higher than their counterparts in the control condition \((b = 4.12, 95\% \text{ CI} = [-2.3, 8.47], SE = 2.16, t(50) = 1.90, p = .063)\). We did not find significant differences for the Jewish students in Study 1 \((b = 2.45, 95\% \text{ CI} = [-1.22, 6.11], SE = 1.87, t(731) = 1.31, p = .191)\) and found marginally significant differences in Study 2 \((b = 1.94, 95\% \text{ CI} = [-.06, 3.95], SE = 1.00, t(50) = 1.95, p = .057)\). Given that we obtained differences for both Palestinian and Jewish students in Study 2, we wanted to make sure that these differences were the result of the intervention, rather than an inherent difference between the departments assigned to the experimental condition. To do so, we compared the departmental average from the previous year (spring 2020), and did not find a significant difference between experimental and control departments in the 2020 grades \((b_{Palestinians} = 2.94, 95\% \text{ CI}_{Palestinians} = [-.60, 6.47], SE_{Palestinians} = 1.76, t_{Palestinians}(50) = 1.67, p_{Palestinians} = 0.101; b_{Jews} = 1.07, 95\% \text{ CI}_{Jews} = [.76, 2.91], SE_{Jews} = .92, t_{Jews}(50) = 1.17, p_{Jews} = 0.246)\), providing further evidence that the difference we obtained was the result of the intervention rather than unbalanced assignment of departments into treatment.

Next, we examined what might explain these differences. We first turn to test whether the intervention affected student’s behavior. We measured behavior in Study 1 using two measures: First, we coded student’s attendance in five lectures (one prior to the intervention, and four after its administration). Second, we coded whether students had their camera turned on in these five lectures. For the Palestinian students, we interpret the behavioral outcomes cautiously given the relatively low number of participants \((n = 96)\) (see Fig. 4). For student attendance, we find that in the lecture before the start of our intervention 19 students in the experimental condition attended, and 27 students in the control condition attended. For the experimental condition we obtain an increase in attendance that is maintained throughout the four
lectures ($n_{\text{class1}} = 28; n_{\text{class2}} = 28; n_{\text{class3}} = 28; n_{\text{class4}} = 27$). For the control condition we obtain a gradual decline in attendance over the next four lectures ($n_{\text{class1}} = 21; n_{\text{class2}} = 21; n_{\text{class3}} = 18; n_{\text{class4}} = 22$). For camera usage we obtain a similar trend. We find that in the lecture before the start of our intervention nine students in the experimental condition and eight students in the control condition had their camera turned on. For the experimental condition we obtain an increase in camera usage that is maintained to some extent throughout the four lectures ($n_{\text{class1}} = 13; n_{\text{class2}} = 19; n_{\text{class3}} = 11; n_{\text{class4}} = 17$). For the control condition we obtain a gradual decline in camera usage over the next four lectures ($n_{\text{class1}} = 6; n_{\text{class2}} = 6; n_{\text{class3}} = 6; n_{\text{class4}} = 2$).

We conducted a regression analysis to test whether these differences were statistically significant and found a significant condition by time interaction for students' attendance ($b = .33, 95\% \text{ CI} = [.18, .49], SE = .08, t(467) = 4.18, p < .001$), and camera usage ($b = .16, 95\% \text{ CI} = [.04,.28], SE = .06, t(468) = 2.71, p = .007$). Turning to the simple effects, we found no pre-treatment differences between conditions in attendance rates and camera usage for Palestinian students' attendance rates ($b = -.18, 95\% \text{ CI} = [-.39, .03], SE = .11, t(467) = -1.63, p = .09$), or camera usage ($b = .02, 95\% \text{ CI} = [-.19,.23], SE = .11, t(468) = .17, p = .87$). Next, we examined the slope for each condition separately. We found a 21% significant increase in Palestinian attendance following the lecturer's name change on Zoom ($b = .21, 95\% \text{ CI} = [.10, .32], SE = .06, t(467) = 3.60, p < .001$). For the Palestinian students in the control condition, we found a significant decrease in attendance ($b = -.12, 95\% \text{ CI} = [-.23,-.02], SE = -.05, t(468) = -2.26, p = .02$). Similarly, following the intervention, we observed an 9.5% significant increase in camera usage among Palestinian students in the experimental condition ($b = .10, 95\% \text{ CI} = [.02, .19], SE = .04, t(468) = 2.39 p = .02$), while for the Palestinian students in the control condition, no significant difference was found ($b = -.06, 95\% \text{ CI} = [-.13, .02], SE = .04, t(468) = -1.41, p = .16$). However, when examining post-treatment differences between conditions, we did not find significant differences in attendance ($b = .15, 95\% \text{ CI} = [.02, .32] SE = .09, t(467) = 1.68, p = .09$) or in camera usage ($b = .18, 95\% \text{ CI} = [.01, .37] SE = .10, t(468) = 1.86, p = .09$) (see Fig. 5).

For the Jewish students, we found a significant condition by time interaction on attendance ($b = -.07, 95\% \text{ CI} = [-.12, -.01], SE = .03, t(3572) = -2.46, p = .014$) but not camera usage ($b = 0.01, 95\% \text{ CI} = [.05, .06], SE = .03, t(3585) = 0.28, p = .779$). Looking at the simple effects, we did not find pre-treatment differences across conditions in attendance rates ($b = .03, 95\% \text{ CI} = [-.10, 0.15], SE = .06, t(3572) = .45, p = .65$) or camera usage ($b = .03, 95\% \text{ CI} = [.11, .18], SE = .07, t(3585) = .46, p = .65$). However, following the intervention, for Jewish students in the experimental condition, we found an 8.5% significant drop in attendance ($b = -.08, 95\% \text{ CI} = [-.12,-.04], SE = .02, t(3572) = -3.94, p < .001$), while for the Jewish students in the control condition, no significant difference was found ($b = -.01, 95\% \text{ CI} = [-.05, 0.03], SE = .02, t(3572) = -.55, p = .58$). This is the only indication that we have for a backlash effect for the Jewish students. We did not observe significant differences following the intervention in camera usage for the Jewish students in both conditions. Finally, when examining post-treatment differences between Jewish participants in the control and experimental condition, no differences were found in attendance ($b = -.05, 95\% \text{ CI} = [.16, 0.08], SE = 0.06, t(3572) = -.68, p = .5$) or camera usage ($b = .04, 95\% \text{ CI} = [-.10, .18], SE
Given the magnitude of Study 2 (n = 20,733 students), we could not code for student's behavior in class (i.e., whether they attended or had their camera turned on).

Finally, we turn to the psychological mechanism we hypothesized that would drive these effects - Palestinian students' sense of belonging. We found that in both studies Palestinian students in the experimental condition felt a significantly higher sense of belonging following the intervention, compared with Palestinian students in the control condition (Study 1: \( b = .83, \text{95% CI} = [.34, 1.33], SE = .25, t(49) = 3.37, p = .001 \); Study 2: \( b = .31, \text{95% CI} = [.05, .56], SE = .13, t(362) = 2.38, p = .018 \)). However, in Study 2 this effect did not persist after the violent events (\( b = .17; \text{95% CI} = [-.06, .41], SE = .12, t(291) = 1.47, p = .142 \)).

Finally, Jewish students' sense of belonging was not affected by the intervention in both studies (see Fig. 6).

**Discussion**

In this paper, we introduced a new approach to improving minority achievements in higher education, the psycho-institutional approach, which addresses minority student's psychological needs and challenges by employing institutional changes. We tested this approach in two large-scale field experiments in Israel, where we developed an intervention based on our approach in which we modified the online environment where classes were held in a manner that signals to Palestinian minority students that they belong. This was achieved by having lecturer's change the way their name appeared on Zoom so that instead of the default option (i.e., Hebrew or English only), it would appear in three languages: Arabic, Hebrew, and English. The results of these experiments demonstrate that our intervention had a positive impact on minority students' sense of belonging, class participation, and grades. In the first experiment, minority students' class grades improved by 7.5 points, while in the second experiment, minority students' average semester grades improved by 4 points.

Our work contributes to both how we understand the problem of minorities in higher education and how we address it. With regards to how we understand the problem, existing approaches have defined the problem from strictly a psychological or institutional perspective. The psycho-institutional approach acknowledges and addresses both the unique psychological challenges alongside the structural inequality embedded within higher education institutions. With regards to how to address the problem, the psycho-institutional approach provides a holistic framework to designing effective interventions. We propose targeting institutions rather than individual minority group members, and modifying elements within the institution, rather than individual's mental model or mindset. Moreover, the psycho-institutional approach focuses on designing institutional changes in a manner that directly targets psychological barriers or challenges (e.g., lack of belonging), and is easy to implement and scale up in terms of its design and cost.

Our work has implications for scholars interested in developing scalable interventions in higher education and beyond. While we tested the psycho-institutional approach within the context of higher education, we believe it can be easily adapted to other contexts and other social goals. With regards to the latter, the
psycho-institutional approach can easily address other psychological constructs other than belonging like support for democratic values or the reduction of affective polarization. With regards to the former, one can easily imagine how psycho-institutional interventions can be designed to fit different spaces that are shared by minority and majority group members other than institutions for higher education. These can be the workplace, the hospital, or even the mall.

Our work also has implications for higher education institutions that are looking for practical and scalable interventions that will improve minority achievements in their respective institutions. Our easy-to-implement intervention can be implemented in other higher education institutions beyond Israel. For example, a US-based institution that wishes to signal to Latino populations that they are valued and included could easily implement this intervention by having lecturers write their name in Spanish as well.

At the same time, when addressing the scalability and generalizability of this research, several limitations emerge, that highlight areas for future research. First, our intervention targeted lack of belongingness Palestinian students often experience in Israeli higher education institutions. Our qualitative fieldwork suggested that one route to making these students feel included would be to restructure academic spaces so that they include the Arabic language. We utilized the outbreak of COVID-19 which shifted all academic activities to virtual online spaces. This made it quite easy to intervene in a similar manner across the university campuses. However how would a psycho-institutional intervention aiming to restructure spaces that are not virtual look like? Would, for example, changing the signs on campus so they include the Arabic language have the same positive impact? What would happen if some faculty members changed the signs outside their doors to include Arabic? Would that yield the positive effects we observed or perhaps, would the fact that only some faculty changed the way their name appeared would signal to Palestinian students that they are only accepted by some but not by others?

Second, our experiments tested outcomes over time. With regards to the main outcome variable, student's grades, we find an encouraging long-term effect in both experiments. In fact, even though in the second experiment the intervention was short in duration, we still find differences in Palestinian student's grades more than three months after the intervention had ended. With regards to the psychological mechanisms at play, in the first experiment, questionnaires were administered a few days after the end of the semester. In the second experiment questionnaires were administered two weeks after classes came back to campus. In both experiments we find that in the immediate term the intervention positively affected Palestinian students’ sense of belonging. In the second experiment we had hoped to measure the durability of the intervention's effects by employing questionnaires at the end of the semester (more then 12 weeks after the intervention had ended). Our 12-weeks-post-intervention-findings indicate that the enhanced feeling of belonging among Palestinian students in the experimental condition dissipated. Thus, while we observe long-term effects on the main outcome of interest, the effect on our mechanism seems to erode. There are several possible explanations to why this occurred. First, it may be that because the intervention lasted for only two weeks, it was just too weak of a signal to last over time. Second, it may be the case that when students went off Zoom and back to classes, the fact that they did not receive any new signal of inclusion in their new non-virtual space was a negative signal of its own.
Third, and most realistically, the violent events probably took their own tole eliminating any positive progress in Palestinians feelings of belonging to an institution controlled by the Jewish majority. Thus, more research is needed to fully understand how durable the effects of such interventions.

Third, our experiments tested one intervention designed under the principles of the psycho-institutional approach – restructuring academic spaces in a manner that signals inclusion. The intervention was customized to address the Palestinian minority in the Israeli context. However, this is just one way of implementing the psycho-institutional approach. Future research should examine other forms of structural changes that target other unique psychological barriers. This requires creativity and strong cooperation between researchers and institutions to identify both the psychological barriers that minorities face in their respective institutions, and the changes the institution can make to directly target those barriers.

Four, an alternative explanation to some of the findings could be that the intervention effected Palestinian students by impacting the lecturers themselves. It is possible that our intervention created an indirect process in which the lecturers behaved more inclusively following their name change on Zoom. This could be an additional mechanism by which Palestinian students were affected by the intervention. While our work does not address this mechanism, it is important for future work to consider this path of influence as well.

Taken together, the psycho-institutional approach represents a promising new approach for designing interventions that can meaningfully impact minority students' outcomes in higher education. By targeting psychological factors through institutional changes, this approach can address the complex interplay between individual and contextual factors that contribute to minority students' success. In this way, the psycho-institutional approach can help to create more inclusive environments that better support minority students' academic and personal growth.

### Table 1
Description of outcomes measured separately for Palestinian and Jewish students.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Measurement</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades</td>
<td>Final course grade for each student (Study 1)</td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>Average departmental grade (Study 2)</td>
<td></td>
</tr>
<tr>
<td>Class attendance</td>
<td>Individual student attendance rates in five classes (one class before the administration of the intervention, and four classes after) (Study 1).</td>
<td>Population</td>
</tr>
<tr>
<td>Camera usage</td>
<td>Individual information regarding camera usage in five classes (one class before the administration of the intervention, and four classes after) (Study 1).</td>
<td>Population</td>
</tr>
<tr>
<td>Sense of belonging</td>
<td>Nine self-report ratings in study 1, and three self-report ratings in study 2 (e.g., “I feel that I belong to my university”)</td>
<td>Sample</td>
</tr>
</tbody>
</table>
Methods

The research was approved by the Hebrew University’s IRB committee (#103120). Informed consent was obtained from all participants.

Random Assignment Procedures:

In Study 1 we randomly assigned 34 classes (up to 50 students each) from two faculties at the Hebrew University, to the experimental (n = 17) and the control (n = 17) condition. Assignment was done on the classroom level making this a clustered design. To reduce the chance of imbalance assignment to conditions and to improve precision, we blocked classes by size (small classes of up to 15 students, medium classes of 16–30 students, and large classes of 31–50 students), and by the proportion of Palestinian students in each class.

In study 2 we randomly assigned 54 departments, from 12 faculties at the Hebrew Universities’ six campuses (“Har Hatzofim”, “Givat ram”, “Rechovot”, “Asaf Harofe”, “Kaplan” and “Ein Carem”), to the experimental (n = 27) and the control (n = 27) condition. Assignment was done on the department level, making this a clustered design. We blocked classes by department size (small departments from up to 200 students, medium departments from 200–400 students and large departments from 400 students) to improve precision of our estimates.

Timeline Of Studies 1 & 2

Timeline – Study 1

Study 1 was conducted during Spring Semester 2020. Due to the outbreak of COVID-19, the Hebrew University announced on the 03.13.2020, prior to the start of the semester, that all classes will be fully remote and held via Zoom. On week Five (04.30.2020), we emailed 110 lecturers in two faculties (Social Science and Humanities) who teach relatively small classes (10–50 students) and asked them to participate in our experiment, letting them know that the experiment was approved by the University ethics committee (see Figure S-1 for the full timeline of both studies). Given that recordings on Zoom was limited to up to 50 faces, we restricted ourselves to classes that had less than 50 students so that we could code for their behavior during class. Overall, 34 lecturers agreed to participate in our experiment.

On week seven, we randomly assigned the 34 classes whose lecturers agreed to participate to either the experimental or control condition (see random assignment procedures bellow for more details). Lecturers in both conditions were asked to notify their students on week eight that the class was participating in a study, and as part of the study their class recordings would be watched and analyzed by researchers. On week 10, we assessed student's baseline participation in classes. During that week, after teaching their class, lecturers in the experimental condition were asked to change the way their name appeared on Zoom (see Fig. 2 for a depiction of the experimental and control conditions). To assist the lecturers, we
sent them detailed instructions on how to change the way their name appeared on Zoom, and we provided them with the Arabic transcript of their name. In cases where lectures asked for further assistance, the first author provided technical guidance by phone. We also ensured through the Zoom class recordings that all lecturers in the experimental condition indeed changed the way their name appeared on Zoom, and that those in the control condition did not.

On week 14, during the last class of the semester, lecturers provided their students with a link to the questionnaire that included the self-report outcome measures. A week after the semester had ended, (07.07.2020), we asked the lecturers to post on the course website the link to the same questionnaire and a reminder to students who did not answer yet, to fill it out. In addition, given the small number of Palestinian students who participated in the experiment (n = 96), three weeks after the semester had ended (07.21.2020) we sent Palestinian students only (in both experimental and control groups) the link directly with a remainder and a request in Arabic to participate.

**Timeline – Study 2**

Study 2 was conducted during Spring Semester 2021 that began on 03.14.2021 one week after the end of Israel's third COVID-19 lockdown. However, given the strict gathering restrictions, the University announced that the first two weeks of the semester will be fully remote. Four weeks before the start of the semester, we randomly assigned 54 departments (out of 90 departments in total at the Hebrew University) that consist of both Palestinian and Jewish students, to either experimental (n = 27) or control (n = 27) condition. Three weeks before the start of the semester (02.23.2021), the university’s Rector sent an email requesting all lecturers in departments assigned to the experimental condition to modify how their names appeared on Zoom (For the Rector’s letter see Fig. 1 in the supplementary information). As in Study 1, we provided lecturers in the experimental condition with instructions on how to change their names as well as the Arabic transcript of their name. When necessary, the first author assisted lecturers who requested help in modifying their names. Before the first week of the semester, we asked the lecturers in the experimental condition to inform us via email if they had modified their name as requested.

In Study 2, the intervention began during the first meeting of the semester, as opposed to Study 1, where the intervention was introduced more than midway of the semester. Two weeks after the semester begun, students went on a two-week holiday break (03.24.21–04.03.21), after which almost all undergraduate classes returned to an in-person format. On week four (04/08/2021), the Dean of Students emailed a link for the first questionnaire to all students in the 54 departments participating in our experiment, requesting their participation (For the Dean of Students’ Email see Fig. 3 in the supplementary information). The questionnaire was available to students for three weeks.

On week eight, at the midst of the semester (May 10th through 21, 2021) violence broke out between Palestinian citizens of Israel and Jewish citizens. This cycle of violence included street lynching and nightly riots in cities where Palestinians and Jews live side by side. It also included violent riots at the Hebrew University’s main gate, and online violence and harassment on classes WhatsApp groups and
social network groups. In addition, these events were accompanied by daily Israeli air strikes on the Gaza Strip, a severe lockdown of the Palestinians living in the West Bank, as well as daily firing of rockets from the Gaza Strip to cities across Israel.

On the last week of the semester (week 14), five weeks after these events (06.27.2021), we administered another questionnaire by sending personal emails to students who had participated in the first questionnaire and provided us with their email. We also approached new students on campus from each department and asked them to fill out the questionnaire.

Procedure And Participants

Self-Report Questionnaires

In Study 1, out of 834 Palestinian and Jewish students who were registered to the 34 courses that participated in the study, 415 completed the questionnaire at the end of the semester (50%), 60 Palestinians (62%), and 348 Jews (47%), 30 Palestinian students and 203 Jewish students in the experimental condition; 30 Palestinian students and 145 Jewish students in the control condition.

In Study 2, Overall, out of 20,733 Palestinian and Jewish students in the 54 departments assigned to the study, 5949 students completed the questionnaire. We excluded 540 participants who were part of the “Mechina” program (i.e., young adults participating in a preparatory program to be accepted as students into the University) and are not part of the departments. In addition, and following our preregistration exclusion criteria, we also excluded 839 students who studied in two departments, one of which was assigned to the experimental condition and the other to the control condition. After removing these 1379 students, our sample was composed of 4570 students: 449 Palestinian students and 4121 Jewish students (270 Palestinian students and 2328 Jewish students in the experimental condition, and 179 Palestinian students and 1793 Jewish students in the control condition).

For the second questionnaire, we emailed students who had participated in our first questionnaire a link inviting them to participate in a brief survey. To recruit additional participants, we also approached students on campus and asked them to fill out the questionnaire. Overall, 2879 students completed our questionnaire. We once again excluded 40 students that were part of the “Mechina” program and are not part of the departments at the university, and 403 students that studied in two departments that were assigned to different conditions. This left us with a sample of 2436 students: 323 Palestinian students and 2113 Jewish students (192 Palestinian students and 1279 Jewish students in the experimental condition, and 131 Palestinian students and 834 Jewish students in the control condition).

To ensure that the experimental and control conditions did not differ significantly on pre-treatment covariates, we conducted balance checks in both Studies, comparing participants from the two conditions on their levels of ideology, gender, and degree. we found no differences across conditions on these covariates (see Fig. 4 in the supplementary information).
**Student’s Behavior**

In Study 1 we measured 834 students’ behavior during class by coding their participation in five classes: one class prior to the start of the intervention (i.e., baseline), and the four remaining classes of the semester, after the intervention was implemented. We obtained recordings of these classes and had a team of three coders who watched and coded the classes. Coders were trained by the first author for three weeks. During the first training week, the first author met with each coder separately and provided them with a detailed oral and written instructions on how to code each behavioral measure. During the second week of training, each coder coded together with the first author two lectures. After that, during the third week of training, both each coder and the first author coded the same class and compared their results to ensure they are coding consistently. The coders received feedback and assistance throughout the coding process which lasted for four months.

Overall, we coded 165 lessons - five lessons for each of the 33 courses (the recordings of one of the 34 courses that participated in the study was not available), coding each students’ behavior. Because we had pre-treatment and post-treatment time series data for all these outcomes, we averaged across the four post-treatment weeks. Behavioral outcomes included Attendance (coded as 1 if the student attended class, and 0 if they did not) and Camera usage (coded as 1 if students had their camera turned on during class, 0.5 if turned on partly during class, and 0 if turned off).

Given the magnitude of Study 2, and the fact that this study included classes that had more than 50 students, we could not code student's behavior during class. Instead, in the second questionnaire we asked students to self-report whether they had changed the way their name appeared on Zoom to include the Arabic transcript (coded 1 = no; 2 = yes). We also gave them the opportunity to be redirected to a website that would provide them with their name's Arabic transcript so that they could change how their name appeared on Zoom. We coded this measure as 2 in cases where the student followed the link, and 1 in cases that he/she had not.

**Student’s Achievements:**

In Study 1, student’s achievements were measured by obtaining each student’s final grade for that class. In Study 2, we obtained the department’s average student grade, rather than student’s individualized grades.

**Statistics:**

Our estimation strategy included Multilevel Modeling accounting for clusters (classes in Study 1 and departments in Study 2) with intercept random effect \( Y_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 Z_{ij} + g_i + \mu_{ij} \) Where \( Y \) is an outcome, \( i \) denotes participant, and \( j \) denotes class (Study 1) or department (Study 2). The regression coefficient \( \beta_1 \) represents the average causal effect of the treatment. \( T_{ij} \) represents a binary variable of the experimental manipulation randomly assigned to classes (Study 1) or departments (Study 2), in which \( T_{ij} \)
= 1 refers to the experimental condition, and T\textsubscript{ij} = 0 refers to the control condition. \( g_i \) denotes class (Study 1) or department (Study 2) random effects, and \( \mu \) is the error term. \( Z_{ij} \) is a vector of individual level participant characteristics that are unaffected by treatment (see description of these below). Importantly, for self-report outcomes measured via the questionnaire, we also added weights to the model to account for differential survey attrition.

Study 1 covariates include: gender, Palestinian proportion in the class, class size, self-reported Hebrew proficiency (for Palestinian students), and course content (whether the content of the course involved intergroup relations). For the behavioral measures, we also controlled for coder. Study 2 covariates include: gender, Palestinian proportion in the department, and department size. We used an \( \alpha \) level of 0.05 for all statistical tests. All tests’ statistics were two-tailed.

We compared the results of the MLM accounting for clusters to those of LM robust which clusters the standard errors and decided to choose to show in the main manuscript the MLM using lmer function over the LM robust because of its higher ICC in our outcome variables. In the Supplementary information, we show the replication of our results with a linear model while clustering the standard error (LM robust), at the class level in Study 1 and at the department level in Study 2. Because in most of our behavioral measures in study 1 (except for attendance and camera usage), there was limited participation in class by students, especially the among the few Palestinian students who attended these classes, we conducted a logistic regression using the glmer function which account for these zero-inflation negative binomial variables.

**Declarations**

**Data and Availability:** Preregistration, data and replication codes are available in the OSF repository. Study 1: https://osf.io/c3qxv/?view_only=12c3fa0a947d441fa6ceb6c0c0183395; Study 2: https://osf.io/vsqax/?view_only=742c5c5fa0fa44c897e67f9cd3b3be2a.

**ACKNOWLEDGMENTS.** For excellent research assistance we acknowledge Adina Karpuj, Alona Friedman, Tom Benezer, Yishai Pulvermacher and Ran Dayan. We would also like to thank Micha Mandel, Gal Chen and Yaron Mor for providing statistical consult, and Noam Markovich and Gal Chen for reproduction of our codes. This research was funded by the ISF 2240/21 granted to Eran Halperin and Roni Porat.

**Authors contributions:** KE EH and RP designed the study and oversaw data collection, KE and RP analyzed the data and wrote the paper, EH provided critical revisions.

**Competing interests:** The authors declare that they have no competing interests.

**References**


62. Brainer, J. Hason, N. Palestinians and police confronted at the entrance to the Hebrew University; police are expected to approve the flag parade. *Haaretz* (2020, May 9).


**Figures**
Figure 1

Timelines for study 1 (above figure) and 2 (lower figure).
Figure 2

Illustration of the intervention in studies 1 & 2.

![Control condition](image1) ![Experimental condition](image2)

Figure 3

Student’s grades in Studies 1 & 2. Each point estimate and its corresponding 95% CI are extracted from a separate regression in which we identify the effect of condition on students’ grades. In Study 1 ($n_{Palestinians} = 87; n_{Jews} = 737$) we used MLM regression with cluster (class) random effect. In Study 2 ($n_{Palestinians} = 3,380; n_{Jews} = 17,353$) we used OLS regression using the average grade for each department separately for Palestinian and Jewish students. All outcomes are standardized. The results of Palestinian students appear in blue, while the results of the Jewish students appear in red.
Figure 4

Students’ attendance and camera usage ($n_{Palestinians} = 96; n_{Jews} = 719$) one lecture before the start of the intervention (week 10) and for the four lectures that followed (weeks 11 through 14). Each point represents the number of students that either attended (upper figures) or had their camera turned on (bottom figures). Results for Palestinian students appear in blue, while results for Jewish students appear in red.

Study 1 – Behavioral measures

Figure 5
Students' attendance ($n_{Palestinians} = 96; n_{Jews} = 719$) and camera usage ($n_{Palestinians} = 96; n_{Jews} = 719$). We analyzed 33 out of 34 courses due to a technical problem with the recordings of one of the courses. Each point estimate and its corresponding 95% CI are extracted from a separate MLM regression with cluster (class) random effect. The figure depicts the interaction's simple effect looking at the differences between pre and post intervention in attendance and camera usage for each condition separately. All outcomes are standardized. The results of Palestinian students appear in blue, while the results of the Jewish students appear in red.

Study 1

![Study 1](image)

Study 2 – First questionnaire  

Study 2 – Second questionnaire

![Study 2](image)
Effects of the intervention in Studies 1 (top figure) & 2 (bottom figure). Each point estimate and its corresponding 95% CI are extracted from a separate MLM regression with weights and cluster (class/department) random effects in which we identify the effect of condition on student's sense of belonging (Study 1: $n_{Palestinians} = 60; n_{Jews} = 338$; Study 2: $n_{Palestinians} = 369, 298; n_{Jews} = 3568, 2052$). In Study 1 the outcomes were measured once, at the end of the semester. In Study 2, outcomes were measured twice: four weeks after the implementation of the intervention, and at the end of the semester. All outcomes are standardized. The results of Palestinian students appear in blue, while the results of the Jewish students appear in red.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- SupMatFebruary2023.docx
- rs.pdf