



Zoom out: An intervention on the virtual learning environment improves minority students' grades in two field experiments in Israel

Kinneret Endevelt^a, Eran Halperin^a, and Roni Porat^{b,c,1}

Edited by Geoffrey L. Cohen, Stanford University, Stanford, CA; received August 10, 2023; accepted April 2, 2024 by Editorial Board Member Mark Granovetter

Closing the achievement gap for minority students in higher education requires addressing the lack of belonging these students experience. This paper introduces a psychological intervention that strategically targets key elements within the learning environment to foster the success of minority students. The intervention sought to enhance Palestinian minority student's sense of belonging by increasing the presence of their native language. We tested the effectiveness of the intervention in two field experiments in Israel ($n > 20,000$), at the height of the COVID-19 pandemic when all classes were held via Zoom. Lecturers in the experimental condition added a transcript of their names in Arabic to their default display (English/Hebrew only). Our findings revealed a substantial and positive impact on Palestinian student's sense of belonging, class participation, and overall grades. In experiment 1, Palestinian student's average grade increased by 10 points. In experiment 2, there was an average increase of 4 points among Palestinian students' semester grade. Our intervention demonstrates that small institutional changes when carefully crafted can have a significant impact on minority populations. These results have significant implications for addressing educational disparities and fostering inclusive learning environment.

belonging | learning environment | inequality | psychological interventions | field experiments

How can we improve minority achievements in higher education? Social scientists from various disciplines have studied this question for over two decades. Scholars' focus on higher education stems from findings suggesting that completing a bachelor's degree is associated with a host of socioeconomic benefits like one's position in the labor market (1), productivity (2), career opportunities (3), and lifelong earnings (4–7). This implies that successfully obtaining an academic degree can serve as an equalizer that can meaningfully reduce inequality and increase minorities' social mobility (8). However, studies repeatedly demonstrate that disparities in higher education loom large (9, 10). In the United States, there is a 14-percent difference in bachelor degree completion rates between non-Hispanic White and Black people (11). This is also true in Israel, where there is an 18 percent difference in degree completion rates between Jewish and Palestinian citizens (12). 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.

A major obstacle hindering minorities from successfully completing their academic degrees is the uncertainty they experience in their sense of belonging within the educational environment. Social belonging is the subjective feeling of inclusion or acceptance into a group or community (13). This sense of connectedness is a basic human need that individuals must satisfy to maintain their identity, physical well-being, and mental health (14, 15). Uncertainty about belonging has been found to negatively influence higher education outcomes, especially for underrepresented students facing stereotypes about their abilities and worthiness (15, 16). For example, belonging uncertainty has been associated with lower academic achievements and an increased risk of dropout (16–18).

Interventions to increase one's sense of belonging have shown some success (16, 18–20). However, a recent megastudy by Walton and colleagues (2023) found that such interventions are only effective in institutions that afforded minority students an opportunity to foster a sense of belonging: "... colleges may complement brief online modules with broader efforts to create a culture on campus that normalizes challenges and worries about belonging and emphasizes opportunities for growth..." (380, p. 504) (21).

How can institutions design affordances that enable the necessary changes to make minority students feel that they belong? (22–27). To address this challenge, we differentiate between educational–pedagogical changes and environmental changes. Educational–pedagogical changes involve modifications at the institutional level to foster inclusion. For example, these changes may involve revising the curriculum to incorporate diverse perspectives, histories,

Significance

Closing the achievement gap in higher education requires addressing minority students' lingering sense of belonging uncertainty. How can this be achieved on a large scale? This paper presents a psychological intervention that modifies the institutional learning space to convey a sense of belonging to minority students. The intervention was tested in two field experiments conducted in Israel, among more than 20,000 Jewish and Palestinian students. By manipulating the display of the lecturer's name on Zoom in three languages (including Arabic) instead of the default option, the intervention positively impacted Palestinian students' sense of belonging, class participation, and overall grades. These findings carry significant implications for narrowing educational disparities and promoting inclusivity within the university environment.

Author contributions: K.E., E.H., and R.P. designed research; K.E. and R.P. performed research; K.E. and R.P. analyzed data; and K.E., E.H., and R.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. G.L.C. is a guest editor invited by the Editorial Board.

Copyright © 2024 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: Roni.porat@mail.huji.ac.il.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2313496121/-DCSupplemental>.

Published May 21, 2024.

and experiences or creating services tailored to the unique needs of minority students such as peer mentoring programs (28–32). On the other hand, environmental changes create a supportive campus space that nurtures a sense of belonging. For example, universities can create inclusive spaces that reflect and celebrate the diversity of their student population. This includes designing and renovating the physical and virtual university facilities to incorporate elements that represent various cultures, backgrounds, and identities (33–37).

While research has highlighted the significance of educational–pedagogical processes in improving minority achievements (27, 38), there has been scant scholarly attention devoted to assessing the impact of changes in the institution’s learning environment (35, 39). For example, laboratory studies demonstrate that removing stereotyping cues from the environment, such as eliminating stereotypical objects in computer science classrooms (e.g., electronics, video games and Star Wars and Star Trek items), enhances female student’s sense of belonging (35, 36). However, experimental field evidence on the effects of introducing inclusive changes in the learning environment, rather than merely removing exclusionary or stereotypical cues, remains scarce. One exception is a single experiment (study 2) conducted by Murrar et al., (2020), where posters providing information about student’s prodiversity attitudes and behaviors were hung on classroom walls for a 5-wk period. They found that students exposed to these posters perceived the classroom climate to be more inclusive (32). This suggests that introducing inclusive environmental cues can positively impact minority students’ perceptions of inclusion. Nevertheless, it is unclear whether such changes can positively impact minority student’s sense of belonging, participation in class, and their grades. To address this gap, we designed a psychological intervention that introduced inclusive cues in the university’s environment, and particularly in student’s learning space.

Our experiments were conducted in Israel, focusing on the Palestinian minority. While several social groups within Israel suffer from inequality, there is wide consensus that the Palestinian minority is the most marginalized (40). Palestinians in Israel face both social and institutional discrimination, resulting in deprivation in

almost all life domains (41). One salient indicator of their marginalization is the absence of the Arabic language from the Israeli public sphere, including higher education institutions (42, 43). In fact, open-ended interviews we conducted with Palestinian students at the Hebrew University revealed that the absence of the Arabic language from campus enhances their feelings of exclusion. This qualitative insight aligns with scholarly work indicating that language plays a pivotal role in fostering a sense of belonging (43, 44). It serves as a significant social and cultural marker of identity, reflecting the unique history, culture, and values of a social group, and shaping the way individuals perceive themselves and others (45–48).

Thus, we developed an intervention that introduced the Arabic language to the university’s learning environment. We thought that it could communicate to the Palestinian students that the university accepts, values, and embraces their identity and culture, which would enhance their sense of belonging. Therefore, our intervention focuses on introducing the Arabic language in the academic learning environment to facilitate Palestinian’s sense of belonging. We hypothesized that an increased sense of belonging would lead students to be more active learners which may result in better academic achievements (49).

We tested the intervention in two preregistered field experiments at the Hebrew University where about 16% of students are Palestinians. The experiments were carried out during the COVID-19 pandemic, a time where minority students were disproportionately harmed in higher education as they experienced more technical and learning challenges (50), financial insecurity (51), and higher rates of mental health disorders (52). Given that classes at the Hebrew University were held remotely during this period, the institutional change to the learning environment 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 Figs. 1 and 2 for the experiments timeline and a depiction of the intervention).

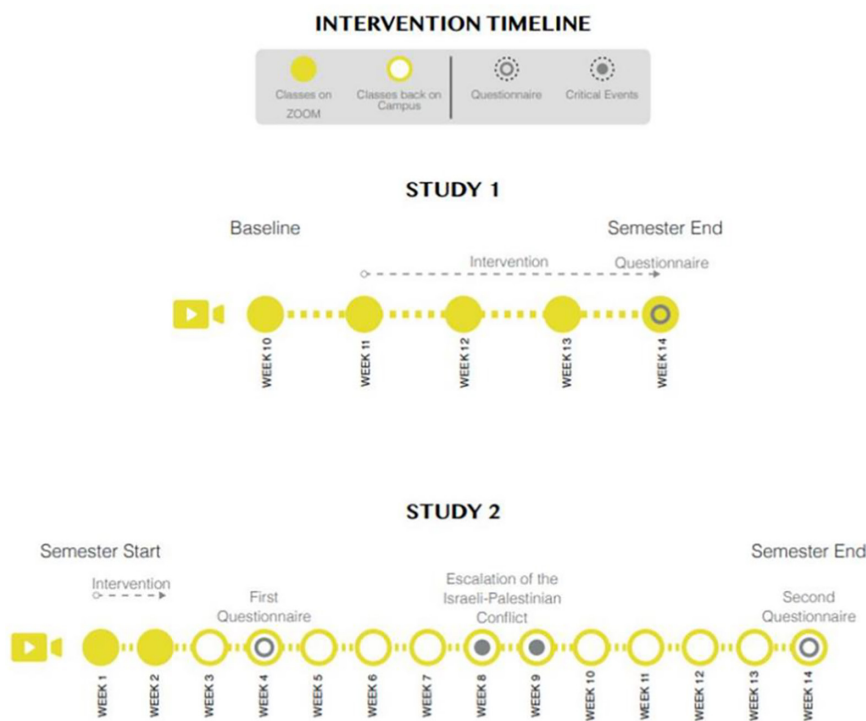
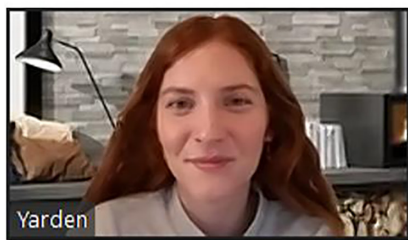


Fig. 1. Timelines for studies 1 (above figure) and 2 (lower figure).

Control condition



Experimental condition

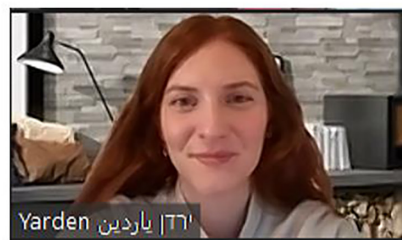


Fig. 2. Illustration of the intervention in studies 1 and 2.

Study 1 was conducted during the spring semester 2020, which was fully remote. We randomly assigned 34 classes ($n = 834$) from two faculties (Social Science and Humanities) into experimental (17 classes) or control (17 classes) condition. Lecturers in the experimental condition were asked to change how their names appeared, 4 wk before the end of the semester. We evaluated student's achievements by examining their final grade for the course and their dropout rates. Furthermore, we assessed their participation by coding attendance and camera usage in the class before and the four classes after the intervention was administered. Finally, we evaluated student's 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 2 wk. We randomly assigned 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 ($n = 27$) condition. 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 later, we administered a questionnaire measuring the same outcomes as in study 1. Importantly, in midsemester (May 10th to 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 (53). 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 (separately for Palestinian and Jews). Details about the recruitment, design, representativeness of the samples, and robustness checks are provided in the *Materials and Methods* section and in *SI Appendix*.

Results

We analyzed the data using linear mixed models with by-subject and by-class/department random intercepts which allowed to account for repeated measurements and differences between clusters (classes in study 1 and departments in study 2). We also controlled for covariates including gender, the proportion of Palestinian students in the class/department, and class/department size (a full list of covariates appears in the methods section). Importantly, whereas certain variables (e.g., achievements, behavior) were obtained through administrative data, enabling the analysis of the entire study sample, other variables (i.e., sense of belonging), were measured through surveys, and as a result, did not encompass the complete study sample. To address potential biases from self-selection into the survey, we employed inverse probability weighting (IPW) (54) (Table 1). To ensure that results were robust to different model specifications, we replicated the

findings using four alternative models and found that results remained consistent throughout these analyses (see *SI Appendix*, pp. 13–26 for more information regarding alternative models and regression outputs).

We begin by reporting findings on student's achievements, the primary outcome for both studies (Fig. 3). We found that the grades of Palestinian students in the experimental condition were higher in both experiments than the grades of those in the control condition. In study 1, we found that Palestinian student's grades in the experimental condition were 10 points* higher than their counterparts in the control condition ($b = 11.29$, 95% CI = [5.42, 17.17], $SE = 2.95$, $t(81) = 3.83$, $P < 0.001$). In study 2, we found a marginally significant effect, where the average departmental grade for Palestinian students in the experimental condition was 4.05 points higher than their counterparts in the control condition ($b = 4.12$, 95% CI = [−0.23, 8.47], $SE = 2.16$, $t(50) = 1.90$, $P = 0.063$). We did not find such differences among the Jewish students in study 1 ($b = 2.40$, 95% CI = [−1.29, 6.09], $SE = 1.88$, $t(731) = 1.28$, $P = 0.202$). We found a marginally significant effect for Jewish students in study 2 such that the average departmental grade for Jewish students in the experimental condition was 1.96 points higher than their counterparts in the control condition ($b = 1.94$, 95% CI = [−0.06, 3.95], $SE = 1.00$, $t(50) = 1.95$, $P = 0.057$).

Given that we found differences for both the 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 or control conditions. To do so, we compared the departmental average grades from the previous year (spring 2020) and did not find a significant difference (Palestinians: $b = 2.94$, 95% CI = [−0.60, 6.47], $SE = 1.76$, $t(50) = 1.67$, $P = 0.101$; Jews: $b = 1.07$, 95% CI = [−0.76, 2.91], $SE = 0.92$, $t(50) = 1.17$, $P = 0.246$), providing further evidence that the grade difference we obtained was the result of the intervention rather than unbalanced assignment of departments into conditions. The intervention appears to have had a substantial impact, closing the achievement gap between Palestinian and Jewish students by 23.9%.

We also assessed in study 1 how many students dropped out at the end of the semester. No significant effects were observed for both Palestinian and Jewish students (Palestinians: $b = -0.06$, 95% CI = [−0.22, 0.09], $SE = 0.08$, $t(87) = -0.80$, $P = 0.424$; Jews: $b = 0.07$, 95% CI = [−0.01, 0.15], $SE = 0.04$, $t(763) = 1.70$, $P = 0.090$). In study 2, due to institutional data access restrictions, we were unable to obtain student's dropout data.

Next, we examined what might explain the differences in grades. We first turn to test whether the intervention affected student's behavior. We measured behavior in study 1 using two measures (given the large number of classes we had in study 2, we could not

*The reported 10-point difference refers to the raw mean difference between conditions.

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

Outcome	Measurement	Data
Grades	Final course grade for each student (study 1) Average departmental grade separately for Palestinian and Jewish students (study 2)	The entire student body within the relevant departments at the Hebrew University
Dropout rates	Dropping out from the degree at the end of the semester (study 1)	The entire student body within the relevant departments at the Hebrew University
Class attendance	Individual student attendance rates in five lessons (one lesson before the administration of the intervention and four lessons after) (study 1).	The entire student body within the relevant departments at the Hebrew University
Camera usage	Individual information regarding camera usage in five lessons (one lesson before the administration of the intervention and four lessons after) (study 1).	The entire student body within the relevant departments at the Hebrew University
Sense of belonging	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")	Sample

code student's behavior in class): First, we coded student's attendance in five lessons (one prior to the intervention, and four after its administration). Second, we coded whether students had their cameras turned on in these five lessons[†]. For the Palestinian students, we interpret the behavioral outcomes cautiously given the relatively low number of participants ($n = 96$) (Fig. 4). For student attendance, we found that in the lesson prior to the intervention, 47.5% of Palestinian students (19 out of 40) in the experimental condition attended, while 50% of Palestinian students (27 students out of 54) in the control condition attended. We next descriptively examine attendance in the four lessons where the intervention was administered to assess whether it had an impact on Palestinian student's attendance. In the control condition, we observed a decline in Palestinian student's attendance over the next four lessons ($n_{\text{class1}} = 21$; $n_{\text{class2}} = 21$; $n_{\text{class3}} = 18$; $n_{\text{class4}} = 22$). However, in the experimental condition, we observed an increase in Palestinian student's attendance that was maintained throughout the four lessons ($n_{\text{class1}} = 28$; $n_{\text{class2}} = 28$; $n_{\text{class3}} = 28$; $n_{\text{class4}} = 27$). For camera usage, we obtain a similar trend. We found that in the lesson before the start of the intervention, nine Palestinian students in the experimental condition (47% of those who attended) and eight Palestinian students in the control condition (30% of those who attended) had their cameras turned on. In the control condition, we observed a decline in camera usage over the next four lessons [$n_{\text{class1}} = 7$ (33%); $n_{\text{class2}} = 6$ (29%); $n_{\text{class3}} = 6$ (33%); $n_{\text{class4}} = 2$ (9%)], whereas in the experimental condition, we observed an increase in camera usage that was maintained to some extent throughout the four lectures [$n_{\text{class1}} = 13$ (46%); $n_{\text{class2}} = 19$ (67%); $n_{\text{class3}} = 11$ (40%); $n_{\text{class4}} = 17$ (63%)].

We performed regression analyses to test whether these differences in attendance and camera usage were statistically significant. Specifically, we examined the interaction between time (pretreatment lesson vs. an aggregate of the four posttreatment lessons) and condition predicting attendance and camera usage among Palestinian and Jewish students. We used linear mixed models with fixed and random effects, including a random intercept for the course variable. This approach effectively accounted for both the repeated measures within students and the clustering of students within courses. We found a significant condition by time interaction for Palestinian students' attendance ($b = 0.33$, 95% CI = [0.18, 0.49], $SE = 0.08$, $t(467) = 4.18$, $P > 0.001$), and camera usage ($b = 0.16$, 95% CI = [0.04, 0.28], $SE = 0.06$, $t(468)$

$= 2.72$, $P = 0.007$). Turning to the simple effects, we found no pretreatment differences between conditions in attendance rates and camera usage (attendance: $b = -0.18$, 95% CI = [-0.39, 0.03], $SE = 0.11$, $t(467) = -1.63$, $P = 0.09$; camera usage: $b = 0.02$, 95% CI = [-0.19, 0.23], $SE = 0.11$, $t(468) = 0.17$, $P = 0.87$). Next, we examined the slope of each condition separately. For the Palestinian students in the control condition, we found a significant decrease in attendance ($b = -0.12$, 95% CI = [-0.23, -0.02], $SE = -0.05$, $t(468) = -2.26$, $P = 0.02$). However, for the Palestinian students in the experiment condition, we found a 21% significant increase following the intervention ($b = 0.21$, 95% CI = [0.10, 0.32], $SE = 0.06$, $t(467) = 3.60$, $P > 0.001$). Similarly, following the intervention, we observed a 9.5% significant increase in camera usage among Palestinian students in the experimental condition ($b = 0.10$, 95% CI = [0.02, 0.19], $SE = 0.04$, $t(468) = 2.39$, $P = 0.02$), while for the Palestinian students in the control condition, no significant difference was found ($b = -0.06$, 95% CI = [-0.13, 0.02], $SE = 0.04$, $t(468) = -1.41$, $P = 0.16$). However, when examining posttreatment differences between conditions, differences in attendance and camera usage were not significant (attendance: $b = 0.15$, 95% CI = [-0.02, .32] $SE = 0.09$, $t(467) = 1.68$, $P = 0.10$; camera usage: $b = 0.18$, 95% CI = [-0.01, 0.37] $SE = 0.10$, $t(468) = 1.86$, $P = 0.09$) (Fig. 5).

For the Jewish students, we found a significant condition by time interaction on attendance ($b = -0.07$, 95% CI = [-0.12, -0.01], $SE = 0.03$, $t(3572) = -2.46$, $P = 0.014$) but not camera usage ($b = 0.01$, 95% CI = [-0.05, 0.06], $SE = 0.03$, $t(3585) = 0.28$, $P = 0.779$). Looking at the simple effects, we did not find pretreatment differences across conditions in attendance ($b = 0.03$, 95% CI = [-0.10, 0.15], $SE = 0.06$, $t(3572) = 0.45$, $P = 0.65$) or camera usage ($b = 0.03$, 95% CI = [-0.11, 0.18], $SE = 0.07$, $t(3585) = 0.46$, $P = 0.65$). However, following the intervention, for Jewish students in the experimental condition, we found an 8.5% significant drop in attendance ($b = -0.08$, 95% CI = [-0.12, -0.04], $SE = 0.02$, $t(3572) = -3.94$, $P > 0.001$), while for the Jewish students in the control condition, no significant difference was found ($b = -0.01$, 95% CI = [-0.05, 0.03], $SE = 0.02$, $t(3572) = -0.55$, $P = 0.58$). This is the only indication that we have for a backlash effect among 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 posttreatment differences between Jewish participants in the control and experimental condition, no differences were found in attendance ($b = -0.05$, 95% CI = [-0.16, 0.08], $SE = 0.06$, $t(3572) = -0.68$, $P = 0.50$) or camera usage ($b = 0.04$, 95% CI = [-0.10, 0.18], $SE = 0.07$, $t(3585) = 0.59$, $P = 0.56$) (Fig. 5).

[†]We also coded for various behavioral variables that captured students' engagement during lessons, such as whether they asked questions, the duration of their speaking time, and whether they posted comments in the chat. However, when analyzing the data, we found that the occurrence of these events was rare, making these outcomes zero-inflated and less revealing. As a result, we report on these findings in *SI Appendix*.

Study 1

Study 2

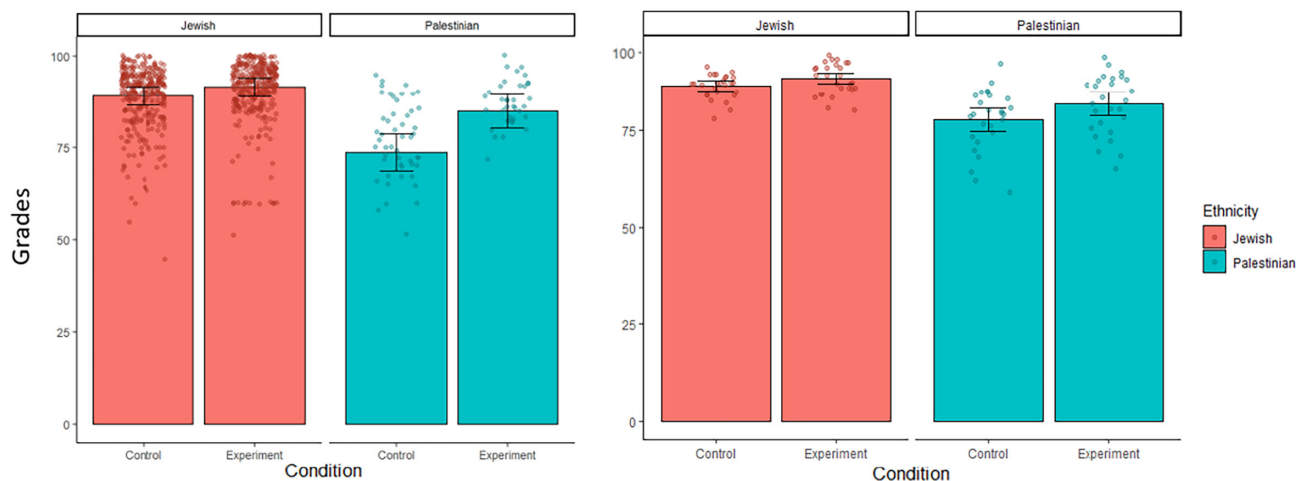


Fig. 3. Student's group mean grades in studies 1 and 2 on a 0- to 100-point scale. Ratings are shown as a function of assignment to condition. Error bars show 95% CI. In study 1 ($n_{\text{Palestinians}} = 87$, $M_{\text{Experiment}} = 87.1$, $SD_{\text{Experiment}} = 6.15$, $M_{\text{Control}} = 77.1$, $SD_{\text{Control}} = 10.2$; $n_{\text{Jews}} = 737$, $M_{\text{Experiment}} = 90.9$, $SD_{\text{Experiment}} = 9.34$, $M_{\text{Control}} = 89.1$, $SD_{\text{Control}} = 8.16$), we used MLM regression with cluster (class) random effect. In study 2 ($n_{\text{Palestinians}} = 3,380$, $M_{\text{Experiment}} = 81.85$, $SD_{\text{Experiment}} = 7.76$; $M_{\text{Control}} = 77.8$, $SD_{\text{Control}} = 7.94$; $n_{\text{Jews}} = 17,353$, $M_{\text{Experiment}} = 88.5$, $SD_{\text{Experiment}} = 4.07$; $M_{\text{Control}} = 86.54$, $SD_{\text{Control}} = 3.14$), we used OLS regression using the average grade for each department separately for Palestinian and Jewish students. The results of Palestinian students appear in blue, while the results of the Jewish students appear in red.

Finally, we turn to the psychological mechanism we hypothesized that would drive these effects, namely, 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: $b = 0.83$, 95% CI = [0.34, 1.33], $SE = 0.25$, $t(49) = 3.37$, $P = 0.001$; study 2: $b = 0.31$, 95% CI = [0.05, 0.56], $SE = 0.13$, $t(362) = 2.38$, $P = 0.018$). However, in study 2, this effect did not persist after the violent events ($b = 0.17$; 95% CI = [-0.06, 0.41], $SE = 0.12$, $t(291) = 1.47$, $P = 0.142$). Finally, Jewish students' sense of belonging was not affected by the intervention in both studies (study 1: $b = -0.16$, 95% CI = [-0.44, 0.12], $SE = 0.14$, $t(330) = -1.12$, $P = 0.265$; study 2: $b = 0.01$, 95% CI = [-0.06, 0.09], $SE = 0.31$, $t(3548) = 0.31$, $P = 0.756$) (Fig. 6).

Discussion

In this paper, we experimentally tested a psychological intervention aimed at improving minority achievements in higher education by changing the learning environment. The intervention was administered in two large-scale preregistered field experiments in Israel. In these experiments, lecturers in the experimental condition changed how their names appeared on Zoom so that instead of the default setting, it appeared in three languages: Arabic, Hebrew, and English. We found that the intervention positively affected Palestinian students' sense of belonging, class participation, and final grades. We do, however, observe one backlash effect for the Jewish students—a decrease in their attendance rates in study 1. This appears to be an isolated finding but still warrants further research.

[‡]In study 2 first questionnaire, we also explored students' sense of belonging to the department to determine whether the enhanced sense of belonging was on the departmental level or the university level. We found no significant differences between the experimental and the control conditions for both Palestinians ($b = 0.14$, 95% CI = [-0.11, 0.44], $SE = 0.14$, $t(366) = 1.19$, $P = 0.235$) and Jews ($b = -0.02$, 95% CI = [-0.16, 0.11], $SE = 0.07$, $t(3556) = -0.34$, $P = 0.734$). This suggests that the rise in Palestinian students' sense of belonging was due to an enhanced sense of belonging in the university rather than the department.

Our findings join others (21, 35) in emphasizing the vital role that institutions can play in supporting the growth and overcoming the challenges minority students face in higher education. Our intervention experimentally tests how changes to the learning environment impact minority's grades and sense of belonging.

Overall, these studies demonstrate the potential of environmental changes (32, 33, 35) (e.g., inclusive learning spaces) as a powerful tool to complement educational–pedagogical interventions (30, 31, 55, 56) (e.g., curriculum revisions, mentoring programs) and create a more holistic and supportive environment for minority students. By incorporating both types of interventions, institutions can work toward reducing disparities in higher education and promote social mobility, ultimately contributing to a more equitable and inclusive society.

We perceive the alteration in the lecturer's name display on Zoom as one instance of a broader concept—actively reshaping the institutional environment to allocate space and attention to minority identities. These changes can go beyond the virtual environment to the physical one. For example, minority language can be embedded in name signs of faculty offices, cafeteria menus, and student lounges. Beyond language, the institution could incorporate minority's identity components in the learning environment, like highlighting minority faculty, culture, and tradition. These efforts send a compelling message of belonging to all who enter the institution's doors. This proactive approach has the potential to actively cultivate a sense of belonging, and foster academic success. In addition, these findings can also be applied to other contexts and spaces beyond higher education like workplaces, healthcare institutions or public spaces such as malls. By subtly integrating minority identities into everyday spaces, we can build a more equitable and welcoming environment.

The findings of these studies also contribute to the literature on organizational authenticity and self-expression (57–59), suggesting that acknowledging and validating minority identities can significantly boost inclusivity. This resonates with broader scholarship highlighting the positive effects of linguistic representation for marginalized groups (44–48, 60), revealing the potential of even subtle interventions to contribute to a more equitable and welcoming learning environment for diverse student populations.

Study 1 – Behavioral measures

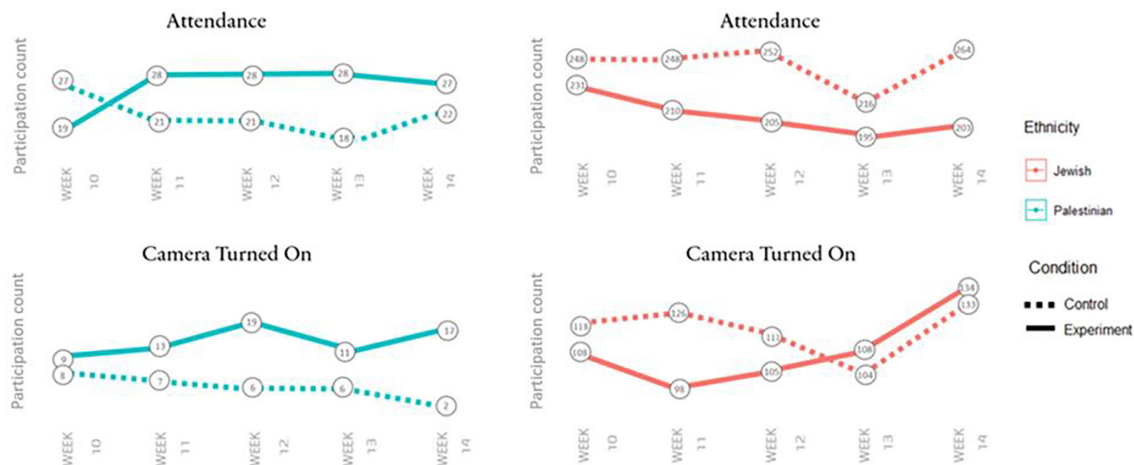


Fig. 4. Students' attendance and camera usage ($n_{\text{Palestinians}} = 96$; $n_{\text{Jews}} = 719$) one lecture before the start of the intervention (week 10) and for the four lectures that followed (weeks 11 through 14). Each point represents the number of students that either attended (*Upper* figures) or had their cameras turned on (*Bottom* figures). Results for Palestinian students appear in blue, while results for Jewish students appear in red.

It is important to acknowledge several limitations in our research that call for more scholarly work. First, while our studies demonstrated promising findings, it is important to acknowledge that there were inconsistencies across outcome measures and studies. With regard to the former, while we do find an increase in GPA for the Palestinian students, we do not find any effect on dropout rates. With regard to the latter, the positive effect on Palestinian grades is less pronounced in study 2. Specifically, the simple effect of condition within Palestinian students is marginal. This may be because study 2 was significantly larger, and there is a high likelihood that included students who were not exposed to the intervention at all. Furthermore, for Palestinian's sense of belonging, while we observe a positive effect in the short term, it did not persist. This may be because the intervention's effect waned over time as students adjusted to the name change. Alternatively,

this decay might be due to the violent events which may have overshadowed the intervention's influence. This highlights a crucial point: These interventions may be susceptible to disruption by external events, and as such, their effects might not be long-lasting without reinforcing mechanisms. Future research should disentangle these possibilities and explore how to design interventions that foster recursive cycles. By acknowledging this, we can move away from viewing such interventions as standalone solutions and recognize the need for a system that continuously strengthens the positive effects over time.

Second, the focus on remote learning during the COVID-19 pandemic may not fully capture the intervention's effectiveness in in-person settings. Future research should explore how the intervention performs in different educational conditions. Third, we acknowledge the possibility that the intervention may have indirectly

Study 1 – Behavioral measures

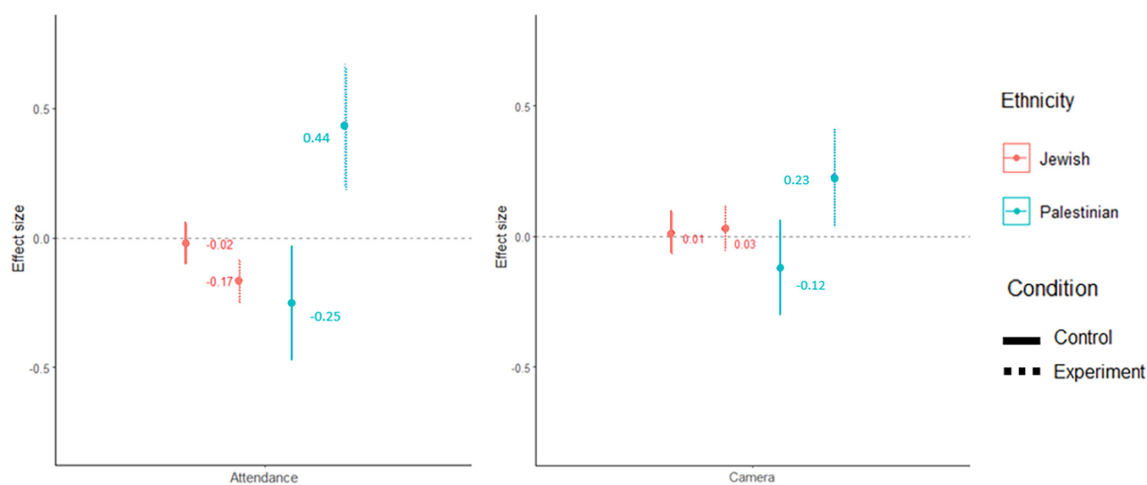
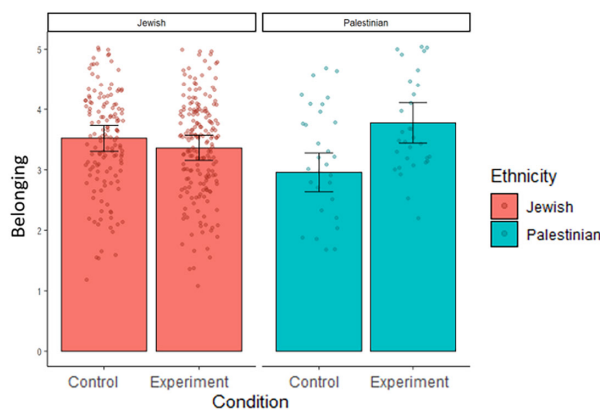
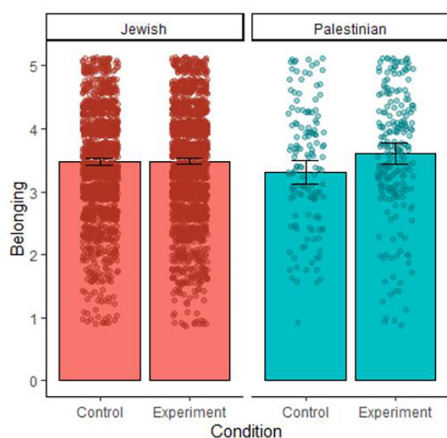


Fig. 5. Students' attendance ($n_{\text{Palestinians}} = 96$; $n_{\text{Jews}} = 719$) and camera usage ($n_{\text{Palestinians}} = 96$; $n_{\text{Jews}} = 719$). We analyzed 33 out of 34 courses due to a technical problem with the recordings of one of the courses. Each point estimate and its corresponding 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 postintervention 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 2 - First questionnaire



Study 2 - Second questionnaire

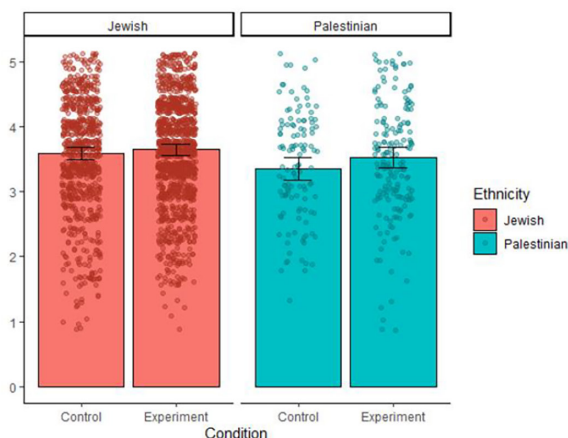


Fig. 6. Student's group mean rating of sense of belonging in studies 1 (Top figure) and 2 (Bottom figure) on a 1 to 5 scale. Ratings are shown as a function of assignment to conditions. Error bars show 95% CI. We used MLM regression with weights and cluster (class/department) random effects in which we identify the effect of condition on student's sense of belonging in study 1 ($n_{\text{Palestinians}} = 60$, $M_{\text{Experiment}} = 3.71$, $SD_{\text{Experiment}} = 0.78$; $M_{\text{Control}} = 3.19$, $SD_{\text{Control}} = 0.93$; $n_{\text{Jews}} = 338$, $M_{\text{Experiment}} = 3.39$, $SD_{\text{Experiment}} = 0.8$; $M_{\text{Control}} = 3.5$, $SD_{\text{Control}} = 0.82$) and in study 2 (first questionnaire: $n_{\text{Palestinians}} = 369$, $M_{\text{Experiment}} = 3.6$, $SD_{\text{Experiment}} = 0.96$; $M_{\text{Control}} = 3.42$, $SD_{\text{Control}} = 0.95$; $n_{\text{Jews}} = 3,568$, $M_{\text{Experiment}} = 3.47$, $SD_{\text{Experiment}} = 0.88$; $M_{\text{Control}} = 3.48$, $SD_{\text{Control}} = 0.87$; second questionnaire: $n_{\text{Palestinians}} = 298$, $M_{\text{Experiment}} = 3.49$, $SD_{\text{Experiment}} = 0.88$; $M_{\text{Control}} = 3.41$, $SD_{\text{Control}} = 0.83$; $n_{\text{Jews}} = 2052$, $M_{\text{Experiment}} = 3.64$, $SD_{\text{Experiment}} = 0.8$; $M_{\text{Control}} = 3.59$, $SD_{\text{Control}} = 0.85$). In study 1, the outcomes were measured once at the end of the semester. In study 2, outcomes were measured twice: 4 wk after the implementation of the intervention and at the end of the semester. The results of Palestinian students appear in blue, while the results of the Jewish students appear in red.

affected Palestinian students by influencing the lecturers. It is possible that our intervention created an indirect process that raised the lecturer's awareness to the hardships their Palestinian v face, leading them to behave more inclusively. While our work does not address this mechanism, it is important for future work to consider this path of influence as well.

Taken together, our research contributes to the literature on reducing disparities in higher education for minority students, highlighting the potential of environmental changes as a powerful tool to foster belonging and improve academic achievements. By cultivating changes in the environment, institutions can create a more equitable and supportive campus for all students, working toward reducing inequality and enhancing social mobility.

Materials and Methods

The research was approved by the Hebrew University's institutional review board committee (#103120). In study 1, informed consent was obtained from all participants prior to their participation. At the beginning of the premeasurement lecture

(1 wk prior to the start of the intervention), lecturers informed the class that they would be participating in a study. Students were given the option to opt out of the study. In study 2, we obtained approval from the ethics committee to proceed with the study without obtaining informed consent directly from participants given its nature and magnitude. However, participants did sign informed consent upon starting the survey questionnaires.

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 improve precision and reduce the chance of imbalance assignment to conditions, we blocked classes by size (small classes of up to 15 students, medium classes of 16 to 30 students, and large classes of 31 to 50 students), and by the number of Palestinian students in each class (0, 1 to 3, 4 to 8, 11 to 22). In addition, 14 classes contained students who were enrolled to more than one class. To avoid a situation whereby students would be assigned to treatment and control conditions, we clustered these classes into five groups (each containing between 2 to 3 classes) and randomly assigned those clusters while taking into consideration the class size and the number of Palestinian students.

In study 2, we randomly assigned 54 departments, from 12 faculties at the Hebrew Universities' six campuses ("Har Hatzofim," "Givat ram," "Rehovot," "Asaf Harofe," "Kaplan," and "Ein Carem"), to the experimental ($n = 27$) and the control ($n = 27$) condition. All departments at the Hebrew University with Palestinian students were included in the study (54 departments). Assignment was done on the department level, making this a clustered design. We blocked departments by department size (small departments from up to 200 students, medium departments from 200 to 400 students and large departments from 400 students) to improve precision of our estimates.

Timeline of Studies 1 and 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 to 50 students) and asked them to participate in our experiment, letting them know that the experiment was approved by the University's ethics committee (see *SI Appendix, Fig. S1* for the full timeline of both studies). Lecturers were not aware of the experiment's hypotheses. Given that Zoom only shows up to 50 faces, we restricted the experiment to classes that had less than 50 students so that we could code student and lecturer's 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. 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 names 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 names appeared, and we provided them with the Arabic transcript of their names. 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 names 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. One week after the end of the semester, we asked the lecturers to remind their students to fill out the questionnaire (by posting a message in the course website). In addition, given the small number of Palestinian students who participated in the experiment ($n = 96$), 3 wk after the semester had ended (07.21.2020), we sent Palestinian students (in both experimental and control groups) the link directly with a reminder and a request in Arabic to participate.

Timeline—study 2. Study 2 was conducted during spring semester 2021 that began on 03.14.2021 1 wk after the end of Israel's third COVID-19 lockdown. However, given the strict gathering restrictions, the University announced that the first 2 wk 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 Rector of the university 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 *SI Appendix, Fig. S1*). Lecturers were not aware of the study's hypotheses. 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 names. 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 names as requested (see more details on the confirmed implementors on *SI Appendix*, pp. 34, 35).

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 2-wk 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 to the first questionnaire to all students in the 54

departments participating in the experiment, requesting their participation (For the Dean of Students' Email, see *SI Appendix, Fig. S3*). The questionnaire was available to students for 3 wk.

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), 5 wk 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 the 834 Palestinian and Jewish students registered in the courses participating in the study, 409 completed the questionnaire (accounting for 49% of the sample). Of these, 60 were Palestinian students (accounting for 62% of the Palestinian student sample), and 349 were Jewish students (accounting for 47% of the Jewish student sample). In terms of conditions, in the experimental condition, 30 Palestinian students and 204 Jewish students filled out the questionnaire. In the control condition, 30 Palestinian students and 145 Jewish students filled out the questionnaires.

In study 2, out of 20,733 Palestinian and Jewish students in the 54 departments assigned to the study, 5,949 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 1,379 students, our sample was composed of 4,570 students: 449 Palestinian students and 4,121 Jewish students (270 Palestinian students and 2,328 Jewish students in the experimental condition, and 179 Palestinian students and 1,793 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, 2,879 students completed our questionnaire. We once again excluded 40 students who were part of the Mechina program and are not part of the departments at the university and 403 students who studied in two departments that were assigned to different conditions. This left us with a sample of 2,436 students: 323 Palestinian students and 2,113 Jewish students (192 Palestinian students and 1,279 Jewish students in the experimental condition and 131 Palestinian students and 834 Jewish students in the control condition). For attrition analyses for both studies, see *SI Appendix*, pages 36–38.

To ensure that the experimental and control conditions did not differ significantly on pretreatment 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 (*SI Appendix, Fig. S4*).

Student's behavior. In study 1, we measured 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 3 wk. During the first training week, the first author met with each coder separately and provided them with 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, each coder and the first author coded the same class and compared their results to ensure that they are coding consistently. The coders received feedback and assistance throughout the coding process which lasted for 4 mo.

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 were not

available), coding each students' behavior. Because we had pretreatment and posttreatment time series data for all these outcomes, we averaged across the four posttreatment weeks. Behavioral outcomes included attendance (coded as 1 if the student attended class and 0 if they did not), camera usage (coded as 1 if students had their cameras turned on during class, 0.5 if turned on partly during class, and 0 if turned off), and other engagement variables (*SI Appendix*).

Student's achievements. In study 1, student's achievements were assessed by obtaining each student's final grade for that class from the University's Student Administration Division. In study 2, we had originally requested individual student semester grades from the University. However, to maintain student's privacy, the University granted us access only to the aggregated departmental average grades for Palestinian and Jewish students separately.

Statistics.

Estimation strategy. Our estimation approach employed linear mixed effects regression (LMER) which is a form of multilevel modeling (MLM). The experimental condition served as a cluster-level variable, with random assignment to experimental or control conditions between clusters (classes in study 1 and departments in study 2). We detail the different model specifications employed for each outcome below.

Model 1—Achievement outcomes. In this multilevel model, clusters were considered with an intercept random effect ($Y_{ij} = \beta_0 + \beta_1T_{ij} + \beta_2Z_{ij} + g_i + \mu_{ij}$). Here, Y represents the outcome, i denotes participant, and j denotes class (study 1) or department (study 2). The regression coefficient β_1 signifies the average causal effect of the treatment. T_{ij} is a binary variable indicating the experimental manipulation, with $T_{ij} = 1$ denoting the experimental condition and $T_{ij} = 0$ denoting the control condition. β_2 , represents the effect of individual-level participant characteristics covariates (Z_{ij}) on the outcome [i.e., gender, Palestinian proportion in the class, class size, self-reported Hebrew proficiency (for Palestinian students), and course content (whether the content involved intergroup relations)] unaffected by treatment. g_i denotes random effects. For the model predicting students' grades in study 2, with grades aggregated at the department level, the analysis was simplified to a single level without departmental clustering.

Model 2—Participation behavioral outcomes. The MLM incorporated clusters with an intercept random effect ($Y_{ij} = \beta_0 + \beta_1T_{ij} + \beta_2Z_{ij} + \beta_3X_{ij} + g_i + \mu_{ij}$). Similar to the previous model, T_{ij} represents the experimental manipulation, with an additional variable X_{ij} indicating the time interaction ($X_{ij} = 1$ for the average of four lessons postintervention and $X_{ij} = 0$ for one lesson preintervention).

Model 3—Sense of belonging outcomes. In this multilevel model, clusters were considered with an intercept random effect ($Y_{ij} = \beta_0 + \beta_1T_{ij} + \beta_2Z_{ij} + \beta_3W_{ij} + g_i + \mu_{ij}$). This model is the same as model 1 with one modification: W_{ij} accounts for survey attrition by assigning different weights to participants based on their likelihood of attrition (Following Gerber & Green, 2012).

Covariates. Study 1 covariates include gender, the proportion of Palestinian students in the class, class size, self-reported Hebrew proficiency (for Palestinian students), and course content (whether the content involved intergroup relations). For behavioral measures, we also controlled for coder. Study 2 covariates include gender, Palestinian proportion in the department, and department size. A significance level of 0.05 was used for all tests, and tests were two tailed.

We conducted robustness checks by comparing these results to a similar model using LM robust that clusters the SE either at the level of the class (study 1) or the department (study 2) (for the full analysis, please see our *SI Appendix* section). Results are unchanged in both models. In addition, for the behavioral outcomes in study 1 (attendance and camera usage), we conducted a logistic regression using the `glmer` function to account for zero inflation.

Descriptive statistics. Study 1: Based on survey responses, our sample included 409 participants (138 men, 269 women, and 2 others), of which 200 were undergraduate students, and 199 were master students. In terms of religion, 54 were Muslim, 9 were Christians, 329 were Jewish, and 4 responded "other". We collected political ideology data only for the Jewish students, as the right-left continuum does not apply for the Palestinian population, especially given that some of our Palestinian students are not citizens of Israel. Our sample included 171 Jewish students who reported they were left leaning, 66 Jewish students who reported they were center leaning, and 44 Jewish students who reported they were right leaning.

Study 2: Based on the first survey responses, our sample included 4,570 participants (1,649 men, 2,906 women, and 15 others), of which 2,998 were undergraduate students, 1,169 were master students, and 398 were graduate students. In terms of religion, 363 were Muslim, 67 were Christians, 3,877 were

Jewish, and 262 responded other. Our Jewish sample included 1,441 students who reported they were left leaning, 841 students who reported they were center leaning, and 1,405 students who reported they were right leaning. We did not measure political ideology among the Palestinian students.

Based on the second survey responses, our sample included 2,436 participants (882 men, 1,550 women, and 4 others), of which 1,628 were undergraduate students, 585 were master students, and 223 were graduate students. In terms of religion, 267 were Muslim, 40 were Christians, 1,987 were Jewish, and 132 responded other. Our Jewish sample included 773 students who reported they were left leaning, 441 students who reported they were center leaning, and 724 students who reported they were right leaning. We did not measure political ideology among the Palestinian students.

Deviations from Preregistration. The preregistration for study 1 was completed on July 2nd, 2020⁵, while study 2's preregistration was finalized on April 5th, 2021⁶. Notably, we deviated from the preregistered statistical approach which specified the use of Linear Regression with robust SE (`lm_robust`) and treated the experimental condition as an individual-level variable. Instead, we used MLM. This shift was prompted by the evident intraclass correlation observed in our data and the nature of our research question and hypotheses. The substantial variance at the cluster level (e.g., classes in study 1, departments in study 2) made MLM more appropriate (63). The adoption of MLM allows us to better capture and account for the hierarchical structure of our data, ensuring a more robust analysis that aligns with the nuances of our study design. Furthermore, for the belonging measures in both studies, we used IPW (54) to adjust for self-selection into the survey (note that in study 2, we preregistered the use of IPW). However, in *SI Appendix*, we conduct the preregistered linear regression with robust SE analyses and show that the pattern of finding is the same (see pages 13–19).

Another deviation from preregistration was with regard to our statistical model. Our analysis focused on planned comparisons within each ethnic group (Palestinian and Jewish) to assess the intervention's effect. This is a deviation from our preregistration that specified a condition \times ethnicity interaction model. We decided to deviate from our preregistered model and analyze the data separately for Palestinian and Jewish students for two reasons. First, when analyzing the data, we realized that there were distinct baseline differences between the Palestinian and Jewish samples with regard to both sociodemographic characteristics, as well as characteristics directly related to the outcome variables. For example, Palestinian students tend to be significantly younger, poorer, and more religious. In addition, we found differences in GPA scores between Palestinian and Jewish students in the year previous to the intervention. Analyzing combined data could mask these differences and hinder understanding of the intervention's effect on each group. Second, survey attrition rates in each classroom (study 1)/department (study 2) were different for Jewish and Palestinian students. This necessitated weighting the data to account for varying participation likelihoods within each cluster for each ethnic group. Since weights differed by ethnicity and by cluster (classrooms/departments), combining the groups would have produced biased results. Separate analyses with group-specific weights allowed for a more accurate assessment of the intervention's impact on each ethnic group. This aligns with Rosenthal and Rosnow's (1995; 2009) (64, 65) advocacy for planned contrasts, especially in cases where group variances differ. Planned contrasts allow for targeted analyses of specific hypotheses, potentially enhancing statistical power and yielding more interpretable results for each population. However, we also conducted the preregistered interaction model and obtained similar results (for a full report of the findings see *SI Appendix*, pp. 27–31).

Results suggest that while the overall pattern of results remained consistent, one notable deviation surfaced in study 2 concerning the final grade outcome. Specifically, we identified a main effect for ethnicity, revealing that Jewish students attained higher grades than their Palestinian counterparts. However, no main effect for the intervention condition or the interaction term emerged. We attribute this inconsistency to the limited sample size (54 of departmental average grades for both Palestinian and Jewish students), potentially impeding the detection of an interactive effect. Additionally, although the time \times condition \times ethnicity interaction predicting students' camera usage was not statistically significant, for Palestinian students, the simple slope of condition from pre- to postintervention and the simple slope of time (postintervention) between the control and experimental conditions were marginally significant.

⁵https://osf.io/c3qkv/?view_only=12c3fa0a947d441fa6ceb6c0c0183395 (61).

⁶https://osf.io/vsqax/?view_only=742c5c5fa0fa44c897e67f9cd3b3be2a (62).

Concerning our preregistered measures, certain measures (perception of norms, prescriptive norms, descriptive norms, and feeling thermometer) were excluded from the analysis as they were unrelated to the core theme of the paper, addressing issues such as intergroup relations and prejudice. These excluded measures will be reported in a separate manuscript.

Data, Materials, and Software Availability. Preregistration, data, and replication codes are available in the OSF repository. Study 1: https://osf.io/c3qxv/?view_only=12c3fa0a947d441fa6ceb6c0c0183395 (61); Study 2: https://osf.io/vsqax/?view_only=742c5c5fa0fa44c897e67f9cd3b3be2a (62).

1. F. T. Pfeffer, F. R. Hertel, How has educational expansion shaped social mobility trends in the United States? *Soc. Forces* **94**, 143–180 (2015).
2. S. Baum, J. Ma, K. Payea, "Education pays 2013: The benefits of higher education for individuals and society" (The College Board, New York, NY, 2013).
3. J. Ma, M. Pender, M. Welch, "Education pays 2016: The benefits of higher education for individuals and society" (Trends in Higher Education Series, College Board, 2016).
4. P. Oreopoulos, U. Petronijevic, Making college worth it: A review of research on the returns to higher education. *The Fut. Child*. **23**, 41–65 (2015).
5. J. Angrist, D. Autor, A. Pallais, Marginal effects of merit aid for low-income students. *Q. J. Eco.* **137**, 1039–1090 (2022).
6. J. J. Heckman, J. E. Humphries, G. Veramendi, Returns to education: The causal effects of education on earnings, health and smoking. *J. Polit. Econ.* **126**, S197–S246 (2016).
7. L. J. Turner, The returns to higher education for marginal students: Evidence from Colorado Welfare recipients. *Econ. Educ. Rev.* **51**, 169–184 (2016).
8. F. Bernardi, G. Ballarino, "Education as the great equalizer: A theoretical framework" in *Education, Occupation and Social Origin*, F. Bernardi, G. Ballarino, Eds. (Edward Elgar, Cheltenham, 2016), pp. 1–19.
9. M. J. Bailey, S. M. Dynarski, "Inequality in postsecondary education" in *Whither Opportunity*, G. J. Duncan, R. J. Murnane, Eds. (Russell Sage, 2011), pp. 117–132.
10. A. M. Garibaldi, The expanding gender and racial gap in American higher education. *J. Negro Educ.* **83**, 371–384 (2014).
11. F. Alonzo, "Census Bureau releases new educational attainment data" (United States Census Bureau, 2022), 24 February.
12. D. Zaken, More Israeli Arabs earning degrees, but inequalities still wide. *Globes*, 17 March 2022. <https://en.globes.co.il/en/article-more-israeli-arabs-earning-degrees-but-inequalities-still-wide-1001406055>. Accessed 11 January 2020.
13. K.-A. Allen, M. L. Kern, C. S. Rozek, D. M. McInerney, G. M. Slavich, Belonging: A review of conceptual issues, an integrative framework, and directions for future research. *Aust. J. Psychol.* **73**, 87–102 (2021).
14. R. F. Baumeister, M. R. Leary, The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* **117**, 497–529 (1995).
15. C. M. Steele, "The psychology of self-affirmation: Sustaining the integrity of the self" in *Advances in Experimental Social Psychology*, L. Berkowitz, Ed. (Academic Press, 1988), vol. **21**, pp. 261–302.
16. G. M. Walton, G. L. Cohen, A question of belonging: Race, social fit, and achievement. *J. Pers. Soc. Psychol.* **92**, 82–96 (2007).
17. D. S. Yeager *et al.*, Teaching a lay theory before college narrows achievement gaps at scale. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3341–E3348 (2016).
18. M. C. Murphy *et al.*, A customized belonging intervention improves retention of socially disadvantaged students at a broad-access university. *Sci. Adv.* **6**, eaba4677 (2020).
19. G. M. Walton, G. L. Cohen, A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* **331**, 1447–1451 (2011).
20. G. M. Walton, A. J. Crum, *Handbook of Wise Interventions* (Guilford Publications, 2020).
21. G. M. Walton *et al.*, Where and with whom does a brief social-belonging intervention promote progress in college? *Science* **380**, 499–505 (2023).
22. C. H. Dupree, C. M. Boykin, Racial inequality in academia: Systemic origins, modern challenges, and policy recommendations. *Policy Insights Behav. Brain Sci.* **8**, 11–18 (2021).
23. C. H. Dupree, M. W. Kraus, Psychological science is not race neutral. *Perspect. Psychol. Sci.* **17**, 270–275 (2022).
24. E. L. Paluck, R. Porat, C. S. Clark, D. P. Green, Prejudice reduction: Progress and challenges. *Annu. Rev. Psychol.* **72**, 533–560 (2021).
25. M. W. Kraus, B. Torrez, A psychology of power that is embedded in societal structures. *Curr. Opin. Psychol.* **33**, 86–90 (2020).
26. P. S. Salter, G. Adams, M. J. Perez, Racism in the structure of everyday worlds: A cultural-psychological perspective. *Curr. Direct. Psychol. Sci.* **27**, 150–155 (2018).
27. D. L. Gray, E. C. Hope, J. S. Matthews, Black and belonging at school: A case for interpersonal, instructional, and institutional opportunity structures. *Educ. Psychol.* **53**, 97–113 (2018).
28. L. Thomas, "Developing inclusive learning to improve the engagement, belonging, retention, and success of students from diverse groups" in *Widening Higher Education Participation*, M. Shah, A. Bennett, E. Southgate, Eds. (Chandos Publishing, 2016), pp. 135–159, 10.1016/B978-0-08-100213-1.00009-3.
29. T. Strayhorn, *College Students' Sense of Belonging: A Key to Educational Success for All Students* (Routledge, 2018).
30. M. R. Maimon, K. Howansky, D. T. Sanchez, Fostering inclusivity: Exploring the impact of identity safety cues and instructor gender on students' impressions and belonging. *Teach. Psychol.* **50**, 105–111 (2023).
31. T. C. Dennehy, N. Dasgupta, Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5964–5969 (2017).
32. S. Murrar, M. R. Campbell, M. Brauer, Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap. *Nat. Hum. Behav.* **4**, 889–897 (2020).
33. B. R. Meagher, Ecologizing social psychology: The physical environment as a necessary constituent of social processes. *Personal. Soc. Psychol. Rev.* **24**, 3–23 (2020).
34. R. Bronkema, N. A. Bowman, A residential Paradox?: Residence hall attributes and college student outcomes. *J. Coll. Stud. Dev.* **58**, 624–630 (2017).
35. S. Cheryan, A. N. Meltzoff, S. Kim, Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Comput. Educ.* **57**, 1825–1835 (2011).
36. S. Cheryan, V. C. Plaut, P. G. Davies, C. M. Steele, Ambient belonging: How stereotypical cues impact gender participation in computer science. *J. Pers. Soc. Psychol.* **97**, 1045–1060 (2009).
37. M. C. Murphy, K. M. Kroeper, E. M. Ozier, Prejudiced places: How contexts shape inequality and how policy can change them. *Policy Insights Behav. Brain Sci.* **5**, 66–74 (2018).
38. M. A. Armstrong, Small world: Crafting an inclusive classroom (no matter what you teach). *Thought Action Fall*, 51–61 (2011).
39. K. E. Chaney, D. T. Sanchez, Gender-inclusive bathrooms signal fairness across identity dimensions. *Soc. Psychol. Personal. Sci.* **9**, 245–253 (2018).
40. A. Bashor, A. Yagur, "Gaps between Jews and Arabs in well-being indicators in Israel" (The Central Bureau of Statistics, 2020).
41. T. Levy-Daphny, Arab Society in Israel. <https://www.vanleer.org.il/en/publications/arab-society-in-israel/>. Accessed 1 November 2020.
42. M. Pinto, The impact of the basic law: Israel as the Nation State of the Jewish people on the status of the Arabic language in Israel. *Minnesota J. Int. Law* **30**, 1–42 (2021).
43. O. Mizel, 'I lost my identity in the halls of academia': Arab students on the use of Arabic in Israeli higher education. *Issues Educ. Res.* **31**, 930–951 (2021).
44. M. N. Craith, *Narratives of Place, Belonging and Language: An Intercultural Perspective* (Springer, 2012).
45. J. Edwards, *Language and Identity: An Introduction* (Cambridge University Press, 2009).
46. K. D. Kinzler, Language as a social cue. *Annu. Rev. Psychol.* **72**, 241–264 (2021).
47. D. Gorter, H. F. Marten, L. Van Mensel, "Linguistic landscapes and minority languages" in *The Palgrave Handbook of Minority Languages and Communities*, G. Hogan-Brun, B. O'Rourke, Eds. (Palgrave Macmillan, London, UK, 2019), pp. 481–506, 10.1057/978-1-137-54066-9_19.
48. V. Banham, "Language: An important social and cultural marker of identity" in *Paper Presented at the Annual Meeting of the Language as a Social Justice Issue* (Edith Cowan University, Perth, WA, 2014).
49. T. L. Strayhorn, *A Key to Educational Success for All Students* (Routledge, New York, ed. 2, 2018).
50. N. Haddad Haj-Yahia, "In the academia: Covid-19 intensified the gaps between Palestinian and Jewish students" (The Israel Democracy Institute, 2020), 18 October.
51. P. H. Barber *et al.*, Disparities in remote learning faced by first-generation and underrepresented minority students during COVID-19: Insights and opportunities from a remote research experience. *J. Microbiol. Biol. Educ.* **22**, 22.1.54 (2021), 10.1128/jmbe.v22i1.2457.
52. R. C. O'Connor *et al.*, Mental health and well-being during the COVID-19 pandemic: Longitudinal analyses of adults in the UK COVID-19 mental health & wellbeing study. *Br. J. Psychiatry* **218**, 326–333 (2021).
53. J. Brainer, N. Hason, Palestinians and police confronted at the entrance to the Hebrew University: Police are expected to approve the flag parade. *Haaretz*, 9 May 2020. <https://www.haaretz.co.il/news/law/2021-05-09/ty-article/0000017f-e038-d3ff-a7ff-f1b803330000>. Accessed 7 July 2022.
54. R. Gomila, C. S. Clark, Missing data in experiments: Challenges and solutions. *Psychol. Methods* **27**, 143–155 (2020), 10.1037/met0000361.
55. T. Liu, Y. Chen, M. Hamilton, K. Harris, Peer mentoring to enhance graduate students' sense of belonging and academic success. *Kinesiol. Rev.* **11**, 285–296 (2022).
56. D. J. Wu, K. C. Thiem, N. Dasgupta, Female peer mentors early in college have lasting positive impacts on female engineering students that persist beyond graduation. *Nat. Commun.* **13**, 6837 (2022).
57. J. E. Dutton, L. M. Roberts, J. Bednar, Pathways for positive identity construction at work: Four types of positive identity and the building of social resources. *Acad. Manage. Rev.* **35**, 265–293 (2010).
58. S. E. Cha *et al.*, Being your true self at work: Integrating the fragmented research on authenticity in organizations. *Acad. Manage. Ann.* **13**, 633–671 (2019).
59. L. M. Roberts, J. E. Dutton, *Exploring Positive Identities and Organizations: Building a Theoretical and Research Foundation* (Psychology Press, 2009).
60. T. J. Yosso, Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* **8**, 69–91 (2005).
61. K. Endevelt, E. Halperin, R. Porat, Zoom out: An intervention on the virtual learning environment improves minority students' grades in two field experiments in Israel – Study 1. Open Science Framework. https://osf.io/c3qxv/?view_only=12c3fa0a947d441fa6ceb6c0c0183395. Deposited 20 June 2023.
62. K. Endevelt, E. Halperin, R. Porat, Zoom out: An intervention on the virtual learning environment improves minority students' grades in two field experiments in Israel – Study 2. Open Science Framework. https://osf.io/vsqax/?view_only=742c5c5fa0fa44c897e67f9cd3b3be2a. Deposited 20 June 2023.
63. D. McNeish, A practical guide to selecting and blending approaches for clustered data: Clustered errors, multilevel models, and fixed-effect models. *Psychol. Methods*, 10.1037/met0000620 (2023).
64. R. L. Rosnow, R. Rosenthal, "Some things you learn aren't so": Cohen's Paradox, Asch's Paradigm, and the interpretation of interaction. *Psychol. Sci.* **6**, 3–9 (1995).
65. R. Rosenthal, R. L. Rosnow, *Artifacts in Behavioral Research: Robert Rosenthal and Ralph L. Rosnow's Classic Books* (Oxford University Press, USA, 2009).

ACKNOWLEDGMENTS. For excellent research assistance, we acknowledge Adina Karpuj, Alona Friedman, Tom Benezzer, Yishai Pulvermacher, and Ran Dayan. We would also like to thank Micha Mandel and Yaron Mor for providing statistical consult and Noam Markovich and Gal Chen for reproducing our code. This research was funded by the ISF 2240/21 granted to E.H. and R.P.

Author affiliations: ^aDepartment of Psychology, The Hebrew University of Jerusalem, Jerusalem 9190501, Israel; ^bThe Department of Political Science, The Hebrew University of Jerusalem, Jerusalem 9190501, Israel; and ^cDepartment of International Relations, The Hebrew University of Jerusalem, Jerusalem 9190501, Israel