



Effects of a brief self-affirmation writing intervention among 7th graders in Germany: Testing for variations by heritage group, discrimination experiences and classroom diversity climate

Linda P. Juang¹ · Maja K. Schachner² · Tuğçe Aral¹ ·
Miriam Schwarzenthal³ · David Kunyu¹ · Hanna Löhmannsröben¹

Received: 17 November 2022 / Accepted: 24 April 2023 / Published online: 23 May 2023
© The Author(s) 2023

Abstract

We tested whether a brief self-affirmation writing intervention protected against identity-threats (i.e., stereotyping and discrimination) for adolescents' school-related adjustment. The longitudinal study followed 639 adolescents in Germany (65% of immigrant descent, 50% female, $M_{age}=12.35$ years, $SD_{age}=.69$) from 7th grade (pre-intervention at T1, five to six months post-intervention at T2) to the end of 8th grade (one-year follow-up at T3). We tested for direct and moderated (by heritage group, discrimination, classroom cultural diversity climate) effects using regression and latent change models. The self-affirmation intervention did not promote grades or math competence. However, in the short-term and for adolescents of immigrant descent, the intervention prevented a downward trajectory in mastery reactions to academic challenges for those experiencing greater discrimination. Further, it protected against a decline in behavioral school engagement for those in positive classroom cultural diversity climates. In the long-term and for all adolescents, the intervention lessened an upward trajectory in disruptive behavior. Overall, the self-affirmation intervention benefited some aspects of school-related adjustment for adolescents of immigrant and non-immigrant descent. The intervention context is important, with classroom cultural diversity climate acting as a psychological affordance enhancing affirmation effects. Our study supports the ongoing call for theorizing and empirically testing student and context heterogeneity to better understand for whom and under which conditions this intervention may work.

Keywords Brief self-affirmation writing intervention · Adolescents of immigrant descent · School-related adjustment · Classroom cultural diversity climate · Germany

1 Introduction

For adolescents of immigrant descent, being stereotyped and experiencing discrimination can undermine academic and socioemotional adjustment (Berry et al., 2006; Kunyu et al., 2020; Titzmann et al., 2011). One approach to understanding this harmful process is through considering social identity. Social identity theory emphasizes that individuals can derive a positive sense of self through the identification and sense of belonging to a particular social group (Tajfel & Turner, 1986). When social identities are devalued and threatened due to stereotyping or discrimination, however, this can compromise adolescents' academic experiences (Verkuyten et al., 2019). Studies consistently show that discrimination based on social identities (i.e., being harassed due to perceived social group membership based on ethnicity-race, gender, disability, etc.) is linked to greater distress and poorer academic performance (Russell et al., 2012).

To counter social identity threats, interventions such as brief self-affirmation writing tasks that focus on affirming *values* were developed (Cohen & Sherman, 2014; Cohen et al., 2006; Steele et al., 2002). Studies testing this type of intervention have proliferated, culminating in a recent meta-analysis of 58 studies in education (Wu et al., 2021). Notably, most studies were from the US except for nine (15%). The results showed that for identity-threatened students, i.e. students potentially experiencing social identity threat because of their group identity and associated negative stereotype, there was an overall positive and significant average affirmation effect (Hedges' $g=0.15$, a medium effect according to benchmarks for educational interventions, Kraft, 2020). For identity non-threatened students, the overall average affirmation effect was small and not significantly different from zero (Hedges' $g=0.01$). Importantly, this meta-analysis also showed a moderate to high heterogeneity of effects for identity-threatened students. Subsequently, the next generation of testing brief self-affirmation writing tasks should continue to pinpoint factors that lie behind these variations (Easterbrook et al., 2021; Hanselman et al., 2017; Wu et al., 2021). The aim of our longitudinal field experiment, then, is to test whether a randomized trial of a brief self-affirmation writing intervention protects against poorer school-related adjustment for identity-threatened students (adolescents of immigrant descent who may be more likely to experience being the target of negative stereotypes and discrimination) vs. identity non-threatened students (adolescents of non-immigrant descent). In addition, tapping both into student characteristics and context conditions, we consider heritage group, experiences of discrimination, and classroom cultural diversity climate as potential moderators that may explain the heterogeneity of effects found in previous research.

1.1 Being a target of stereotypes and discrimination

The experience of being the target of negative stereotypes and living with stereotype threat—the fear of confirming a negative stereotype about one's group (Steele & Aronson, 1995)—can be stressful. One of the most prominent stereotypes of immigrant descent adolescents in Germany include being less competent academically

(Froehlich et al., 2022), even when their competence is not lower than their non-immigrant peers (Zander et al., 2014). Pre-service teachers who read scenarios of students that confirmed stereotypes (e.g., a student of Turkish-heritage performing low academically) vs. disconfirming stereotypes (e.g., a student of Turkish-heritage being the best student in the class) were more likely to agree with statements that were negatively biased against students of immigrant descent (Glock & Krolak-Schwerdt, 2013). Pervasive negative stereotypes of immigrant students in Germany and throughout Europe, particularly of Turkish or Muslim background, increase the risk for these adolescents to experience stereotype threat and also more explicit discrimination (Baysu & Phalet, 2019; Frankenberg et al., 2013; Vedder et al., 2007). When stereotype threat is activated, Turkish-heritage adolescents in Germany performed worse on math tasks (Martiny et al., 2014). However, subsequent meta-analyses show that these effects are not strong, especially under real-life conditions (Appel et al., 2015; Froehlich et al., 2022; Shewach et al., 2019).

Related to (implicit) stereotype threat, explicitly experiencing stereotyping and discrimination in the school context not just from students but also teachers, can also contribute to disengagement from school and negatively affect academic achievement (Baysu & Phalet, 2019; Cohen et al., 2009; Easterbrook et al., 2021). A meta-analysis of 67 studies found that teacher-based racial-ethnic discrimination relates to lower well-being and academic performance among students (Civitillo et al., 2023). Taken together, being the target of stereotypes and discrimination in school by classmates and teachers undermines positive academic and socioemotional adjustment if there is nothing to offset the negative effects.

1.2 Self-affirmation theory and intervention

Self-affirmation theory suggests that if an adolescent is threatened (i.e., devalued) in an important area of the self (such as being the target of pervasive, negative stereotyping concerning academic performance), then affirming a different area of the self (e.g., reminding adolescents of the important values that they hold), can diffuse the threat by reinforcing self-integrity, promoting a sense of belonging, and adopting a more expansive view of the self (Cook et al., 2012; Sherman & Cohen, 2020; Steele et al., 2002). Therefore, self-affirmation is an important psychological intervention, especially for identity-threatened students.

Brief self-affirmation writing interventions are short writing exercises where students reflect on values that are important to them. These self-affirmation interventions have effectively diffused threat to the self and, importantly, prevented academic declines for ethnic minority students in the United States (Cohen et al., 2006, 2009; Sherman et al., 2013; Yeager & Walton, 2011). While subsequent studies did not always show effects (e.g., de Jong et al., 2016; Hanselman et al., 2017), a study in England showed that another “underperforming” stereotyped group—students of low SES—benefited from a self-affirmation writing intervention such that it reduced the academic performance gap with high SES students by 62% (Hadden et al., 2020). As noted above, while effects are not always evident in every study, a recent meta-analysis of this type of intervention in education shows an overall positive,

significant effect for identity-threatened students (Wu et al., 2021). Therefore, in our study, we hypothesize that the intervention will benefit school-related adjustment for adolescents of immigrant descent (who are assumed to experience higher stereotype threat, especially as they are also more likely to have lower SES, SVR-Forschungsbereich, 2016), but not for those of non-immigrant descent (who are assumed to experience lower stereotype threat).

To date, studies and the recent meta-analysis show heterogeneity of effect sizes, suggesting that potential moderators should be taken into account for a more precise understanding for whom and under what conditions the intervention may be effective (Easterbrook et al., 2021; Hanselmann et al., 2017; Wu et al., 2021). For instance, implementation by teachers is important as it conveys that the school environment reinforces students' affirmed views. Nonetheless, European research also shows the opposite: Whereas an intervention administered by researchers was effective amongst Turkish- and Arab-heritage German students, at least for short-term follow-up (8 weeks after the intervention was administered) (Lokhande & Müller, 2019; Müller & Lokhande, 2017), a study targeting Turkish- and Moroccan-Dutch students where the intervention was administered by teachers showed no effect on school achievement or problem behavior (de Jong et al., 2016). Nonetheless, in the meta-analysis, effect sizes were greater when the affirmation task was administered by teachers as part of a regular classroom assignment rather than researchers (Wu et al., 2021). Such variations are important to consider and empirically test.

1.3 Potential moderators producing heterogeneity of effects

Theorized heterogeneity (i.e., explanations for variations in the effectiveness of the intervention based on theory) is necessary to move the field of self-affirmation interventions forward (Easterbrook et al., 2021; Hanselmann et al., 2017). Therefore, we focus on three theoretically important moderators that focus on student characteristics and context conditions: heritage group, experiences of discrimination, and classroom cultural diversity climate.

One important potential moderator is heritage group. Attention to variation within the broad grouping of adolescents of immigrant descent is necessary to better pinpoint for whom protective factors may be more likely to work. Although large-scale studies of academic achievement report that being an immigrant predicts poorer achievement (e.g., PISA, TIMSS), these findings obscure important variations within and between different heritage groups (Brenick & Titzmann, 2015). For example, *Aussiedler* (German heritage immigrants from Russia) and their descendants attain, on average, a higher level of education compared to adolescents from other heritage groups (Fuchs & Sixt, 2007). Turkish-heritage and Arabic-speaking country heritage 7th graders in Germany (but not Eastern European-, or other heritage groups) showed higher math competence for those who participated in a brief writing affirmation intervention compared to a control group (Lokhande & Müller, 2019; Müller & Lokhande, 2017). As different heritage groups are distinguished by different

migration histories and experiences, lumping them together in one group erases important variations. Therefore, in our study we tested whether the effectiveness of the intervention may vary by specific heritage group.

Another important moderator is experiences of discrimination. Focusing on categories of heritage groups is a proxy for group-related experiences (Easterbrook et al., 2021). A more precise way to understand group variations is to directly assess those experiences. For instance, adolescents with Turkish-heritage and heritage from Arabic-speaking countries report higher levels of discrimination compared to other ethnic minority groups (Antidiskriminierungsstelle des Bundes, 2013). Rather than relying solely on heritage group categorization that may gloss over the specific mechanism to explain variation in intervention effectiveness, we include adolescents' self-reported experiences of discrimination to test whether the affirmation intervention would be more consequential for those experiencing more discrimination (and thus more identity threat) compared to those experiencing less.

In addition to student variables, considering the school context is also important (Hanselmann et al., 2014). The classroom cultural diversity climate reflects how teachers and schools approach and manage diversity (Schachner et al., 2021). It is part of the learning environment and may also moderate the effectiveness of the affirmation intervention. Two of the most widely studied classroom climate dimensions are *cultural pluralism* and *equality and inclusion*. The *cultural pluralism* dimension emphasizes that cultural diversity in the classroom should be acknowledged, valued, seen as a resource, and opportunities should be provided to learn about this diversity. The *equality and inclusion* dimension emphasizes that all students from all backgrounds should be treated equally, no one should be excluded, and opportunities for contact, cooperation, and creating common goals among diverse groups, are intentionally supported. There are two ways moderation could work. On the one hand, studies show that a positive diversity climate—both naturally observed and experimentally induced—produces similar effects as self-affirmation interventions by boosting students' sense of belonging and achievement, notably amongst identity threatened groups (Celeste et al., 2019, 2021; Schachner et al., 2019). Thus, intervention effects may be more pronounced in the absence of a positive cultural diversity climate. In our study, we expect that students from contexts with higher contextual threat in their school (e.g., in more negative cultural diversity climates, defined by less emphasis on cultural pluralism or equality and inclusion) may benefit more from the intervention (Hanselmann et al., 2014). On the other hand, drawing on the research above (e.g., Celeste, 2021) and accumulating evidence that supportive contexts afford affirmation effects to take hold and grow (Bailey et al., 2017; Walton & Yeager, 2020), it is plausible that the affirmation intervention may be more pronounced in a more positive cultural diversity climate. Further, we expect that the intervention can be even more effective when it includes implicit messages valuing pluralism and diverse cultural heritages. Therefore, besides the traditional self-affirmation intervention, we included an additional culture-sensitive intervention condition that explicitly included messages valuing cultural pluralism.

1.4 The current study

The current study tests whether students of immigrant descent in 7th grade who participate in a brief self-affirmation writing intervention administered by their teacher would be protected against poorer school-related adjustment in both the short-term (at the end of 7th grade five to six months after the intervention) and longer-term (at the end of 8th grade one year after the intervention). In the federal state of Berlin, where the study was conducted, 7th grade is the year following the transition to secondary school. We chose this grade because interventions targeted at students in transitional periods (and thus periods of higher uncertainty and threat) may be more effective than in periods of more stable development (Sherman et al., 2013).

Self-affirmation studies in education have focused predominantly on academic performance in terms of grades (Cohen et al., 2006), math motivation (Gaspard et al., 2015), or math competence (Lokhande & Müller, 2019; Müller & Lokhande, 2017). Because self-affirmation interventions can affect a range of educational outcomes (Easterbrook et al., 2021), we include grades and math competence as well as other important but less studied school-related indicators of adjustment such as disruptive school behavior, reactions to academic challenges, and behavioral and emotional school engagement.

1.5 Hypotheses

H1 (intervention effects): Adolescents of immigrant descent who receive the self-affirmation writing intervention will show better school-related adjustment compared to adolescents of immigrant descent in the control group who do not receive the intervention, with effects more pronounced in the culturally sensitive intervention condition compared to the traditional self-affirmation condition. Further, these effects will be evident in the short-term (five to six months after the intervention) as well as long-term (one year after the intervention). No intervention effects will be observed for adolescents of non-immigrant descent.

H2 (moderators of intervention effects): Because we expect that the intervention will be more effective for those experiencing more identity threat, we hypothesize that the intervention will show greater effects depending on the heritage group and experiences of discrimination. More specifically, we hypothesize that the self-affirmation intervention would be most effective for adolescents with Turkish heritage and those with South-West Asian and North African (SWANA) heritage¹ (H2a) and those experiencing more discrimination (H2b). For classroom cultural diversity climate, our test of moderation is exploratory as previous research suggests diverging plausible outcomes (H2c).

¹ SWANA heritage includes those from Egypt, Algeria, Iraq, Jordan, Palestine, Kuwait, Lebanon, Libya, Morocco, Mauritania, Tunisia, Chad, Qatar, Saudi Arabia, Oman, Somalia, Sudan, Syria, Bahrain, Djibouti, Eritrea, United Arab Emirates, Kurdish from Iraq and Yemen. They share a commonality in terms of language and religion. Most (but not all) are primarily Arabic-speaking countries.

Our study hypotheses and analyses were pre-registered before we started analyses: https://osf.io/7d4bs/?view_only=187b1e8b251d49119e7019140ebdc486. In this study, we focus on our pre-registered Research Questions 2 (Does the intervention protect against identity threat over time?) and 3 (What factors moderate the effects of the intervention?) and focus on a subset of outcomes (i.e., academic and school-related and not socioemotional adjustment). We did not focus on all outcomes and all research questions from the pre-registration as it was too much to include in one paper.

2 Method

2.1 Participants

All 211 secondary schools in Berlin were contacted by mail and followed up by phone, prioritizing those schools with a high share of students with a non-German heritage language (statistics on which are published on the website of the Berlin Senate Administration for Education, Youth, and Family). Fifteen principals (7%) agreed to participate in the study. The principals were told that it would be desirable for as many 7th grade classrooms as possible from their schools to participate in the study. Overall, 58 classes participated, including 38 at *Integrierte Sekundarschulen* (comprehensive schools or integrated secondary schools), 14 at *Gemeinschaftsschulen* (non-academic track secondary schools) and six at *Gymnasien* (academic track secondary schools).² Next, a parental consent letter was drafted, translated into Albanian, Arabic, English, French, Polish, Russian, and Turkish, and mailed to the schools for distribution amongst parents. Members of our research team introduced the study at school parent nights and were available for questions. In exchange for participation, classrooms received 25€ for their class fund at each time point. Schools received information about the study results through a newsletter.

The sample included 639 7th grade adolescents with data collected in October/November/December 2017 (T1), May/June/July 2018 (T2), and May/June 2019 (T3). Most ($n=413$, 65%) adolescents were of immigrant descent, while the rest were of non-immigrant descent ($M_{\text{age}}=12.35$ years, $SD_{\text{age}}=0.69$, 50% female). Eight percent of immigrant descent adolescents were first-generation (born outside of Germany and immigrated after they began school) and the rest were second (79%) and third (10%) generations. Adolescents represented five heritage groups:

² In our sample, SES varied by school type ($F(2,556)=15.289$, $p<.001$, $\eta_p^2=.079$). Students attending Gymnasium reported the highest SES (factor score $M=.39$, $SD=.84$), Gemeinschaftsschulen the lowest ($M=-.33$, $SD=.85$), and Integrierte Sekundarschule in-between ($M=.04$, $SD=1.05$). Percent of students of immigrant descent also varied by the three school types, $\chi^2(2)=31.87$, $p<.001$. The Gemeinschaftsschule had the highest percentage (84%), Gymnasium the lowest (56%), and Integrierte Sekundarschule in-between (60%). As school type is strongly correlated with SES and percentage of students of immigrant descent, testing for moderation by immigrant descent status/heritage group and SES addressed these potential effects of school type. Moreover, as our analyses are at the individual level and not school or classroom level, we deemed it more appropriate to include individual-level SES than school type.

non-immigrant descent ($n=211$), Turkish heritage ($n=107$), South-West Asian and North African (SWANA) heritage ($n=121$), Eastern European heritage ($n=115$), and other heritage immigrant descent ($n=70$). These groups were based on the birthplace of their parents and grandparents in combination with their self-reported heritage identity.

2.2 Procedure

We received approval from the ethics committee of the University of Potsdam and the Berlin Senate Administration for Education, Youth, and Family. Students with parental consent voluntarily completed the three surveys and two interventions. Each survey was administered by researchers with teacher support during two class periods and the survey materials were available in German, the language of instruction in the schools.

Based on information from the first questionnaire, students were grouped into heritage groups (SWANA, Turkish, Eastern European, other heritage immigrant descent, no immigrant descent). For each heritage group, the students were further classified into immigrant generations (1st vs. 2nd/3rd generation), and then into gender groups (male vs. female). Within each stratified group (heritage group by immigrant generation by gender), the students were then *randomly assigned* across three intervention conditions: traditional self-affirmation ($n=212$), culture-sensitive self-affirmation ($n=197$), and no-affirmation control group ($n=205$). Based on the effect size from a previous experimental self-affirmation study in Germany with a similar sample (Cohen's $d=0.28$, Müller & Lokhande, 2017), with Type I error of 0.05 (power of 0.95), for three groups a sample size of $N=370$ was required. Thus, our sample size had adequate power to detect the estimated main intervention effect. In addition, we conducted post-hoc power analysis for our tests of interactions. Using *semPower* (Moshagen & Erdfelder, 2016), we calculated the actual power with the available sample size to detect misspecifications of a model corresponding to $RMSEA=0.05$ with an alpha error of 0.05.

The interventions were administered by teachers during a single class period and each student received a closed envelope containing either the self-affirmation or control group writing task. The teachers received prior training and were asked to administer the tasks with the same procedure for all conditions (Cohen et al., 2006; Sherman et al., 2013). In this procedure, teachers presented the task as a normal classroom activity. Students, teachers, and researchers did not know the treatment group to which students were assigned. After administering the intervention, teachers were asked to fill out a feedback sheet responding to questions regarding the implementation such as how well students cooperated, whether students asked about why they were doing the writing task, and whether students thought the task was affiliated with the university.

2.3 Intervention and control conditions

Established self-affirmation procedures suggest administering the intervention across several time points with slightly different intervention tasks at each time point. The self-affirmation tasks in our study were structured (intervention 1) and open-ended (Intervention 2) and had previously been implemented by Sherman and colleagues (2013). To encourage greater engagement with the task, we presented the self-affirmation tasks in the form of a comic based on the German study finding that students wrote more when presented in this form (Müller & Lokande, 2017). See Supplemental Material Figures S1 and S2 for examples of the self-affirmation materials. In terms of timeliness, the 2017 school year began on September 4. Of the 58 classes, Intervention 1 was completed in November/December 2017 for 55 classes and 3 classes in March 2018. For Intervention 2, 48 classes were completed in February/March and 10 classes in May/June 2018. We aimed to administer the interventions shortly before a math test. We had also originally planned for a third intervention tailored to students' responses at Intervention 2, but were not able to schedule it because of the long delays to complete Intervention 2 due to illnesses in schools, non-responsiveness of some teachers, and school activities.

2.3.1 Traditional self-affirmation task

The tasks followed the original materials (Cohen et al., 2006) as well as slight modifications (Müller & Lokhande, 2017). For the self-affirmation task for Intervention 1 adolescents read a comic where a friendly alien came to earth and asked: "What is important to you in life? What do you like to do? Please read the following list. Mark the two most important things." Twelve values were presented (e.g., being with family/friends, being free and independent, and being good at art). In the next step, as a writing aid, students were asked to note down seven words related to why the two values were important to them. This was followed by instructions asking them to write an essay about these values, using the words that they noted down. Finally, to reinforce the affirmation, students again were asked to list the two values and asked "How do you feel when you think of these two things? When I think about these two things..." and responded to 5 items such as, "I know that I belong and am not alone."

Intervention 2 was an open-ended task. Adolescents read another comic about the alien who returned to earth and wanted to learn more, and asked them to write about what is important in their lives and why.

2.3.2 Culturally-sensitive self-affirmation task

We slightly modified the instruction of the brief self-affirmation task to explicitly affirm being of various cultural backgrounds. The modified instruction said, "In different parts of the world people find different things important. Please read through the following list. Which of these is important to you in your life? What do you like to do? Please circle the two most important things" "We also added, "You can choose to write in German or your heritage language, whatever feels more

comfortable for you.” As we wanted to explore if this culturally sensitive instruction would improve overall intervention effectiveness, half of the participants in the intervention condition received the culturally sensitive instruction. Even though most of our participants represent the second or third generation of immigrant descent and may not have sufficient writing skills in their heritage language to complete the tasks in this language (indeed, only seven students used their heritage language for Intervention 1 and three for Intervention 2), the assumption was that allowing them to use their heritage language would communicate that their background and heritage identity are valued and can be a resource when completing academic tasks in school.

2.3.3 No affirmation control group

Following the original control group task (Cohen et al., 2006) and modified as with the affirmation tasks (Müller & Lokhande, 2017), students in the no-affirmation control group were presented with a comic of an alien that came to earth and said “Some people find things important that are not so important to you. What is not important to you in life? What do you not like to do? Please read the following list. Mark the two least important things.” The same set of 12 values were presented as in the affirmation task and the rest of the writing instructions were the same.

Intervention 2 was an open-ended task. Adolescents read another comic about the alien who returned to earth and were asked to write down for the alien what time they woke up that morning, what they had for breakfast if they did, and how they got to school.

2.3.4 Fidelity check

As a check on the fidelity of the intervention, we noted whether the intervention was completed before a math test. We also had teachers note the number of students who noticed that there were different written tasks (i.e., different conditions) after doing the intervention. They also rated how much the following statements were true for them “I was motivated to do the writing tasks with the class”, “I believe the writing assignments were helpful to the students”, and “The study and writing assignments were additional tasks that made too much work for me as a teacher”. The response scale for these three items were 1 = not at all, 2 = not quite true, 3 = Neither true or not true, 4 = rather true, 5 = that is true.

2.4 Measures

We calculated mean scores for all scales. Descriptive statistics and bivariate correlations are reported in Tables 1 and 2.

2.4.1 School-related adjustment

2.4.1.1 Math, German, and English Grades Teachers were asked to report student grades in these subjects at all three time points. Because the schools used different

Table 1 Descriptive Information of the Measures for Immigrant and Non-Immigrant Descent Students in the Control and Intervention Groups

Measures	Pre-test (T1)			Post-test (T2)			Follow-up (T3)		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
<i>(Nonimmigrantdescent) Control Group</i>									
Math grades	68.57	15.12	63	61.54	18.02	52	57.38	20.48	42
German grades	69.35	12.33	62	64.30	12.97	50	64.88	12.62	41
English grades	69.60	13.62	63	64.62	17.73	52	65.35	14.93	43
Math competence	6.45	3.13	68	6.73	3.02	61	7.47	3.38	46
Mastery reactions to academic challenges	2.95	.63	67	2.85	.75	61	2.68	.81	48
Disruptive school behavior	2.02	.67	66	2.06	.68	61	2.51	.84	48
Behavioral school engagement	3.80	.71	67	3.54	.89	61	3.44	.74	48
Emotional school engagement	3.42	.83	67	3.19	.88	61	3.10	.91	47
Discrimination	1.17	.46	68	1.19	.56	61	1.17	.61	48
(Un)equal treatment	1.72	.87	67	1.69	.77	60	1.70	.76	48
Heritage and intercultural learning	2.37	.89	67	2.41	.96	60	2.31	1.01	48
<i>(Non-immigrant descent) Intervention Group</i>									
Math grades	69.23	15.49	117	66.97	18.89	107	64.83	18.82	87
German grades	69.40	12.36	117	66.31	14.66	113	66.80	14.48	85
English grades	70.04	14.36	115	66.59	15.96	109	65.30	16.04	88
Math competence	5.95	3.08	137	6.67	2.94	119	6.75	3.53	94
Mastery reactions to academic challenges	3.01	.54	135	2.89	.57	119	2.83	.51	97
Disruptive school behavior	1.88	.50	134	2.18	.82	118	2.15	.69	98
Behavioral school engagement	3.68	.71	135	3.57	.79	119	3.48	.74	97
Emotional school engagement	3.36	.76	134	3.22	.86	119	3.23	.79	97
Discrimination	1.09	.32	135	1.15	.51	117	1.19	.59	97
(Un)equal treatment	1.60	.69	135	1.65	.64	118	1.69	.68	97
Heritage and intercultural learning	2.22	.81	135	2.19	.84	118	2.25	.87	97
<i>(Immigrant descent) Control Group</i>									
Math grades	62.05	17.53	117	61.14	17.10	100	59.24	18.56	82
German grades	64.74	13.71	118	62.28	13.98	103	60.85	15.73	82
English grades	67.63	16.46	118	62.08	18.51	106	63.99	18.01	84
Math competence	4.15	2.78	134	5.31	2.83	116	5.09	3.34	94
Mastery reactions to academic challenges	3.04	.62	132	2.78	.70	116	2.83	.68	94
Disruptive school behavior	2.06	.68	133	2.17	.76	117	2.31	.77	94
Behavioral school engagement	3.83	.81	132	3.58	.80	116	3.62	.88	94
Emotional school engagement	3.47	.82	132	3.27	.81	116	3.25	.84	94
Discrimination	1.28	.61	132	1.29	.58	113	1.26	.62	94
(Un)equal treatment	1.65	.80	129	1.77	.82	111	1.73	.75	93
Heritage and intercultural learning	2.23	.92	130	2.26	.97	113	2.29	.97	93
<i>(Immigrant descent) Intervention Group</i>									
Math grades	61.03	16.94	227	59.54	19.14	192	60.14	17.58	145
German grades	64.74	13.71	229	62.02	15.35	196	61.89	15.73	141
English grades	65.72	15.56	221	62.59	17.64	198	62.44	17.34	144

Table 1 (continued)

Measures	Pre-test (T1)			Post-test (T2)			Follow-up (T3)		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Math competence	4.45	2.92	258	5.71	2.88	221	5.70	3.34	165
Mastery reactions to academic challenges	3.01	.59	252	2.85	.62	220	2.83	.61	172
Disruptive school behavior	2.08	.68	255	2.31	.80	220	2.20	.79	171
Behavioral school engagement	3.70	.77	256	3.49	.86	221	3.48	.79	172
Emotional school engagement	3.48	.80	255	3.25	.86	220	3.23	.79	171
Discrimination	1.29	.61	255	1.13	.62	219	1.36	.66	171
(Un)equal treatment	1.71	.81	254	1.71	.75	217	1.83	.80	170
Heritage and intercultural learning	2.30	.91	254	2.37	.97	217	2.44	.91	170

grading assessments (i.e., “regular” grades with 1 as *outstanding* and 6 as *fail*, points, percentages, colors, and certificates using ratings of expert, basic, etc.), we converted all forms into percentages to make them comparable (see Supplemental Material, S1 for a chart on how grades were converted into percentages).

2.4.1.2 Math competence A German standardized mathematics test, the DEMAT 6+ designed for sixth and seventh graders, was used to assess mathematical competence at T1, T2, and T3 (Götz et al., 2013). The math items of subscales A and C of the DEMAT were divided across waves. Each wave contained 15 items, seven of which remained the same while eight differed across waves. The average item difficulty was kept constant across waves, following the procedure of Müller and Lokhande (2017). The range of possible scores was 0 to 15.

2.4.1.3 Mastery reaction to academic challenges A subscale of Skinner et al. (2009) motivational perspective on engagement and disaffection measure was used to assess how persistent adolescents were in the face of academic challenges (i.e., showing a mastery reaction). An example item is, “If I can’t get a problem right the first time, I just keep trying”. The four items were rated on a 4-point Likert scale (1 = that’s not right to 4 = that’s completely right). Cronbach’s alphas were 0.73, 0.78, and 0.79 for T1, T2, and T3, respectively.

2.4.1.4 Disruptive school behavior Jenkins’ (1995) scale on school delinquency and school commitment was used to assess disruptive school behavior. Five items that were age-appropriate and used in previous research in Germany were selected (Schachner et al., 2018). An example item is, “How often did you refuse to do a task given by the teacher in the last four weeks?”. The items were rated on a 5-point Likert scale (1 = almost never to 5 = very often). Cronbach’s alphas were 0.61, 0.73, and 0.70 for T1, T2, and T3, respectively.

2.4.1.5 Behavioral and emotional school engagement Two subscales from Skinner and colleagues (2009) were used where adolescents reported on their effort, attention, and persistence during their participation in classroom activities. The measures

Table 2 Bivariate correlations of the outcomes across time

	1	2	3	4	5	6	7	8	9	10	11
1 Math grades T1	–	.52**	.43**	.49**	.18**	–.25**	.20**	.09	.46**	.15*	.22**
2 German grades T1	.53**	–	.53**	.31**	.17**	–.20**	.17**	0.07	.16**	.17**	.26**
3 English grades T1	.51**	.55**	–	.20**	.22**	–.27**	.25**	.14*	.14*	.16**	.48**
4 Math competence T1	.58**	.34**	.38**	–	.05	–.09	.05	–.04	.36**	.11*	.22**
5 Mastery react to acad. ch. T1	.31**	.27**	.29**	.34**	–	–.35**	.61**	.52**	.23**	.16**	.13*
6 Disruptive school behavior T1	–.19**	–.21**	–.16*	–.10	–.36**	–	–.47**	–.33**	–.28**	–.12*	–.22**
7 Behavioral school engagement T1	.33**	.31**	.28**	.23**	.65**	–.41**	–	.69**	.27**	.18**	.11*
8 Emotional school engagement T1	.24**	.32**	.25**	.20**	.63**	–.32**	.74**	–	.20**	.02	–.01
9 Math grades T2	.34**	.23**	.19*	.34**	.14	–.16*	.18*	.22**	–	.36**	.33**
10 German grades T2	.26**	.38**	.32**	.25**	.18*	–.14	.20**	.09	.44**	–	.34**
11 English grades T2	.06	.14	.29**	.11	.09	–.08	.16*	.07	.30**	.43**	–
12 Math competence T2	.55**	.33**	.29**	.67**	.28**	–.21**	.34**	.28**	.35**	.18*	.14
13 Mastery reactions to acad. ch. T2	.28**	.30**	.32**	.26**	.50**	–.34**	.57**	.61**	.28**	.19*	.20**
14 Disruptive school behavior T2	–.22**	–.21**	–.26**	–.18*	–.39**	.51**	–.46**	–.39**	–.19*	–.18*	–.11
15 Behavioral school engagement T2	.23**	.22**	.28**	.17*	.54**	–.36**	.71**	.62**	.25**	.24**	.21**
16 Emotional school engagement T2	.23**	.27**	.23**	.18*	.50**	–.32**	.60**	.72**	.29**	.18*	.13
17 Math grades T3	.46**	.29**	.31**	.39**	.15	–.17	.27**	.22*	.47**	.28**	.06
18 German grades T3	.36**	.49**	.27**	.25**	.20*	–.19*	.32**	.15	.32**	.54**	.19*
19 English grades T3	.16	.27**	.37**	.18*	.11	–.20*	.20*	.08	.11	.19*	.45**
20 Math competence T3	.48**	.35**	.34**	.66**	.30**	–.17*	.29**	.22**	.30**	.14	.03
21 Mastery reactions to acad. ch. T3	.33**	.28**	.24**	.29**	.40**	–.36**	.40**	.48**	.20*	.08	–.01
22 Disruptive school behavior T3	–.23**	–.16	–.19*	–.15	–.31**	.49**	–.36**	–.24**	–.13	–.07	.01
23 Behavioral school engagement T3	.34**	.34**	.25**	.29**	.36**	–.31**	.48**	.42**	.19*	.20*	–.01
24 Emotional school engagement T3	.27**	.30**	.25**	.17*	.34**	–.24**	.30**	.51**	.26**	.17	–.01

Table 2 (continued)

	12	13	14	15	16	17	18	19	20	21	22	23	24
1 Math grades T1	.50**	.19**	-.23**	.20**	.05	.44**	.28***	.27**	.45**	.20**	-.16*	.19**	.04
2 German grades T1	.32**	.22**	-.17**	.19**	0.07	.24**	.39**	.34**	.25**	.23**	-.20**	.24**	.07
3 English grades T1	.33**	.23**	-.16**	.26**	.14*	.16*	.31**	.40**	.23**	.16*	-.10	.14*	.09
4 Math competence T1	.65**	.08	-.14**	.10	-.03	.40**	.24**	.27**	.70**	.11	-.13*	.08	.03
5 Mastery react to acad. ch. T1	.07	.51**	-.38**	.47**	.35**	.08	.16*	.09	.06	.44**	-.30**	.40**	.31**
6 Disruptive school behavior T1	-.10	-.39**	.56**	-.42**	-.30**	-.07	-.11	-.010	-.10	-.31**	.41**	-.38**	-.26**
7 Behavioral school engagement T1	.07	.54**	-.38**	.63**	.53**	.09	.14*	.07	.09	.37**	-.31**	.48**	.36**
8 Emotional school engagement T1	.04	.48**	-.19**	.52**	.60**	.06	.06	-.04	.06	.37**	-.19**	.41**	.44**
9 Math grades T2	.35**	.26**	-.27**	.29**	.18**	.46**	.21**	.16*	.40**	.23**	-.24**	.24**	.12
10 German grades T2	.11	.19**	-.23**	.21**	.14*	.25**	.32**	.21**	.27**	.14*	-.17*	.23**	.07
11 English grades T2	.32**	.16**	-.19**	.18**	.11	.19**	.34**	.56**	.33**	.11	-.08	.08	.11
12 Math competence T2	-	.16**	-.16**	.16**	.04	.43**	.38**	.32**	.69**	.14*	-.29**	.20**	.12
13 Mastery reactions to acad. ch. T2	.31**	-	-.49**	.75**	.64**	.15*	.25**	.10	.19**	.65**	-.47**	.61**	.45**
14 Disruptive school behavior T2	-.23**	-.39**	-	-.54**	-.34**	-.10	-.21**	-.03	-.19**	-.39**	.59**	-.46**	-.27**
15 Behavioral school engagement T2	.26**	.67**	-.50**	-	.71**	.17*	.27**	.08	.19**	.53**	-.46**	.67**	.45**
16 Emotional school engagement T2	.30**	.68**	-.43**	.74**	-	.05	.21**	-.02	.05	.47**	-.31**	.51**	.56**
17 Math grades T3	.38**	.15	-.08	.24**	.22*	-	.33**	.30**	.43**	.20**	-.20**	.20**	.12
18 German grades T3	.18*	.17	-.21*	.32**	.20*	.39**	-	.42**	.33**	.27**	-.27**	.30**	.18**
19 English grades T3	.31**	.11	-.24**	.26**	.08	.20*	.22*	-	.31**	.16*	-.10	.10	.07

Table 2 (continued)

	12	13	14	15	16	17	18	19	20	21	22	23	24
20 Math competence T3	.61**	.23**	-.24**	.28**	.27**	.26**	.19*	.25**	-	.15*	-.24**	.15*	.09
21 Mastery reactions to acad. ch T3	.24**	.52**	-.31**	.37**	.47**	.30**	.22*	.05	.26**	-	-.43**	.69**	.59**
22 Disruptive school behavior T3	-.15	-.38**	.50**	-.37**	-.39**	-.17*	-.20*	-.11	-.19*	-.45**	-	-.51**	-.30**
23 Behavioral school engagement T3	.19*	.43**	-.25**	.44**	.50**	.31**	.35**	.12	.27**	.72**	-.45**	-	.66**
24 Emotional school engagement T3	.13	.46**	-.13	.41**	.56**	.32**	.28**	.10	.17*	.68**	-.34**	.79**	-

Correlations for adolescents of non-immigrant descent below the diagonal, correlations for adolescents of immigrant descent above the diagonal. Note: * $p < .05$, ** $p < .001$

have been used with German adolescents (Bakadorova & Raufelder, 2014). A sample item for behavioral school engagement is, “When I’m in class, I participate in class discussions” and a sample item for emotional school engagement is, “When we work on something in class, I feel interested”. Each subscale had five items that were rated on a 5-point Likert scale (1 = No, that’s not right to 5 = Yes, that’s right). Cronbach’s alphas for behavioral engagement were 0.82, 0.86, and 0.84, and for emotional engagement, 0.81, 0.83, and 0.84 for T1, T2, and T3, respectively.

2.4.2 Moderators

2.4.2.1 Heritage group As noted above, adolescents were categorized into five heritage groups: Turkish, SWANA, Eastern European, other heritage immigrant descent, no immigrant descent.

2.4.2.2 Discrimination We used the perceived discrimination subscale from a measure for migration-specific hassles that was initially developed for adolescents of immigrant descent in Germany (Titzmann et al., 2011). Six items assessed the events of unfair or negative treatment experienced by adolescents that were attributed to their ethnic background in the last year. Although this scale was developed for adolescents of immigrant descent, all adolescents responded to the measure. A sample item for discrimination is, “My schoolmates laughed at me because of my heritage culture”. The items were rated on a 5-point Likert-scale (1 = never, 2 = 1 to 2 times, 3 = 3 to 5 times, 4 = 6 to 10 times, and 5 = more than 10 times). Cronbach’s alphas for discrimination were 0.86, 0.88, and 0.87 for T1, T2, and T3, respectively.

2.4.2.3 Classroom cultural diversity climate We used two subscales of the Classroom Cultural Diversity Climate Scale (Schachner et al., 2021): the 5-item unequal treatment subscale to measure the opposite of equality and inclusion, and the 7-item heritage culture and intercultural learning subscale to measure cultural pluralism. A sample item for unequal treatment is “Students from certain heritage cultures are teased more often than students from other heritage cultures”. A sample item for the heritage and intercultural learning subscale is “During class we learn about the heritage cultures of fellow students”. Items were rated on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree). Cronbach’s alphas for unequal treatment were 0.81, 0.74, and 0.75 for T1, T2, and T3, respectively. For heritage culture and intercultural learning, Cronbach’s alphas were 0.87, 0.88, and 0.89 for T1, T2, and T3, respectively. Higher scores indicated higher unequal treatment and greater heritage and intercultural learning.

3 Results

3.1 Randomization check

To ensure randomization was effective, we tested for pre-intervention (T1) differences on demographics (age, SES, importance of religion, German writing, German reading, cognitive test) and main study variables between students in the control and intervention conditions. The adolescents in the control, traditional intervention, and culture-sensitive intervention groups did not differ in their scores at pre-test (see Supplemental Material, Table S2 for a more detailed description of how demographics were measured and results of this randomization check).

3.2 Opportunity gap check

We tested baseline (T1) differences on study variables between immigrant and non-immigrant descent students across all treatment conditions as an indication of the degree of “achievement disparity” (Wu et al., 2021) and varying contextual conditions. Adolescents of immigrant descent differed from those of non-immigrant descent in math competence, $F(1,621)=49.724$, $p<.001$, $\eta_p^2=0.074$, and disruptive behavior, $F(1,611)=6.448$, $p=.011$, $\eta_p^2=0.010$, but not on mastery reactions to academic challenges or behavioral and emotional school engagement across all treatment conditions. Non-immigrant descent students scored higher on math competence ($M=6.09$, $SD=3.08$) than immigrant descent students ($M=4.33$, $SD=2.89$). They also reported lower disruptive school behavior ($M=1.94$, $SD=0.57$) than immigrant descent students ($M=2.08$, $SD=0.67$).

We also tested whether students of immigrant descent differed from those of non-immigrant descent regarding contextual moderators. At baseline, students of immigrant descent ($M=1.28$, $SD=0.61$) reported higher discrimination than students of non-immigrant descent ($M=1.11$, $SD=0.37$), $F(1,614)=13.270$, $p<.001$, $\eta_p^2=0.021$. Students of immigrant descent ($M=2.27$, $SD=0.91$) and non-immigrant descent ($M=2.26$, $SD=0.83$) did not differ in heritage and intercultural learning classroom climate, $F(1, 610)=0.005$, $p=.942$, $\eta_p^2=0.000$. Students of immigrant descent ($M=1.69$, $SD=0.80$) and non-immigrant descent ($M=1.64$, $SD=0.76$) also did not differ in unequal treatment classroom climate, $F(1, 609)=0.468$, $p=.494$, $\eta_p^2=0.001$.

3.3 Manipulation check

As a manipulation check we based our procedures on Hanselmann et al. (2017). Two trained independent coders who were blind to the experimental condition coded all essays (first independently coding then discussing and reaching consensus on codes not in agreement) to examine whether students in the intervention group wrote more

affirming responses (e.g., by expressing either the importance of the value for themselves, by expressing that they are “good in” the valued domain, or by expressing positive emotions) than students in the control group. Essays of the students in the intervention group were coded as more self-affirming than essays of the students in the control group for both the first, $\chi^2(2)=437.77$, $p < .001$, and second intervention, $\chi^2(2)=20.180$, $p < .001$. Most of the essays in Intervention 1 (95%) and Intervention 2 (87%) in the intervention group were coded as expressing self-affirmation.

3.4 Fidelity check

A majority of students (67%) did Intervention 1 before a math test while (33%) did Intervention 1 one day after.³ For Intervention 2, all students did the intervention before a math test. Most of the teachers ($n=45$ out of 47 teachers) who gave the intervention responded to fidelity check questions. After both interventions, 43% of teachers reported that no students noticed different written tasks and the mean number of students the teachers reported noticing was 2.54. Thus, on average, two students in a class noticed differences.⁴ Of those who filled out the teacher survey, 28% of the teachers reported either “not quite true” or “not at all” to being motivated to do the writing task while 39% reported either “rather true or “true”. Thirty-four percent reported either “not quite true” or “not at all” to seeing the writing assignments as beneficial for the students, 11% reported “rather true” and none reported “true”. Finally, 43% of the teachers reported either “rather true” or “true” that the intervention was an additional task that caused extra workload, while 32% reported “not quite true” or “not at all”.

3.5 Attrition and missing data

A total of 639 adolescents participated in the pre-test (T1) and at least one of the two intervention tasks. Among those, 86% ($n=549$) participated in the post-test (T2). Among the adolescents present in both pre-test (T1) and post-test (T2), 73% ($n=399$) participated in the follow-up (T3). The attrition rates in the intervention and control groups between the pre-test (T1) and the post-test (T2) were similar

³ There was no difference in math results between students who did intervention 1 before or after a math test. For instance, after selecting those who did intervention 1 before a math test, the intervention and control groups did not differ regarding math grades from T1-T2 (Wald $\chi^2(1, n=316)=.602$, $p=.43$) or T2-T3 (Wald $\chi^2(1, n=316)=3.354$, $p=.07$). They also did not differ regarding math competence from T1-T2 (Wald $\chi^2(1, n=374)=.240$, $p=.62$) or from T2-T3 (Wald $\chi^2(1, n=374)=.140$, $p=.70$).

⁴ In additional analyses, we selected those students in classrooms where no one noticed the different intervention conditions and tested the direct effect of the intervention on the study outcomes. The sample size decreased in these analyses, $n=250$. The main effects of the intervention were the same as with the larger sample. Namely, the main effects were significant only for disruptive school behavior and not for other outcomes. For disruptive school behavior the change scores differed between the intervention and control groups for T1-T2 (Wald $\chi^2(1, n=250)=3.694$, $p=.054$) and for T2-T3 (Wald $\chi^2(1, n=250)=4.53$, $p=.033$). Model fit was acceptable, $\chi^2(250)=738.020$, $p < .001$; CFI = .958, TLI = .949, RMSEA = .044 (90% CI = .003-.066), SRMR = .09).

(control: 12%, traditional intervention: 14%, culture-sensitive intervention: 15%, $\chi^2(2)=0.82$, $p=.66$, $\eta_p^2=0.001$). The attrition rates in the intervention and control groups between the post-test and follow-up were also similar (control: 26%, traditional intervention: 29%, culture-sensitive intervention: 23%, $\chi^2(2)=1.50$, $p=.47$, $\eta_p^2=0.001$).

For further attrition analyses, we followed the same approach from another longitudinal study (Miklikowska et al., 2019). To test whether attrition from T1 to T2 ($n=90$) or from T1 to T3 ($n=204$) was related to the demographic and the study variables, we conducted logistic regression testing whether attrition (0=dropped, 1=retained) was predicted by demographic (e.g., gender) or the study variables at T1. The results showed that attrition from T1 to T2 was related to the cognitive skills test, behavioral school engagement and discrimination. Adolescents who scored higher at the cognitive skills test were more likely to participate at both time points (T1 and T2) than adolescents who scored lower at the cognitive skills test, $F(1,637)=8.856$, $p=.003$, $\eta_p^2=0.030$. Adolescents who reported lower discrimination were more likely to participate at both time points (T1 and T2) than those who reported higher discrimination, $F(1,627)=5.697$, $p=.017$, $\eta_p^2=0.013$. Adolescents who reported higher behavioral school engagement were more likely to participate at both time points (T1 and T2) than those who reported lower behavioral school engagement, $F(1,627)=16.035$, $p<.001$, $\eta_p^2=0.038$. Nagelkerke R^2 of the regression model was 0.08.⁵

Attrition from T1 to T3 was not significantly related to the demographic (e.g., gender) or study variables except the cognitive skills test. Again, adolescents who scored higher on the cognitive skills test were more likely to participate in all time points (T1, T2, and T3) than adolescents who scored lower on the cognitive skills test, $F(1, 637)=15.354$, $p<.001$. Nagelkerke R^2 was 0.03. We added the cognitive skills variable to the models as an auxiliary variable (Graham, 2003).

Missing data for the study variables were less than 7% within each time point. Participants with and without complete data were compared on the study variables using Little's (1988) missing completely at random test within each time point. Results suggested that missing values were missing completely at random for T1, $\chi^2(78, n=639)=67.114$, $p=.80$, not missing completely at random for T2, $\chi^2(14, n=536)=27.147$, $p=.018$, and again completely at random for T3, $\chi^2(21, n=399)=17.950$, $p=.652$. Due to the small amount of missing data overall, we used full information maximum likelihood (FIML) to handle missing data.

3.6 Measurement invariance

We tested measurement invariance for all study measures across time and immigration status groups (immigrant descent vs non-immigrant descent; see Supplemental

⁵ Nagelkerke R^2 refers to an adjusted version of Cox and Snell R-square that are usually used in logistic regression analyses (Allison, 2013). An adjusted Nagelkerke R^2 ranges between 0 to 1 and refers to the variance explained by the independent variable. Therefore, our percentages of 0.8% and 0.3% are small and not likely to affect subsequent analyses.

Material, Tables S3-S9) as measurement invariance is necessary to meaningfully interpret latent difference scores (Geiser, 2012). We first tested configural invariance (specifying the same measurement model in all groups, without constraining factor loadings or intercepts to be equal across groups), metric invariance (factor loadings were constrained to be equal across groups), and scalar invariance (intercepts were constrained to be equal across groups). Metric invariance implies that associations can be compared across groups and scalar invariance implies that means can be compared across groups.

All measures demonstrated scalar invariance across time and immigrant status group except for one measure. One item (i.e., How often have you been too late for class in the morning or after lunch break in the last four weeks?) of the disruptive school behavior measure was removed to establish scalar measurement invariance (with four items) to proceed to testing a latent change model.

3.7 Steps to test hypotheses

We built multigroup latent change models using Mplus, using intervention vs. control condition as a grouping variable. In each model, we allowed the same items to correlate with each other across time. We then generated latent change variables/scores for the change between T1 and T2 and the change between T2 and T3. In a latent change model, a latent state variable at T2 is a function of the latent state variable at T1 plus a variable capturing the change from T1 to T2 (Geiser, 2012). In other words, because T1 scores between intervention and controls groups are the same, the calculated latent state variable at T2 reflects only the change in latent scores between T1 and T2. Likewise, the calculated latent state variable at T3 reflects only the change in latent scores between T2 and T3.

We then used the “model test” function in multigroup analyses to compare whether the change score for T1 to T2 and T2 to T3 significantly differed between intervention and control groups. We did this separately for students of immigrant descent and non-immigrant descent (H1). We did not add any covariates to the models as groups had been randomized by participants’ heritage group, generation, and gender within classes. However, based on preliminary analysis regarding attrition, we added cognitive skills as an auxiliary variable to the models.

Regarding moderators, to test whether the intervention was effective for specific heritage groups among immigrant descent adolescents, we used a categorical variable for heritage groups and tested the direct effects of the intervention by selecting sub-samples (H2a). However, the sample size for each heritage group was small (e.g., $n = 114$ for SWANA, $n = 107$ for Turkish heritage).

To test the moderators of discrimination and classroom cultural diversity climate among immigrant descent adolescents, we used a single group model where change scores for T1 to T2 and T2 to T3 were regressed on the treatment condition (0 = control, 1 = intervention), the moderator, and the interaction term between the treatment condition and the respective moderator (H2b and H2c).

Against our hypotheses, we note that there were no differences between the traditional and culture-sensitive self-affirmation tasks in any of the main analyses.

Therefore, we combined the two intervention conditions into one group and report main results based on this combined intervention group vs. control group, below.

3.8 Effects of the intervention on school-related adjustment

With regards to Math, German, and English grades, we could not use latent change models because each grade was assessed with just one item. Therefore, we tested for an intervention effect with two steps. First, we used a multigroup (intervention vs. control) regression model where grade (each tested separately) at T2 was regressed on T1 and grade (each tested separately) at T3 was regressed on T2. Using the model test function, we examined whether the regression parameters were equal across the two groups. Based on the best fitting model, we calculated the predicted values (the estimated model parameters) for T1, T2, and T3 to test for an intervention effect. Second, we combined the intervention and the control groups. Controlling for grade at T1, grade at T2 was regressed on a categorical variable representing the intervention conditions (0 = control, 1 = intervention). Controlling for grades at T1 and T2, grade at T3 was regressed on the intervention conditions. Based on these steps, there was no intervention effect for students of immigrant descent, non-immigrant descent (H1), or specific heritage groups (H2a). Discrimination and two dimensions of classroom cultural diversity climate did not moderate intervention effects on grades for students of immigrant descent (H2b & H2c).

We could not use latent change models for math competence because the scale was one ordinal item ranging from 0 to 15. Therefore, we followed the same analytic approach with grades. We found no intervention effect on math competence for students of immigrant descent, non-immigrant descent (H1), or specific heritage groups (H2a). Discrimination and two dimensions of classroom cultural diversity climate did not moderate intervention effects for math competence for students of immigrant descent (H2b & H2c).

With regards to mastery reactions to academic challenges, change scores did not differ between the control and intervention groups for T1-T2 and T2-T3 for students of immigrant descent, non-immigrant descent (H1), or specific heritage groups (H2a). However, students of immigrant descent who experienced high levels of discrimination may have benefited more from the intervention, at least in the short term (H2b). More precisely, they showed a smaller decline in their mastery reactions to academic challenges compared to students who experienced high levels of discrimination in the control group (interaction term approached significance, $\beta = 0.243$, $p = .064$) at T2. But this effect was not significant in the long term, meaning that students of immigrant descent who experienced high levels of discrimination ended up with similar scores on mastery reactions to academic challenges at T3 regardless if they received the intervention or not (see Fig. 1 and Table 3). The post hoc power analysis suggested that with the available sample size ($n = 394$), we had sufficient power (99.88%) to reject a poor-fitting model ($df = 85$). The two dimensions of classroom diversity climate did not moderate the intervention effects on mastery reactions to academic challenges in the short- or long-term for students of immigrant descent (H2c).

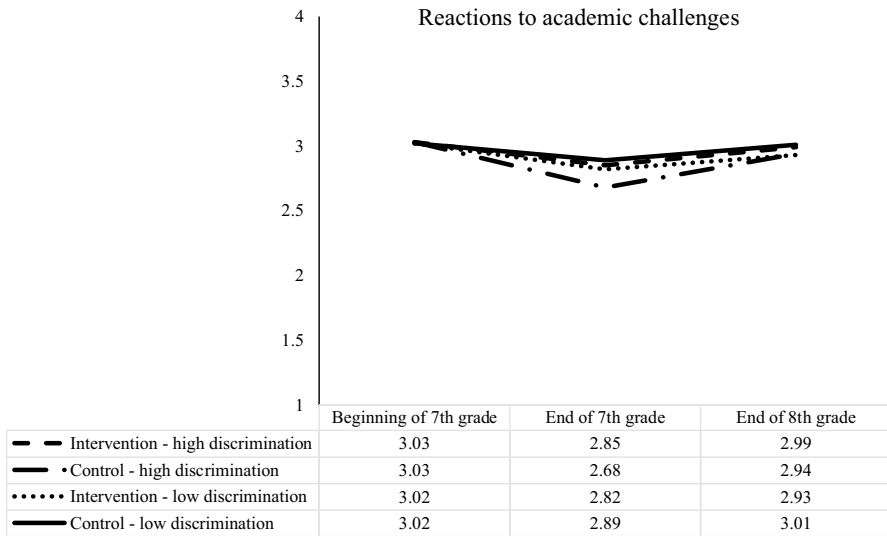


Fig. 1 Discrimination as a Moderator of the Intervention Effect on the Reactions to Academic Challenges for Students of Immigrant Descent ($n=413$, low = 1 *SD* below the mean, high = 1 above the mean)

With regards to disruptive school behavior, change scores differed between the intervention and control groups for T1-T2 and T2-T3 for both students of immigrant descent (Wald χ^2 (1, $n=413$)=4.489, $p<.05$ for T1-T2; Wald χ^2 (1, $n=413$)=7.257, $p<.01$ for T2-T3) and non-immigrant descent (Wald χ^2 (1, $n=226$)=4.535, $p<.05$ for T1-T2; Wald χ^2 (1, $n=226$)=14.214, $p<.001$ for T2-T3) (H1). More specifically, from T1 to T2, the intervention group initially showed greater disruptive behavior than the control group, but from T2 to T3 this reversed. By T3, the intervention group showed less disruptive behavior compared to the control group for students of non-immigrant descent (see Figs. 2 and 3 and Table 3). In other words, the intervention lessened an upward trajectory in disruptive behavior for both groups. This finding did not support our hypothesis as the intervention benefitted *both* students of immigrant and non-immigrant descent concerning disruptive behavior, at least in the long term. The change scores in disruptive behavior for T1-T2 and T2-T3 did not differ by intervention condition for specific heritage groups (H2a). Discrimination and the two dimensions of classroom climate did not moderate the intervention effects on disruptive behavior for students of immigrant descent (H2b and H2c).

With regards to behavioral school engagement, change scores did not differ between the intervention and control groups for T1-T2 or T2-T3 for students of immigrant descent, non-immigrant descent (H1) or specific heritage groups (H2a). Discrimination and unequal treatment climate did not moderate the intervention effects among students of immigrant descent (H2b and H2c). But there was a significant interaction for the intervention effect and heritage/intercultural learning classroom climate between T1 and T2 ($\beta=0.281$, $p=.013$) and T2 and T3 ($\beta=0.288$, $p=.025$) for students of immigrant descent (H2c, see Fig. 4 and

Table 3 Model fit indices for the latent change models that show significant effects
Mastery reactions to academic challenges

Discrimination as moderator for immigrant descent students (H2b)	$\chi^2(413) = 176.835, p < .001$; CFI = .935, TLI = .922, RMSEA = .051 (90% CI = .040—.062), SRMR = .126
<i>Disruptive school behavior</i>	
Direct effects for immigrant descent students (H1)	$\chi^2(394) = 147.790, p < .01$; CFI = .960, TLI = .951, RMSEA = .043 (90% CI = .024—.060), SRMR = .07
Direct effects for non-immigrant descent students (H1)	$\chi^2(205) = 152.210, p < .01$; CFI = .931, TLI = .916, RMSEA = .063 (90% CI = .037—.085), SRMR = .100
<i>Behavioral school engagement</i>	
Heritage/intercultural learning norms as moderator for immigrant descent students (H2c)	$\chi^2(413) = 424.008, p < .001$; CFI = .895, TLI = .876, RMSEA = .075 (90% CI = .067—.083), SRMR = .152

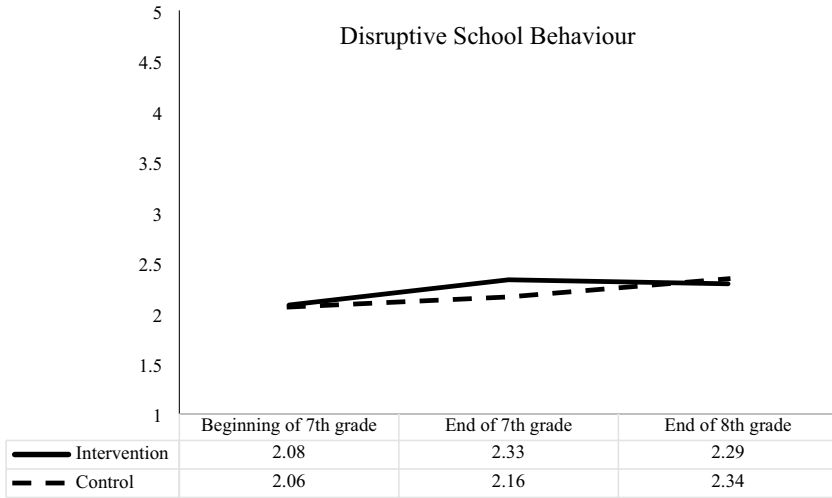


Fig. 2 Direct Effects of the Intervention on Disruptive School Behavior for Students of Immigrant Descent ($n=413$, low = 1 SD below the mean, high = 1 SD above the mean)

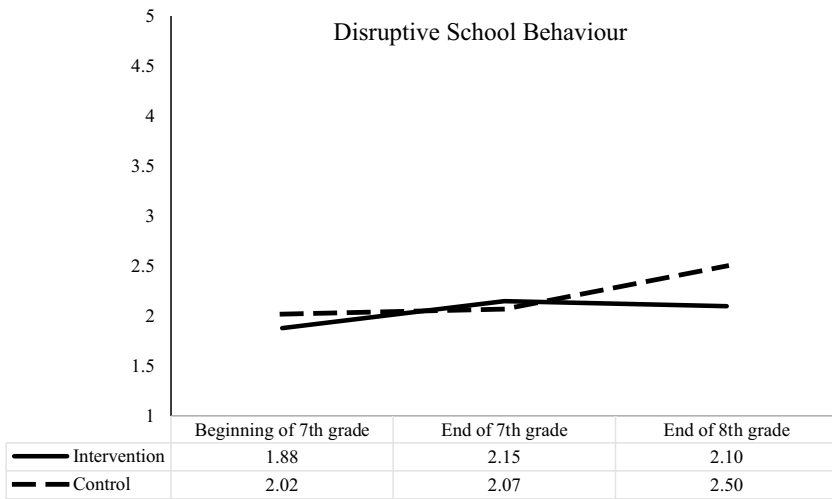


Fig. 3 Direct Effects of the Intervention on Disruptive School Behavior for Students of Non-Immigrant Descent ($n=226$, low = 1 SD below the mean, high = 1 SD above the mean)

Table 3). In the short-term (T2), students of immigrant descent who perceived high heritage and intercultural learning in the classroom benefitted more from the intervention (had a smaller decline in behavioral engagement) compared to those who perceived high heritage and intercultural learning in the classroom but

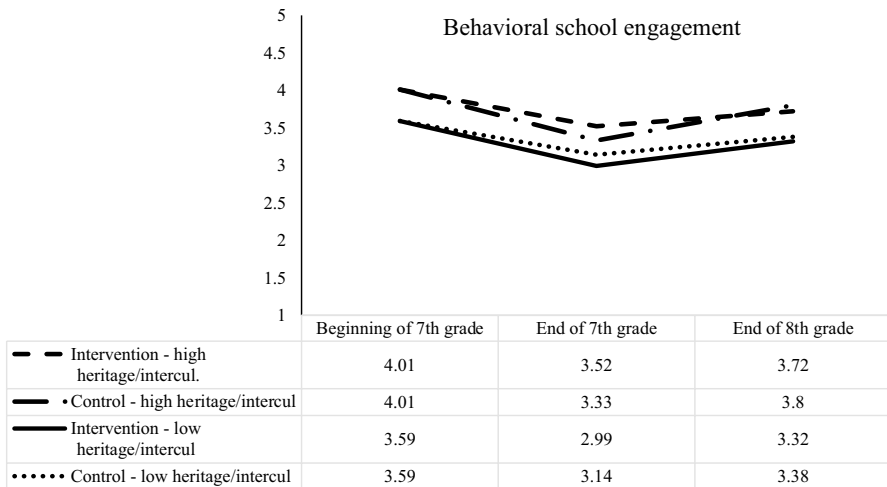


Fig. 4 Heritage and Intercultural Learning Climate as a Moderator of the Intervention Effect on Behavioral School Engagement for Students of Immigrant Descent ($n=413$, low = 1 *SD* below the mean, high = 1 *SD* above the mean)

did not receive the intervention. At T3, the intervention effect was not sustained such that the intervention and control groups both increased in behavioral school engagement similarly. The post hoc power analysis suggested that with the available sample size ($n=413$), we had sufficient power (99%) to reject a poor-fitting model ($df=127$). Overall, students in the higher heritage and intercultural learning classroom showed greater behavioral school engagement compared to those in lower heritage and intercultural learning classrooms across all time points.

With regards to emotional school engagement, change scores did not differ between the control and intervention group for T1-T2 or T2-T3 for adolescents of immigrant descent, non-immigrant descent (H1), or specific heritage groups (H2a). Discrimination and two dimensions of classroom climate did not moderate intervention effects for adolescents of immigrant descent (H2b and H2c).⁶

⁶ We also conducted additional exploratory, non-pre-registered analyses to test whether gender and SES moderated the relations between condition and adjustment indicators. SES did not moderate the relations. Gender was a significant moderator for disruptive school behavior. For students of immigrant descent, the interaction of gender and the effect of the intervention was significant between T1 to T2 ($\beta=.311$, $p<.01$) and not significant between T2 to T3 ($\beta=. -170$, $p=.173$, see Figure S6). For students of non-immigrant descent, the interaction of gender and the effect of the intervention was marginally significant between T1 to T2 ($\beta=.322$, $p=.065$) and significant between T2 to T3 ($\beta=. -078$, $p=.046$, see Figure S7). The results suggest that the intervention worked differently by gender. More specifically, compared to the control group, boys in the intervention group (but not girls) increased in disruptive behavior in the short-term, but in the long-term showed less disruptive behavior. We added the two figures to Supplemental Materials with a note that model fit for students of non-immigrant descent was not very good.

4 Discussion

Adolescents of immigrant descent may experience stereotyping and discrimination, creating threats to their identity that may undermine their school-related adjustment. In our study we tested whether engaging in a brief self-affirmation writing task would protect adolescents against a decline in school-related adjustment from the beginning of 7th to the end of 8th grade. We also considered whether heritage group, experiences of discrimination, and classroom cultural diversity climate would contribute to heterogeneity of effects. Our overall conclusion is that the brief self-affirmation writing intervention shows a small effect on some aspects of school-related adjustment for adolescents of immigrant descent and may sometimes be beneficial also for adolescents of non-immigrant descent. The *context* of the intervention is also important, with classroom cultural diversity climate acting as a psychological affordance to enhance affirmation effects, at least in the short term.

We did not find that the self-affirmation writing intervention benefited students of immigrant descent regarding grades or math competence, two of the most widely studied academic outcomes in education (Wu et al., 2021). Our findings contrast to Müller and Lokhande's (2017; Lokhande & Müller, 2019) study that had a very similar sample drawn from the same city and found their intervention to be effective for boosting math competence using the same measure. Several differences with their study may explain our divergent findings. In their study the intervention was administered by the researchers and the effects were assessed immediately after the intervention and again eight weeks later. In our study the intervention was administered by teachers and effects were assessed five to six months and one year later. We trained teachers to implement the intervention based on the theoretical importance of having affirmation clearly linked to school (Easterbrook et al., 2021). While we tried to be consistent in the training, teachers also varied in motivations. Some were enthusiastic about the intervention and others less so. Perhaps this variation in teacher training compliance and effectiveness contributed to fewer intervention effects compared to the previous study. Indeed, implementation fidelity has been identified as a likely moderator of self-affirmation intervention effectiveness also in previous research, and this particularly applies for interventions implemented by teachers (Bradley et al., 2016). For teacher-implemented studies with smaller samples, procedures can be followed and monitored more carefully and close collaborations can be established with teachers. This was difficult to attempt to do in a larger-scale study such as ours, and indeed some of the larger self-affirmation studies with the goal of scaling up the intervention were less successful than the initial, smaller studies.

In addition, especially in the East German federal states including Berlin, many schools are under-resourced and there is a drastic shortage of teachers. This becomes even more problematic when facing additional challenges, such as providing schooling for large numbers of refugee students who arrived in the years of our study. It means that there is less time available for teacher training and careful implementation of an intervention in addition to their regular teaching content. Scaling up

paired with teacher implementation may therefore cumulate in additional challenges for intervention effectiveness.

Despite the absence of a main intervention effect on grades and math competence, we did find effects on a few outcomes. Our findings for mastery reactions to academic challenges are somewhat in line with self-affirmation findings showing that the intervention prevented a decline in academic motivation and performance (Cohen et al., 2009; Sherman et al., 2013), but under specific conditions. Our findings show (tentatively as these results approached significance at $p = .064$) that adolescents of immigrant descent who participated in the intervention were more likely to show less decline in their mastery reactions to academic challenges, at least in the short term, from T1 to T2, but only those who experienced higher (not lower) levels of discrimination. These results suggest that those who experience the greatest direct threat to their sense of self could be targeted for affirmation interventions to reinforce and make clear that they are valued. Finding that these effects last only in the short-term demonstrates the importance of studying longer-term effects to draw conclusions about overall effectiveness. Yet finding a short-term effect also suggests there may be windows of opportunity for intervention effectiveness that might make sense to target. The non-immigrant descent adolescents did not show any of these effects for mastery reactions to academic challenges, in line with theoretical reasoning that the intervention would not affect adolescents assumed to experience less identity threat, especially in an educational context.

Regarding disruptive school behavior, we found that the self-affirmation writing intervention prevented an upward trajectory in disruptive school behavior from the 7th to end of 8th grade for all adolescents, regardless of immigrant status. While disruptive school behavior can increase throughout secondary school (Zimmermann et al., 2013), the intervention appears to have prevented some of this increase. Engaging in disruptive behavior is consequential for relationships with teachers, school grades, and perceptions of the self (Zimmermann et al., 2013), and thus an important outcome to pay attention to.

Our results do not align with findings with sixth and seventh grade Dutch students, where a brief self-affirmation writing task did not reduce problem behaviors across an academic school year (de Jong et al., 2016). Yet our results do align with findings with U.S. students, where multiple brief self-affirmation writing tasks across three years reduced discipline incidents (taken from school records) in 7th and 8th grade, with no difference by ethnic group (Binning et al., 2019). Similarly, contrary to hypotheses, the intervention in their study did not solely benefit identity-threatened adolescents, but all adolescents.

There is some evidence that self-affirmation tasks can reduce effects of identity threat in general and not only threats tied to a specific group identity (Cohen & Sherman, 2014). In Germany, the often-used categories of immigrant descent vs. not is contested as a relevant or meaningful identity as it is a label that the majority of people in Germany officially designated as “migration background”, do not self-identify with (Nesterko & Glaesmer, 2019). The risk for susceptibility to stereotype threat has been argued to be greater for those who identify more strongly with the group that is being targeted, but there are very few studies in Europe on this topic and the few show mixed evidence for immigrant groups (Appel et al., 2015). This

may be a reason why we did not find consistent differences in the effectiveness of the intervention for these two groups.

Classroom cultural diversity climate moderated the effect of the self-affirmation writing task on behavioral school engagement. In the short term, students of immigrant descent who perceived high heritage and intercultural learning in the classroom (*less* identity threat) showed a smaller decline in behavioral school engagement (i.e., benefitted more from the intervention) compared to students who perceived low heritage and intercultural learning in the classroom. In the long-term, however, there were increases in behavioral school engagement for all students from the end of 7th into 8th grade, regardless of treatment condition. Thus, a positive classroom diversity climate acted as a protective factor for those participating in the intervention in the short-term, and later, as an enhancing factor for all adolescents in the long-term. These results illustrate the importance of defining both vulnerabilities and opportunities when considering interventions within real-life social contexts (Walter & Yeager, 2020). Our results suggest that a positive diversity climate in the classroom could be considered a *psychological affordance*, a feature of the context that can act as a *resource* to allow an affirmation intervention to take hold (Walter & Yeager, 2020) as well as foster positive school adjustment beyond the intervention. Our findings are also in line with findings from a previous study showing that trust in teachers and school, and a more positive school climate may contribute to positive affirmation intervention effects (Binning et al., 2019).

Against our hypothesis, we did not find that the intervention was effective in changing emotional school engagement. While negative stereotypes for immigrant descent students and families in Germany, especially for those of Turkish- or Arab-heritage regarding school engagement, for instance in terms of being *bildungsfern* (“academically distant”) (Emmerich & Hormel, 2013), perhaps the stereotype is not as pronounced towards the entire group of students of immigrant descent. Alternatively, not all members belonging to the same negatively stereotyped group are equally affected because they may not be aware of the stereotype, do not endorse the stereotype, or do not identify with the group (Spencer et al., 2016). We did not find other studies that included this dependent variable so we await future studies to test this further.

Regarding heritage group as moderator, we did not find it moderated the effectiveness of the intervention. Heritage groups are broad categories that do not directly assess potentially important experiences, such as the degree of discrimination experienced. Discrimination indeed moderated some of the intervention effects, suggesting it is more important to consider such experiences rather than solely rely on categorizations.

Based on theory and empirical evidence, the self-affirmation writing intervention should be more effective when the “achievement gap” between identity-threatened adolescents and identity non-threatened adolescents is large (Wu et al., 2021). In our study for the whole sample at pre-test, we found that adolescents of immigrant descent differed from those of non-immigrant descent in math performance and disruptive behavior, but not on mastery reactions to academic challenges or behavioral and emotional school engagement. Further, these differences were small. The fact that the two groups did not show substantial school-related adjustment gaps in most

of the study variables, could also explain why we found few effects. Despite persistent educational disparities between students of immigrant descent and non-immigrant descent in Germany, the school system in the federal state of Berlin is more inclusive than elsewhere in Germany, with a longer period of primary schooling and more within-school (as opposed to between-school) tracking in integrated secondary schools (Teltemann & Schunck, 2016).

Overall, our results speak to the importance of considering a range of conditions to clarify self-affirmation effects in educational settings as the quality and timing of implementation, demographics and experiences of students and teachers, and availability of school resources can all contribute to heterogeneity of effects (Easterbrook et al., 2021; Hanselmann et al., 2017). Notably, in the Wu et al. (2021) meta-analysis, of the nine studies that were conducted outside the US, only four included adolescents, with two studies finding effects (England and Germany) and two that did not (Netherlands). Although the overall meta-analytic effect is positive, significant, and medium size for identity-threatened students, which is encouraging, the authors clearly note that because of the great heterogeneity in effect sizes, it will be the case that single studies may show a null effect (Wu et al., 2021). As noted earlier, future studies should continue to detail moderators on various levels to uncover when and for whom this intervention may be effective.

4.1 Limitations

There were some limitations to our study that could have affected our study findings. One such limitation is that our timing and dosage were not ideal. We had planned to implement the first intervention at the very beginning of the year, directly after students transitioned to secondary school, based on the importance of starting before negative recursive processes begin, especially during or after a transition (Sherman et al., 2013). However, a delay in receiving ethics approval prevented us from doing so and thus we started further into the school year (eight weeks) than desired. Indeed, a 4-week delay can significantly decrease the effectiveness of the affirmation intervention (Cook et al., 2012). We also originally planned for three interventions but were only able to complete two. Nonetheless, dosage may be less important than timing (Wu et al., 2021) and other aspects of our study were well-timed in that we focused specifically on 7th graders after a period of transitioning into secondary school and attempted to schedule interventions before math tests.

4.2 Strengths and avenues for future research

One strength is that we tested for theorized heterogeneity and focused on student (heritage group, experiences of discrimination) and contextual (classroom quality climate) moderators. Other moderators such as implementation quality (Bradley et al., 2016), teacher-related (e.g., growth mindset), and other school-related aspects (Easterbrook et al., 2021; Hanselmann et al., 2017; Wu et al., 2021) are also important. In future studies we plan to also test mediators to explain why self-affirmation writing interventions work in terms of the underlying psychological processes

(Easterbrook et al., 2021; Wu et al., 2021). Based on self-determination theory (Ryan & Deci, 2020), we included measures of sense of belonging, autonomy, and competence. There is evidence that a key mechanism for why self-affirmation tasks work is enhancing a sense of belonging (Celeste et al., 2021; Cook et al., 2012; Shnabel et al., 2013). A greater sense of belonging seems to provide a context for doing well in school academically (Celeste et al., 2019). Relatedly, self-affirmation tasks can promote a sense of trust with teachers and school and contribute to better school behavior among adolescents (Binning et al., 2019). Less known is whether self-affirmation tasks enhance the two other basic needs of self-determination—a sense of autonomy and competence—so we plan to test this in the future.

Our findings are useful in that the intervention seems to offer some protection for school-related adjustment, for both students of immigrant and non-immigrant descent, depending on the particular aspect of adjustment and context. Although we did not find strong support for all hypotheses with our wide range of outcomes, the strengths of the study are the theoretical foundation, adequate sample size, real-life conditions and longer-term school-related adjustment measured. Our findings are still important to demonstrate under what conditions and for what outcomes the intervention does and does not seem to work.

We also would like to highlight our agreement with recent studies that recognize the limitations of only focusing on affirming students with an individual-level intervention (Easterbrook et al., 2021; Wu et al., 2021). Broader societal and structural level supports, as well as within the school, will be necessary to provide the resources and nurturing environment that will allow students to be affirmed, valued, and supported for who they are (Walton & Yeager, 2020).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11218-023-09789-9>.

Acknowledgements This research is funded by the German Research Foundation (DFG), Project Number 335746752, awarded to Linda P. Juang and Maja K. Schachner. We thank Mohini Lokhande, Tim Müller, and David Sherman for generously sharing their study materials and experiences. We thank our collaborators GÜseli Baysu, Karen Phalet, and Ruth Ditzmann as well as the students, teachers, and schools that participated in and supported the study.

Author Contributions LJ and MKS: conceptualized, designed, acquired funding for, and supervised the study. Material preparation, school and participant recruitment, and data collection were coordinated by MS and performed by all authors but primarily by MS, TA, DK, and HL. Teacher training was conducted primarily by HL. Data analyses were performed by TA with input from other authors. The first version of the full manuscript was drafted by LJ based on reports written by all authors, the methods and results were primarily edited by TA and DK, and all authors commented on and edited previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability statement The data that support the findings of this study are available on request from the corresponding author [L.J.]. The data are not publicly available as we do not have permission from the participants.

Declarations

Conflicts of interest The authors do not have any conflicts of interest related to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allison, P. (2013, February 13). What's the best R-squared for logistic regression? *Statistical Horizons*. <https://statisticalhorizons.com/r2logistic/>
- Antidiskriminierungsstelle des Bundes [Federal Anti-Discrimination Agency] (2013). Diskriminierung im Bildungsbereich und im Arbeitsleben [Discrimination in education and employment]. Berlin. https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/BT_Bericht/gemeinsamer_bericht_zweiter_2013.pdf?__blob=publicationFile&v=4
- Appel, M., Weber, S., & Kronberger, N. (2015). The influence of stereotype threat on immigrants: Review and meta-analysis. *Frontiers in Psychology*, *6*, 1–15. <https://doi.org/10.3389/fpsyg.2015.00900>
- Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*(1), 7–39. <https://doi.org/10.1080/19345747.2016.1232459>
- Bakadorova, O., & Raufelder, D. (2014). The mediating role of socio-motivational support in the association between individual school self-concept and achievement motivation amongst adolescent students. *European Journal of Psychology of Education*, *29*(3), 347–366. <https://doi.org/10.1007/s10212-013-0202-5>
- Baysu, G., & Phalet, K. (2019). The up- and downside of dual identity: Stereotype threat and minority performance. *Journal of Social Issues*, *75*(2), 568–591. <https://doi.org/10.1111/josi.12330>
- Berry, J. W., Phinney, J. S., Sam, D. L., & Vedder, P. (2006). *Immigrant youth in cultural transition: Acculturation, identity, and adaptation across national contexts*. Lawrence Erlbaum.
- Binning, K. R., Cook, J. E., Purdie-Greenaway, V., Garcia, J., Chen, S., Apfel, N., Sherman, D. K., & Cohen, G. L. (2019). Bolstering trust and reducing discipline incidents at a diverse middle school: How self-affirmation affects behavioral conduct during the transition to adolescence. *Journal of School Psychology*, *75*, 74–88. <https://doi.org/10.1016/j.jsp.2019.07.007>
- Bradley, D. N., Crawford, E. P., & Dahill-Brown, S. E. (2016). Defining and assessing FoI in a large-scale randomized trial: Core components of values affirmation. *Studies in Educational Evaluation*, *49*, 51–65. <https://doi.org/10.1016/j.stueduc.2016.04.002>
- Brenick, A., & Titzmann, P. F. (2015). Internationally comparative research on minority and immigrant adolescents' social, emotional, and behavioral development. *Journal of Adolescent Health*, *57*(6), 571–573. <https://doi.org/10.1016/j.jadohealth.2015.09.005>
- Celeste, L., Baysu, G., Phalet, K., & Brown, R. (2021). Self-affirmation and test performance in ethnically diverse schools: A new dual-identity affirmation intervention. *Journal of Social Issues*, *77*(3), 824–850. <https://doi.org/10.1111/josi.12418>
- Celeste, L., Baysu, G., Phalet, K., Meeussen, L., & Kende, J. (2019). Can school diversity policies reduce belonging and achievement gaps between minority and majority youth? Multiculturalism, color-blindness, and assimilationism assessed. *Personality and Social Psychology Bulletin*, *45*(11), 1603–1618. <https://doi.org/10.1177/0146167219838577>
- Civitillo, S., Mayer, A.-M., & Jugert, P. (2023). A systematic review and meta-analysis of the associations between perceived teacher-based racial-ethnic discrimination with student well-being and academic outcomes. *Manuscript under Review*. <https://psyarxiv.com/d4rva/>
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, *324*(5925), 400–403. <https://doi.org/10.1126/science.1170769>

- Cohen, G. L., Garica, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*(5791), 1307–1310. <https://doi.org/10.1126/science.1128317>
- Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*(1), 333–371. <https://doi.org/10.1146/annurev-psych-010213-115137>
- Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: Protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology*, *102*(3), 479–496. <https://doi.org/10.1037/a0026312>
- Creswell, J. D., Welch, W. T., Taylor, S. E., Sherman, D. K., Gruenewald, T. L., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science*, *16*(11), 846–851. <https://doi.org/10.1111/j.1467-9280.2005.01624.x>
- de Jong, E. M., Jellesma, F. C., Koomen, H. M. Y., & de Jong, P. F. (2016). A values-affirmation intervention does not benefit negatively stereotyped immigrant students in the Netherlands. *Frontiers in Psychology*, *7*(691), 1–20. <https://doi.org/10.3389/fpsyg.2016.00691>
- Easterbrook, M. J., Harris, P. R., & Sherman, D. K. (2021). Self-affirmation theory in educational contexts. *Journal of Social Issues*, *77*(3), 683–701. <https://doi.org/10.1111/josi.12459>
- Emmerich, M., & Hormel, U. (2013). Ungleichheit als Systemeffekt? [Inequality as a system effect?] In F. Dietrich, M. Heinrich, & N. Thieme (Eds.), *Bildungsgerechtigkeit Jenseits Von Chancengleichheit: Theoretische und Empirische Ergänzungen und Alternativen zu 'PISA' [Educational equity beyond equal opportunity: Theoretical and empirical extensions and alternatives to PISA]* (pp. 137–158). Springer Fachmedien Wiesbaden.
- Frankenberg, E., Kupper, K., Wagner, R., & Bongard, S. (2013). Immigrant youth in Germany: Psychological and sociocultural adaptation. *European Psychologist*, *18*(3), 158–168. <https://doi.org/10.1027/1016-9040/a000154>
- Froehlich, L., Mok, S. Y., Martiny, S. E., & Deaux, K. (2022). Stereotype threat-effects for Turkish-origin migrants in Germany: Taking stock of cumulative research evidence. *European Educational Research Journal*, *21*(2), 330–354. <https://doi.org/10.1177/1474904118807539>
- Fuchs, M., & Sixt, M. (2007). Zur Nachhaltigkeit von Bildungsaufstiegen. Soziale Vererbung von Bildungserfolgen über mehrere Generationen. [On the sustainability of educational improvement. Social heredity transmission of educational success over multiple generations]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie Cologne Journal for Sociology and Social Psychology*, *59*(1), 1–29. <https://doi.org/10.1007/s11577-007-0001-6>
- Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology*, *51*(9), 1226–1240. <https://doi.org/10.1037/dev0000028>
- Geiser, C. (2012). *Data analysis with Mplus*. Guilford Press.
- Glock, S., & Krolak-Schwerdt, S. (2013). Does nationality matter? The impact of stereotypical expectations on student teachers' judgments. *Social Psychology of Education*, *16*(1), 111–127. <https://doi.org/10.1007/s11218-012-9197-z>
- Götz, L., Lingel, K., & Schneider, W. (2013). DEMAT 6+. Deutscher Mathematiktest für sechste Klassen [German math test for 6th grade]. Hogrefe.
- Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 80–100. https://doi.org/10.1207/S15328007SEM1001_4
- Hadden, I. R., Easterbrook, M. J., Nieuwenhuis, M., Fox, K. J., & Dolan, P. (2020). Self-affirmation reduces the socioeconomic attainment gap in schools in England. *British Journal of Educational Psychology*, *90*(2), 517–536. <https://doi.org/10.1111/bjep.12291>
- Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education*, *87*(2), 106–124. <https://doi.org/10.1177/0038040714525970>
- Hanselman, P., Rozek, C. S., Grigg, J., & Borman, G. D. (2017). New evidence on self-affirmation effects and theorized sources of heterogeneity from large-scale replications. *Journal of Educational Psychology*, *109*(3), 405–424. <https://doi.org/10.1037/edu0000141>
- Jenkins, P. H. (1995). School delinquency and school commitment. *Sociology of Education*, *68*, 221–239. <https://doi.org/10.2307/2112686>
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. <https://doi.org/10.3102/0013189X20912798>

- Kunyu, D., Juang, L. P., Schachner, M. K., & Schwarzenthal, M. (2020). Discrimination among youth of immigrant descent in Germany. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 52(3–4), 88–102. <https://doi.org/10.1026/0049-8637/a000231>
- Lokhande, M., & Müller, T. (2019). Double jeopardy – Double remedy? The effectiveness of self-affirmation for improving doubly disadvantaged students' mathematical performance. *Journal of School Psychology*, 75, 58–73. <https://doi.org/10.1016/j.jsp.2019.07.006>
- Martiny, S. E., Mok, S. Y., Deaux, K., & Froehlich, L. (2014). Effects of activating negative stereotypes about Turkish-origin students on performance and identity management in German high schools. *Revue Internationale De Psychologie Sociale*, 27(3–4), 205–225.
- Miklikowska, M., Bohman, A., & Titzmann, P. F. (2019). Driven by context? The interrelated effects of parents, peers, classrooms on development of prejudice among Swedish majority adolescents. *Developmental Psychology*, 55(11), 2451–2463. <https://doi.org/10.1037/dev0000809>
- Moshagen, M., & Erdfelder, E. (2016). A new strategy for testing structural equation models. *Structural Equation Modeling*, 23(1), 54–60. <https://doi.org/10.1080/10705511.2014.950896>
- Müller, T., & Lokhande, M. (2017). *Wider die Stereotypisierung: Bessere Schulleistung durch Selbstbestätigung [Countering stereotype threat: Increasing achievement in school by self-affirmation]*. Vielfalt Im Klassenzimmer. Wie Lehrkräfte Gute Leistung Fördern Können [Diversity in the classroom. How teachers can promote good student achievements], (pp. 38–56). Berlin Institute for Integration and Migration Research (BIM) and Research Unit of the Expert Council of German Foundations on Integration and Migration (SVR Research Unit). www.svr-migration.de/publikationen/vielfalt-im-klassenzimmer.
- Nesterko, Y., & Glaesmer, H. (2019). Warum fragen wir nicht direkt nach?: Eine Analyse zur subjektiven Zuschreibung des Migrationshintergrundes [Why don't we ask directly?: An analysis of the subjective attribution of migration background]. *Psychologische Rundschau*, 70(2), 101–108. <https://doi.org/10.1026/0033-3042/a000399>
- Russell, S. T., Sinclair, K. O., Poterat, V. P., & Koenig, B. W. (2012). Adolescent health and harassment based on discriminatory bias. *American Journal of Public Health*, 102(3), 493–495. <https://doi.org/10.2105/AJPH.2011.300430>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Schachner, M.K., Van de Vijver, F., & Noack, P. (2018). Acculturation and school adjustment of early-adolescent immigrant boys and girls in Germany: Conditions in school, family, and ethnic group. *The Journal of Early Adolescence* 38(3) 352–384. <https://doi.org/10.1177/0272431616670991>
- Schachner, M. K., Schwarzenthal, M., van de Vijver, F. J. R., & Noack, P. (2019). How all students can belong and achieve: Effects of the cultural diversity climate amongst students of immigrant and nonimmigrant background in Germany. *Journal of Educational Psychology*, 111(4), 703–716. <https://doi.org/10.1037/edu0000303>
- Schachner, M. K., Schwarzenthal, M., Moffitt, U., Civitillo, S., & Juang, L. (2021). Capturing a nuanced picture of classroom cultural diversity climate: Multigroup and multilevel analyses among secondary school students in Germany. *Contemporary Educational Psychology*, 65, 101971. <https://doi.org/10.1016/j.cedpsych.2021.101971>
- Sherman, D. K., & Cohen, G. L. (2020). Self-affirmation: Understanding the effects. In J. Aronson & E. Aronson (Eds.), *Readings about the Social Animal* (12th ed., pp. 94–113). New York: Worth Publishers.
- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., & Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, 104(4), 591–618. <https://doi.org/10.1037/a0031495>
- Shewach, O. R., Sackett, P. R., & Quint, S. (2019). Stereotype threat effects in settings with features likely versus unlikely in operational test settings: A meta-analysis. *Journal of Applied Psychology*, 104(12), 1514–1534. <https://doi.org/10.1037/apl0000420>
- Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia, J., & Cohen, G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a key to buffering against identity threat. *Personality and Social Psychology Bulletin*, 39(5), 663–676. <https://doi.org/10.1177/0146167213480816>
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional

- participation in academic activities in the classroom. *Educational and Psychological Measurement*, 69(3), 493–525. <https://doi.org/10.1177/0013164408323233>
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67(1), 415–437. <https://doi.org/10.1146/annurev-psych-073115-103235>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). Academic Press.
- SVR-Forschungsbereich [SVR-Research Area]. (2016). Doppelt benachteiligt? Kinder und Jugendliche mit Migrationshintergrund im deutschen Bildungssystem. Eine Expertise im Auftrag der Stiftung Mercator [Doubly disadvantaged? Children and youth with a migration background in the German education system. A report commissioned by the Mercator Foundation]. www.svr-migration.de/wp-content/uploads/2017/07/SVR-FB_Doppelt_benachteiligt.pdf
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In W. Austin & S. Worchel (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7–24). Nelson-Hall.
- Teltemann, J., & Schunck, R. (2016). Education systems, school segregation, and second-generation immigrants' educational success: Evidence from a country-fixed effects approach using three waves of PISA. *Journal of Comparative Sociology*, 57, 401–424. <https://doi.org/10.1177/0020715216687348>
- Titzmann, P. F., Silbereisen, R. K., Mesch, G. S., & Schmitt-Rodermund, E. (2011). Migration-specific hassles among adolescent immigrants from the former Soviet Union in Germany and Israel. *Journal of Cross-Cultural Psychology*, 42(5), 777–794. <https://doi.org/10.1177/0022022110362756>
- Vedder, P., Sam, D. L., & Liebkind, K. (2007). The acculturation and adaptation of Turkish adolescents in North-Western Europe. *Applied Developmental Science*, 11(3), 126–136. <https://doi.org/10.1080/10888690701454617>
- Verkuyten, M., Thijs, J., & Gharaei, N. (2019). Discrimination and academic (dis)engagement of ethnic-racial minority students: A social identity threat perspective. *Social Psychology of Education*, 22(2), 267–290. <https://doi.org/10.1007/s11218-018-09476-0>
- Walton, G. M., & Yeager, D. S. (2020). Seed and soil: Psychological affordances in contexts help to explain where wise interventions succeed or fail. *Current Directions in Psychological Science*, 29(3), 219–226. <https://doi.org/10.1177/0963721420904453>
- Wu, Z., Spreckelsen, T. F., & Cohen, G. L. (2021). A meta-analysis of the effect of values affirmation on academic achievement. *Journal of Social Issues*, 77(3), 702–750. <https://doi.org/10.1111/josi.12415>
- Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, 81(2), 267–301. <https://doi.org/10.3102/0034654311405999>
- Zander, L., Webster, G. D., & Hannover, B. (2014). Better than me?! How adolescents with and without migration background in Germany perceive each others' performance in class. *Psychology*, 5(1), 62–69. <https://doi.org/10.4236/psych.2014.51011>
- Zimmermann, F., Schütte, K., Taskinen, P., & Köller, O. (2013). Reciprocal effects between adolescent externalizing problems and measures of achievement. *Journal of Educational Psychology*, 105(3), 747–761. <https://doi.org/10.1037/a0032793>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Linda P. Juang is a Professor in the Department of Inclusive Education at the University of Potsdam, Germany. Her work focuses on how experiences of immigration relate to minoritized adolescents' development and adjustment in school, family, and community contexts.

Maja Schachner is a Professor of Educational Psychology with an emphasis on Socialisation and Culture at Martin-Luther-University Halle-Wittenberg in Germany. Her main research interests include the

cultural diversity climate in schools, school-based cultural diversity-related interventions, and acculturation and adjustment of adolescents of immigrant descent. She has received funding from the German Research Foundation (DFG), the Jacobs Foundation, and the European Social Funds (ESF).

Tuğçe Aral is a Ph.D. student in the Department of Inclusive Education at the University of Potsdam, Germany. Her research focuses on ethnic-racial socialization in family and school contexts and their relations to identity development and intergroup relations.

Miriam Schwarzenthal is a junior professor of School Socialization Research in the School of Education at the University of Wuppertal in Germany. Her research is situated between the fields of psychology and education and focuses on socialization processes with regards to cultural diversity and social inequity in school, family, and peer contexts. Research topics include intercultural competence and critical consciousness among secondary school students and (preservice) teachers, as well as the school diversity climate.

David Kunyu is a Ph.D. student in the Department of Inclusive Education at the University of Potsdam, Germany. His research focuses on experiences of discrimination and academic and socioemotional adjustment among immigrant and refugee youth in Germany and Kenya.

Hanna Löhmannsröben is a Professor of Inclusive Education at the University of Potsdam, Germany. She is a theologian, former church Superintendent, and special education teacher.

Authors and Affiliations

Linda P. Juang¹  · Maja K. Schachner²  · Tuğçe Aral¹  ·
Miriam Schwarzenthal³  · David Kunyu¹  · Hanna Löhmannsröben¹

✉ Linda P. Juang
juang@uni-potsdam.de

Maja K. Schachner
maja.schachner@paedagogik.uni-halle.de

Tuğçe Aral
tugce.aral@uni-potsdam.de

Miriam Schwarzenthal
schwarzenthal@uni-wuppertal.de

David Kunyu
david.kunyu.khisoni@uni-potsdam.de

Hanna Löhmannsröben
hloeh@uni-potsdam.de

- ¹ Universität Potsdam, Inklusionspädagogik, 24-25 Karl-Liebknechtstrasse, 14476 Potsdam, Germany
- ² Institut Für Bildungsforschung, Martin Luther Universität Halle-Wittenberg, Franckeplatz 1, 06110 Halle, Germany
- ³ Institut Für Bildungsforschung, Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany