



The classroom social environment in mixed secondary school classes: The role of ethnic classroom composition and classmates' explicitly and implicitly measured ethnic attitudes

Lian van Vemde¹ · Jochem Thijs¹ · Lisette Hornstra²

Received: 14 June 2022 / Accepted: 6 December 2022 / Published online: 19 January 2023
© The Author(s) 2023

Abstract

The ethnic classroom composition and classmates' ethnic attitudes can affect how students experience their classroom social environment (CSE). Following the imbalance of power thesis and prior research on ethnic attitudes, this cross-sectional study examined if ethnic classroom composition (i.e., proportion of in-group and Herfindahl Index) and classmates' explicitly and implicitly measured ethnic attitudes predicted secondary school students' ($M_{age} = 13.31$ years; 58.1% female) classroom belonging, popularity and likability, classroom cohesion and conflict in mixed classes in the Netherlands. Differences between non-ethnic Dutch ($n = 248$) versus ethnic Dutch students ($n = 141$) were examined as well. Ethnic Dutch students report an overall more negative CSE than their non-ethnic Dutch classmates. Multilevel analyses indicated that a higher proportion of in-group peers affected non-ethnic Dutch students' popularity and likability negatively. Moreover, classmates' explicitly measured ethnic attitudes were predictive of student popularity while classmates' implicitly measured ethnic attitudes were predictive of student likability. Finally, classmates implicitly measured ethnic attitudes moderated the effect of proportion in-group peers on students' shared experience of classroom belonging. These results show that promoting classroom diversity is not enough to create a positive CSE for all students. Classmates' ethnic attitudes are also important to consider.

Keywords Ethnic classroom composition · Explicitly measured attitudes · Implicit association test · Peer relations · Classroom climate

✉ Lian van Vemde
l.vanvemde@uu.nl

¹ Department of Interdisciplinary Social Science, Utrecht University, Padualaan 14, 3584CH Utrecht, The Netherlands

² Department of Education, Utrecht University, Heidelberglaan 1, 3584CS Utrecht, The Netherlands

1 Introduction

With societies becoming increasingly diverse, there has been a plea for implementing policies to counter school segregation and create ethnically diverse classrooms (Peters & Walraven, 2011). Classrooms can be seen as micro-societies that represent society at large (Dewey, 1966), and social cohesion in diverse classrooms may transfer to more cohesion in society later on. The underlying assumption is that when students with different ethnic backgrounds attend the same classroom, they will probably develop positive interethnic relationships. However, in diverse classrooms, social relationships are not always harmonious and there can be misunderstandings and conflict between students due to differences in ethnic background. Research has found, for example, that both ethnic majority and minority students report more (ethnic) peer victimization in ethnically heterogeneous classrooms and school settings (Durkin et al., 2012; Vervoort et al., 2010). Such findings may be explained by the existing ethnic attitudes in the classroom (Thijs et al., 2014). That is, when students hold negative attitudes toward other ethnic groups, there may be more misunderstandings and conflict between students from different groups. However, even though there have been many studies on the effect of intergroup contact on students' ethnic attitudes (cf. Pettigrew & Tropp, 2006), there is little research on how classmates' ethnic attitudes are related to students' experiences of their (ethnically mixed) classroom social environment. The present study examines how ethnic classroom composition together with classmates' ethnic attitudes affect perceived classroom climate and peer relationships among Dutch secondary school students with diverse backgrounds. It goes beyond previous research by incorporating both implicit (reaction time) and explicit (questionnaires) measures of students' ethnic attitudes. This may lead to a better understanding of how ethnic attitudes affect the classroom social environment because both measures tend to predict different types of behaviors and outcomes (Greenwald et al., 2009; Hahn & Gawronski, 2017; Kurdi et al., 2019). Thereby, this study aims to provide insights into factors that help to create a positive social environment in diverse classes. Ultimately, such knowledge can help design more tailored interventions to create a positive classroom social environment for all students.

1.1 Students' classroom social environment

Experiencing a positive classroom social environment is vital for students' academic achievement and motivation (e.g., Skinner & Belmont, 1993), engagement (e.g., Patrick et al., 2007), well-being (e.g., Wang et al., 2020), and socio-emotional adjustment (e.g., Sakiz et al., 2012; Stewart, 2016), and can help to prevent early dropout (e.g., McMahan et al., 2009). The classroom social environment (CSE) consists of multiple interrelated aspects: Individual aspects referring to students' personal position in class and the quality of their peer relationships (i.e., students' likability, popularity, and sense of belonging), and collective aspects that pertain to the classroom as a whole and its climate in general (i.e., classroom cohesion and conflict; Patrick et al., 2007, 2011; Ryan & Patrick, 2001).

Students' likability and popularity are two individual aspects of the CSE and can be examined using peer nominations that indicate who is popular or well-liked by the other students in the classroom (Graham & Echols, 2018). *Likability* refers to acceptance while *popularity* connotes power, prestige, and visibility of the student. Thus, someone can be nominated as popular, but not necessarily as likable (Cillesen & Marks, 2011; Graham & Echols, 2018). In addition, students' *sense of belonging* reflects their personal perception of their relationships with classmates and their feelings that they are included, accepted, supported, and respected by them (Juvonen, 2006).

Whereas classroom positions and peer relations can differ between individual students within the same class, classrooms are also collective units. Collective aspects of a classroom are important to consider as students in the same classrooms are not only a set of individuals, but also a group. The overall quality of the social relations within this unit can be referred to as the classroom climate (Fraser et al., 1982). Two main dimensions of this climate are cohesion and conflict (Fraser et al., 1982). *Cohesion* reflects the degree to which classroom members are connected to each other and the extent to which there are positive relationships in the class (Wilson et al., 2011). *Conflict* refers to the amount of friction and tension between students, and whether students often fight with each other (Fraser et al., 1982). Both dimensions tend to be negatively related to each other (McMahon et al., 2009). Yet they are conceptually independent, as the absence of conflict does not guarantee a cohesive classroom.

The collective aspects cohesion and conflict are typically assessed via students' perceptions, and these subjective perceptions can be aggregated across children in the classroom to obtain a shared and relatively objective measure of classroom climate (see Lüdtke et al., 2009). However, complete consensus among classmates is very unlikely, which means that, although one part of their subjective perceptions of cohesion and conflict may be shared, another part will be not. Conversely, although students' aforementioned sense of belonging refers to their individual experiences of the CSE, perceptions of belonging can be shared among classmates as well. Most research on the CSE has not distinguished between students' shared perceptions and their unique perceptions (Lüdtke et al., 2009; Marsh et al., 2012; Thijs & Verkuyten, 2016). However, the effects of shared classroom perceptions may differ from the effects of unique classroom perceptions (Marsh et al., 2012), making it important to study both of them separately and to disentangle effects at the *within* and *between* classroom level.

1.2 Ethnic classroom composition and CSE experiences

Ethnic classroom composition can be operationalized in different ways. On the one hand, studies have used proportion measures for the presence of specific groups in the classroom, such as the proportion of minority students (e.g., Hornstra et al., 2015; Jackson et al., 2006; Thijs et al., 2014), or the proportion of students with the same ethnicity as the participants (proportion in-group students: e.g., Bubritzki et al., 2018; Madsen et al., 2016; Wilson & Rodkin, 2011). On the other hand,

studies have also examined ethnic classroom composition through general diversity measures that take into account the number of different groups together with their relative proportions, such as the Herfindahl index (e.g., Bubritzki et al., 2018; Hirschman, 1964), which is also known as the Simpson index (e.g., Benner & Graham, 2009; Juvonen et al., 2006; Simpson, 1949). The Herfindahl index is a standardized measure which indicates the likelihood that any two individuals randomly selected from the classroom will be from the same category (Putnam, 2007). It captures both the number of different ethnic groups in a setting as well as the relative representation of each group (Bayram Özdemir et al., 2018; Benner & Graham, 2009). Whereas the proportion of in- or out-group classmates can differ within a class depending on students' own ethnicity, diversity measures like Herfindahl's and Simpson's are the same for all students. Thus, ethnic classroom composition can be investigated at the *within* and *between* class level.

Existing research suggests that the proportion of in-group classmates has mixed but generally positive effects on students' social position and classroom belonging. It has been found, for example, that African-American youth were considered more popular, cool, and likable when they had more in-group classmates (Jackson et al., 2006; Rock et al., 2011; Wilson & Rodkin, 2011). A reason for this could be that students tend to nominate peers with the same ethnic background more favorably than peers with a different ethnic background (Jackson et al., 2006). Likewise, ethnic majority students in Dutch classrooms were found to be more popular when they had more in-group peers, while ethnic minority students were less popular when they had more in-group peers (Stevens et al., 2020). Conversely, studies have also shown that students report more victimization when they have less in-group peers in their classroom (e.g., Agirdag et al., 2011; Felix & You, 2011; Graham & Juvonen, 2002; Hoglund & Hosan, 2013). Furthermore, other studies found that both ethnic majority (Rjosk et al., 2017) and ethnic minority students (Benner & Graham, 2009; Hornstra et al., 2015) experienced more classroom belonging when they had a larger share of same-ethnic peers in the classroom (but see Peetsma et al., (2006) for non-significant results). Also, having a larger share of same-ethnic classmates reduces minority adolescents' chances of feeling lonely in school (Madsen et al., 2016), their experience of social exclusion, and their risk of isolation (Madsen et al., 2016; Plenty & Jonsson, 2017). This is in line with the notion that it can be easier to connect with in-group as compared to out-group others (Baumeister & Leary, 1995; Benner & Crosnoe, 2011; Rjosk et al., 2017). Taken together, these findings are consistent with the idea that an ethnically congruent learning environment (i.e., having more classmates "like me") helps students to be and feel connected to their classrooms and peers (Benner & Graham, 2009; Graham & Echols, 2018).

Studies using general diversity measures like Herfindahl's or Simpson's are scarce. However, such general *between* class measures are needed to understand how classroom composition may be related to shared aspects of the CSE. The available research seems to support the so-called *imbalance of power thesis* (Graham, 2006; Juvonen et al., 2006), the idea that ethnic diversity balances power distributions between groups, which reduces incidents of peer harassment and improves overall relations. For instance, Graham (2006) found that in classrooms with more ethnic diversity, students reported less victimization by peers and less loneliness.

Moreover, minority students have been found to feel more safe and less lonely when the classroom was more ethnically diverse (Juvonen et al., 2006).

1.3 Ethnic attitudes

As mentioned, ethnically diverse classrooms may be beneficial for students' shared CSE experiences, but this probably depends on the ethnic attitudes in the classroom. For example, the *imbalance of power thesis* (Graham, 2006; Juvonen et al., 2006) seems to assume that students of different ethnic groups tend to be biased against each other (which makes it important to have an equal representation of these groups in the classroom). This suggests that less diverse classes may not be conducive to students' CSE. However, this assumption is not often tested directly. Students' ethnic attitudes involve their evaluations of certain ethnic groups which can be associated with their interpersonal behavior toward members of these groups (e.g., Aboud et al., 2003; Hamm et al., 2005; Wagner et al., 2008). Thus, students' social experiences in the classroom may be affected by the attitudes that others in their classroom hold about their ethnic groups. To date, very few studies have examined the impact of ethnic classroom composition in combination with the existing ethnic attitudes in the classroom, but their findings suggest there is at least a negative correlation between quality of contact and ethnic attitudes (König et al., 2022) and that both factors interact (e.g., Bayram Özdemir et al., 2018; Thijs et al., 2014). Therefore, the present study also examined whether the ethnic attitudes of students' classmates predicted their own CSE experiences and interacted with ethnic composition in doing so.

1.3.1 Measuring students' ethnic attitudes

Many studies on ethnic relations in classroom or school settings have used self-report measures to assess students' evaluations of different ethnic groups (e.g., thermometer scales, Bubritzki et al., 2018; smiley faces, Jugert et al., 2011; general perceptions of immigrants, Bayram Özdemir et al., 2018). These are also referred to as explicit attitude measures (e.g., Bayram Özdemir et al., 2018; Bubritzki et al., 2018; Jugert et al., 2011; Özdemir & Bayram Özdemir, 2017). Explicit measures are easy to administer and give students the opportunity and motivation to engage in deliberate processing, and are therefore mostly predictive of deliberate and consciously monitored behaviors, such as verbal behaviors (Azjen et al., 2018; Dovidio et al., 1996; Gawronski & Creighton, 2013).

Implicit instruments, such as the Implicit Association Test (IAT; Greenwald et al., 1998) try to capture automatic or implicit processes that underlie the effect of attitudes on behavior (Greenwald & Lai, 2020). Therefore implicit attitude measures, contrary to explicit measures, are especially predictive of behaviors that are more difficult to control, such as nonverbal behaviors or facial expressions (Azjen et al., 2018; Dovidio et al., 1996). Implicit measures limit participants' opportunity to engage in effortful processing, which reduces the social desirability of the answers. This may be why implicit attitude measures are in general somewhat more

predictive of behaviors in (socially) sensitive domains, like prejudice and discrimination than explicit measures (Greenwald & Lai, 2020). Since both types of attitude measures tend to be predictive of different types of behavior (i.e., deliberate versus spontaneous; Azjen et al., 2018; Dovidio et al., 1996) that can affect students' interpersonal relations, each of them might explain unique variance in different aspects of students' CSE experiences. Therefore, the present study incorporates both types of attitude measures. To date, most studies using implicit attitude measures have been performed in (experimental) laboratory settings with undergraduate students and focused on how attitudes affect judgements and behavior of the person holding them. Additionally, very few studies examined the effects of ethnic attitudes in real-life settings, or studied the effects of the attitude on the target of the attitude (i.e., members of the ethnic groups involved) rather than effects on the behavior of the attitude holder (Hahn & Gawronski, 2017; Kurdi et al., 2019; Madva & Brownstein, 2018). The present study aims to unveil whether findings of previous studies transfer to youngsters in actual classroom situations and how the attitudes affect the experiences of the attitude receiver.

1.4 Classmates' ethnic attitudes and CSE experiences

Although many studies have examined how students' *own* interethnic attitudes affect their own behaviors and evaluations (e.g., Binder et al., 2009; Jugert et al., 2011; Thijs, 2017), there are only a few studies that focused on the impact of attitudes of *others* (e.g., Thijs et al., 2014) on students' experiences of the CSE. Nevertheless, the former studies provide useful insights in the potential effects of classmates' attitudes on students' CSE. For example, regarding peer relations, research has found that explicitly measured negative ethnic attitudes predicted a stronger preference for same-ethnic friendships amongst adolescents (Binder et al., 2009; Hamm et al., 2005) and children (Aboud et al., 2003; Jugert et al., 2011), and thus less positive relationships between students of different ethnic groups. Moreover, higher explicitly measured prejudicial attitudes of German adults (Wagner et al., 2008) and Swedish youth (Bayram Özdemir et al., 2016, 2018) led them to engage in discriminatory behavior toward and ethnic harassment of non-native others. Thus, negative ethnic attitudes can not only form a barrier to the establishment of positive interethnic relations, but also stimulate negative behaviors toward ethnic out-groups (Bayram Özdemir et al., 2016). Furthermore, ethnic harassment is more likely when one's group is a numerical minority (Bayram Özdemir et al., 2018). This could indicate that having more in-group classmates could serve as a protective factor against the negative attitudes of the out-group peers and thus that having more in-group classmates would be especially beneficial for students' CSE experiences when classmates' attitudes towards the students' group are more negative.

Besides the attitudes of classmates toward the students' in-group, the classroom-average ethnic attitude may also be predictive of students' CSE. For example, previous research found that in classrooms where students are generally more biased toward their in-group, discriminatory behaviors toward out-group members may occur more often (Salmivalli et al., 1996; Thijs et al., 2014), while in classes where

peers have a strong anti-bias norm, students may be less likely to show discriminatory behavior (see Rutland et al., 2005). This is in line with the *imbalance of power thesis* (cf. Graham, 2006), which seems to assume that students tend to be more biased in favor of their own group. Therefore, diversity would be more beneficial for students' CSE experiences classrooms with a strong average degree of in-group bias.

1.5 The present study

The goal of the present study was to examine how classroom ethnic composition and classmates' ethnic attitudes were associated with both collective (i.e., classroom climate) and individual aspects (i.e., peer relations) of the CSE of secondary school students in ethnically diverse classrooms in the Netherlands. In doing so, the present study focuses on ethnic Dutch students and students from the largest ethnic minority groups in the Netherlands, that is, students from Turkish, Moroccan, Surinamese, and Antillean descent (Statistics Netherlands, 2016).

Based on the literature discussed so far, we formulated eight hypotheses to be evaluated in the present study. Regarding classroom diversity at the within classroom level, it was hypothesized that students with relatively more ethnic in-group classmates reported a stronger sense of belonging, and were considered more popular and likable by their peers (Hypothesis 1) and that students with more in-group classmates would perceive more cohesion and less conflict than their classmates (Hypothesis 2). In addition, at the between classroom level (involving students' shared perceptions), we hypothesized that more classroom diversity would be associated with less conflict, but also with more cohesion and a stronger shared sense of belonging (Hypothesis 3). With regards to classmates' ethnic attitudes, at the within classroom level, we hypothesized that students experienced more belonging and were considered more popular or likable when their classmates had more positive attitudes toward the *students'* ethnic group (Hypothesis 4) and that students whose classmates held less positive attitudes toward the *students'* in-group experienced less cohesion and more conflict (Hypothesis 5). Furthermore, we hypothesized that the effect of the presence of in-group classmates on students' CSE experiences would be stronger when classmates' attitudes toward the in-group were less positive (Hypothesis 6). In addition, at the between classroom level it was hypothesized that a stronger average in-group bias (a more positive evaluation of the in-group versus the out-group) would be associated with the collective experience of more classroom conflict, less cohesion, and less classroom belonging (Hypothesis 7). Lastly, we hypothesized that the anticipated effects of classroom diversity were stronger in classrooms with a stronger average degree of in-group bias (Hypothesis 8). The proposed model is depicted in Fig. 1.

In evaluating these hypotheses, we included student socio-economic status (SES) as a covariate, as it is often confounded with ethnicity (Kalter et al., 2018). Thereby, we excluded the possibility that differences in CSE experiences are the result of SES rather than ethnic background. We further controlled for students' age and gender. Finally, we explored whether the effects differed for students with a non-ethnic Dutch versus an ethnic Dutch background.

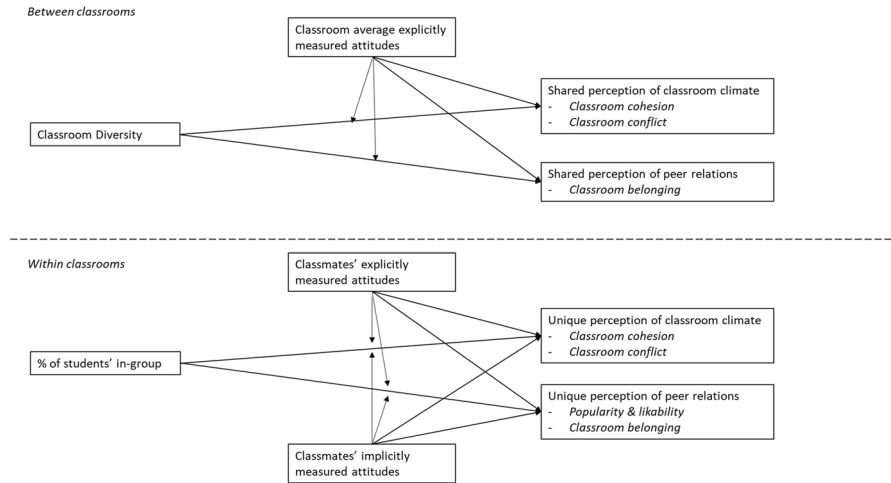


Fig. 1 Proposed model of the relationships between the collective and individual aspects of the CSE, ethnic classroom composition and classmates' ethnic attitudes

2 Methods

2.1 Procedure

Data were collected in February 2014 at three secondary schools in the Netherlands. First and second year students were recruited through their schools and teachers, and parents and students were informed about the procedure of the study. Parents and students provided passive informed consent. Only five students (0.9%) did not participate as their parents did not give consent. The present study was approved by the institutional ethical review board of the University of Amsterdam (the Netherlands).

Students filled out a questionnaire and completed two Implicit Association Tests (IAT) on a laptop during regular class hours. This took about one hour. All students first completed an IAT on their implicit attitude toward people with a white versus dark skin color and then completed an IAT on their implicit attitude toward Turkish and Moroccan children versus children of Dutch origin. The primary goal of the present study was to examine to what extent attitude differences between participants were predictive of certain outcomes. As counterbalancing is only recommended when the goal is to report on the average degree of bias of the group as a whole (i.e., the overall mean scores), we did not employ counterbalancing of the blocks of the IATs (Greenwald et al., 2022). Hence, all blocks within the IAT were administered in the same order to all students. Upon completion of both IATs, students filled out an online questionnaire measuring their explicit ethnic attitudes, their perception on classroom climate and conflict, their sense of classroom belonging and nominated peers in terms of popularity and likability.

2.2 Participants

The original sample consisted of 535 students in the first and second year of secondary school. They were from three different schools located in urban areas of the Netherlands, and from 25 classes, with an average of 21 participating students ($SD=4.4$ students) per classroom. One school participated with two classes, one with 11 classes, and one with 12 classes.

Students were classified into ethnic groups based on self-reports of their family background, following the guidelines of Statistics Netherlands (2016). For students to be classified as ethnic Dutch, both the mother's and father's family background should be Dutch. When one of the parents' family background was Dutch and the others' was of a different ethnicity, students were classified as a member of that other ethnic group. For students whose mother's and father's ethnicity was non-Dutch, the mothers' ethnicity was used to classify the student. When one of the parent's ethnicity was missing and the other was Dutch or other-ethnic, students' were classified as Dutch ($n=6$) or as a member of that other ethnic group ($n=4$). When both the mother's and father's ethnicities were missing ($n=12$), the mothers' country of birth was used. If that information was also missing ($n=7$), the fathers' country of birth was used to classify the student. In the end, information on students' background was missing for seven students, due to missing information about the family background and parents' country of birth. These students' were excluded from the analyses.

After this categorization, 26.7% of the students were identified as ethnic Dutch, 6.4% as Turkish, 13.8% as Moroccan, 23.7% as Surinamese, 3.0% as Antillean, and 26.4% as "other" (e.g., Brazilian, Chinese, Ghanaian, Indonesian, or Pakistani). Most of these ethnic minority students were born in the Netherlands (89.7%). Across the classes, the total percentage non-ethnic Dutch students varied from 23.8% to 90.5% ($M=71.3\%$, $SD=15.9$). For the final sample, 139 students whose ethnic background was categorized as "other" were excluded, because there was only information available for Turkish, Moroccan, Surinamese, Antillean, or ethnic Dutch groups regarding classmates' explicitly measured attitudes (see Measures). However, these other students were included in the calculation of the classroom composition measures, in the calculation of the class-level measures of the CSE (see Measures for a more detailed explanation), and they were included as classmates in the classmates' attitude measures, as they are still part of the classroom, are holders of attitudes, and share CSE experiences with other students.

The final sample consisted of 389 students (58.1% female) in 25 classes. Their mean age was 13.31 years ($SD=0.79$ years). Students in the ethnic minority groups (Turkish, Moroccan, Surinamese, and Antillean students) were combined into two overarching categories (Turkish/Moroccan students, $N=107$, and Surinamese/Antillean students, $N=141$) as the four ethnic groups were too small to include separately¹ (respectively, $n=34$, $n=73$, $n=125$, and $n=16$).

¹ The results of a MANOVA with post-hoc comparisons with Tukey correction showed that Turkish and Moroccan students did not significantly differ in their age, gender, and SES (p -values all $>.05$) and that Surinamese and Antillean students did not significantly differ from one another in their age, gender, and SES (p -values all $>.05$).

2.3 Measures

2.3.1 Implicit attitudes

Students implicit ethnic attitudes were measured using two IATs (Greenwald et al., 1998). These response latency measures were administered on a laptop, using the Inquisit Web software (Millisecond Software, 2021). The first IAT was a *race IAT* (Greenwald et al., 2003) and measured the relative strength of the association between skin color (i.e., black versus white using pictures of black and white faces) and the valence of words (i.e., positive versus negative connotations of words). For the present study, the IAT was translated into Dutch and the instruction and positive and negative words were adapted to the sample. For example, the word “glory” was changed into “happy” and the word “horror” into “evil”. The second IAT measured the relative strength of the association between *ethnicity* (i.e., Turkish or Moroccan versus Dutch using names representing these ethnicities) and the valence of words. This IAT was successfully used in a previous Dutch study by van den Bergh et al. (2010). For a detailed description of the IATs, see Online Resource 1.

For the *race IAT*, the average percentage of correct trials was 90.8%, and for the *ethnicity IAT*, the average percentage correct was 88.3%. No participants needed to be excluded due to extremely high error rates. The raw data were transformed using the improved scoring algorithm as proposed by Greenwald et al. (2003).² The standardized score (*D*) was then used as the indicator of students’ implicit attitudes. For the *race IAT*, a positive *D* score indicated a positive attitude toward people with a dark skin color relative to people with a white skin color and a negative *D* score indicated a positive attitude toward people with a white skin color relative to people with a dark skin color. For the *ethnicity IAT*, a positive *D* score indicated a positive attitude toward Turkish/Moroccan children relative to Dutch-origin children. A negative *D* score indicated a negative attitude toward Turkish/Moroccan children relative to Dutch-origin children.

Finally, as the focus of the present study was on *classmates’* attitudes, the *D* score was transformed to create a measure reflecting classmates’ attitudes at the *within classroom* level: *classmates’ implicit attitudes*. This variable was calculated by aggregating the IAT scores toward specific racial/ethnic groups at the class level and correcting this score for students’ own IAT score. This newly calculated variable represented the attitudes of classmates toward the students’ own group (e.g., for Turkish and Moroccan students, the score reflected the average attitude of classmates toward Turkish/Moroccan versus Dutch people). While computing this measure, the *race IAT* was only used for students with a Surinamese or Antillean background as they more often experience discrimination based on skin color in the Netherlands (Andriesen et al., 2020). The *ethnicity IAT* was only used for students with a Turkish and Moroccan background as they were the target groups in this IAT.

² This means that latencies greater than 10,000 ms were removed from analysis. Next, the inclusive standard deviation was computed for all trials in Block 3 and 5. The difference between the mean latency for Block 3 and Block 5 was then divided by the inclusive standard deviation.

For students with an ethnic Dutch background the average scores on both IATs was used. Because the IATs targeted specific racial or ethnic groups a class-average score would be meaningless, and therefore we did not compute a *between classroom* level measure.

2.3.2 Explicit attitudes

Students' explicit ethnic attitudes were assessed based on earlier research by Bakker et al. (2007). That is, students were asked to indicate to what extent they would like to be friends with someone from, respectively, an ethnic Dutch, Turkish, Moroccan, Surinamese, and Antillean background. Students answered the items on a 5-point Likert scale ranging from 1 (*Not at all*) to 5 (*Very much*). We also added an additional item to measure students' in-group preference: Students indicated whether they would rather be friends with someone from their own ethnic background compared to someone with a different ethnic background. This item was also answered on a 5-point Likert scale ranging from 1 (*Not at all*) to 5 (*Very much*). Based on these questions, two types of explicitly measured ethnic attitudes were derived. The first measure was a measure at the *within classroom* level: *classmates' explicit attitudes* pertained to the individual student and indicated how classmates evaluated the students' ethnic in-group. For example, for a Moroccan student, this score indicated how their classmates evaluated Moroccan people. A higher score on this new measure indicated a more positive attitude of classmates toward the in-group. The second measure, *explicit classroom in-group bias* represented the average in-group bias of the class and thus was a *between classroom* level measure. This measure was calculated by aggregating the in-group preference scores across all students in the classroom. A higher score indicated that students on average had a stronger preference for being friends with someone from their own as compared to a different ethnic group. Note that, although students with an ethnicity other than our specific target groups were not included as targets of the attitudes at the within classroom level (Level 1), they did report their attitudes toward the target groups as well as their in-group preference. Hence, we used our full sample ($N=535$) to compute our explicit attitude measures.

2.3.3 Classroom climate

Classroom cohesion and *classroom conflict* were measured with the *cohesion* and *friction* scales from the My Class Inventory (MCI; Fisher & Fraser, 1981), which both consist of five items. Example items for cohesion are "Students in this classroom see each other as friends" and "All students in this classroom like each other". For conflict, example items are "The students in this classroom are always fighting with each other" and "Students in this classroom often have conflicts with each other". The items were answered on a 5-point Likert scale ranging from 1 (*Not true at all*) to 5 (*Completely true*). Principal component analysis showed that the cohesion and conflict items loaded on two different factors which explained 57.6% of the variance. The internal consistency for cohesion was sufficient ($\alpha=0.77$) and good ($\alpha=0.84$) for conflict. To create between class level

measures, the scores were aggregated at the classroom level using the scores of the full sample ($N=535$).

2.3.4 Peer relations

2.3.4.1 Popularity and likability Students' popularity and likability was indicated by how often a student was nominated as popular or likable by their peers (van der Linden et al., 2010). The sociometric questions were "Which classmates are the most popular?" and "With whom of your classmates would you like to be friends?". Students could nominate up to five students per question. To correct for the number of students in the classroom and the total amount of nominations given by each student, students' popularity and likability were divided by the total amount of nominations given in a classroom. Hence, the resulting score is a proportion of the number of nominations the student has received relative to the total number of nominations given in a class.

2.3.4.2 Sense of classroom belonging Student's classroom belonging was measured using the *well-being at school with classmates* scale (Peetsma et al., 2001). The scale consisted of six items. Example items are "I have a lot of contact with my classmates" and "I sometimes feel alone in my classroom" (reverse coded). Answers were given on a 5-point Likert scale ranging from 1 (*Not true at all*) to 5 (*Completely true*). Principal component analysis showed that all items loaded on one factor which explained 52.9% of the variance. The internal consistency of the scale was good ($\alpha=0.82$). To reflect classroom belonging at the between class level, the scores were aggregated at the classroom level using the scores of the full sample ($N=535$).

2.3.5 Ethnic classroom composition

2.3.5.1 Proportion in-group classmates The proportion of in-group classmates was calculated by first determining students' own ethnicity as described under *Participants*. Next, the number of students from the participants' own ethnic group was divided by the total number of classmates.

2.3.5.2 Classroom diversity The overall degree of ethnic classroom diversity was determined with the reversed Herfindahl Index (Putnam, 2007; Sincer et al., 2021). The Herfindahl Index takes into account the number and size of different ethnic groups, using the following formula: $(\text{proportion ethnic group } 1)^2 + (\text{proportion ethnic group } 2)^2 + \dots + (\text{proportion ethnic group } n)^2$. The index was subtracted from 1 to indicate the degree of heterogeneity. A higher index score represented a more heterogeneous and balanced classroom. The average diversity index of the classes in the present study was 0.74 ($SD=0.10$, range 0.40 to 0.89), indicating that in general classes were more heterogeneous than homogeneous.

2.3.6 Socio-economic status

A proxy for students' socio-economic status (SES) was included as a covariate at the within student level in the present study. This SES proxy was calculated based on students' zip code (e.g., van Leest et al., 2021). The four digits of the students' zip code were transformed into a status score that indicates the social status of a particular neighborhood based on its inhabitants' education, income, and position on the labor market. Higher scores indicate a higher SES. At the time of the study, the average status score in the Netherlands was 0.17 (Knol et al., 2012). The average status score of the students participating in the present study was -0.56, indicating that their SES was below the national average.

2.4 Data analysis

To test our hypotheses we estimated a set of multilevel linear regression models for each dependent variable (i.e., classroom cohesion, conflict, classroom belonging, popularity, and likability) in Mplus (Version 8.6; Muthén & Muthén, 2021). Class was used as the cluster variable. All Level 1 variables were group-mean centered (based on the final sample including only the 389 students with our targeted ethnic background) and all Level 2 variables were grand-mean centered. Intraclass correlations were calculated to check the amount of variance at the group level (ICC1) and the reliability of the class-mean rating (ICC2; Lüdtke et al., 2009).

In the first step of all regression models, we specified the direct effects of the classroom composition measures and the attitude measures at both Level 1 and Level 2. Only explicit attitude measures were entered as a Level 2 predictor. Classroom cohesion and conflict, and students' sense of classroom belonging served as dependent variables with variance at both Level 1 and 2. Students' popularity and likability served as dependent variables at Level 1 only, given that these are student-specific and had almost no variance at Level 2. In the second step, the interactions between the classroom composition measures and attitude measures were entered.

The tested models were saturated and therefore model fit indices could not be compared and are not reported. Standardized beta's (b^*) were used as a measure for effect size. A value of 0.1 corresponds to a weak effect, 0.3 to a moderate effect, and 0.5 to a strong effect (Cohen, 1988). The assumptions for a multilevel regression (i.e., linearity, normality, homoscedasticity, and no multicollinearity) were checked and (approximately) met for all variables. Missing value analysis indicated that students' SES had the most missing values (19.3%). For all other variables, less than 2% of the data was missing (range 0.0%—1.3%). All missing values were located at the student level, so no classes were excluded due to missing data. Little's MCAR test was not significant, $\chi^2 = 41.10$, $df = 38$, $p = .337$, suggesting that data was missing completely at random. Hence, in addition to the MLR estimation, which uses full information maximum likelihood to handle missing data on the dependent variables, the cases with missing values on the independent variables were excluded by Mplus from the analysis.

3 Results

3.1 Preliminary analyses

Descriptive statistics for the total sample, as well as for the separate ethnic groups are depicted in Table 1. The ICC1s for classroom conflict, cohesion, and belonging were all above 0.10, indicating meaningful differences between classes (LeBreton & Senter, 2008). Moreover, the high ICC2s for these variables (> 0.70) indicate that the classroom aggregates were reliable, as an ICC2 above 0.40 is sufficient for reliably assessing group-level means (Fleiss, 1986). Table 2 shows the correlations at the within (top panel) and between (bottom panel) class level as well as the descriptive statistics for the final sample.

3.2 Predicting students' CSE experiences

Prior to testing the hypotheses, we examined which covariates should be taken into account. It was examined, using a multilevel regression, whether age, gender, and SES predicted students' CSE experiences at the *within classroom* level. Results showed that students' gender significantly predicted students' likability ($b^* = 0.18$, $p < .001$, indicating a weak to moderate effect) and that students' age significantly predicted students' experience of conflict ($*b = -0.08$, $p = .033$, indicating a weak effect). It was therefore decided to include gender as a covariate in the models for student's likability and age in the models for conflict. Moreover, students' ethnicity was included as a covariate in all analyses to account for the mean-level differences between these groups in the outcome variables (see Table 1).

3.2.1 Within classroom level

The standardized results for classroom belonging, popularity, and likability can be found in Table 3 (Model 1) and for classroom cohesion and conflict in Table 4 (Model 1). The hypotheses that students would experience a more positive CSE when they had more in-group classmates (Hypotheses 1–2) were not supported. Instead, the effect of proportion in-group classmates was negative for popularity and likability (respectively, $b^* = -0.17$, $p = .017$; and $b^* = -0.11$, $p = .044$; both indicating a weak effect). Thus, students were considered less popular or likable when they had more same-ethnic peers. Moreover, Hypotheses 4 and 5 were only partially supported. Classmates' explicitly measured ethnic attitudes were significantly and positively related to students' popularity ($b^* = 0.24$, $p < .001$, indicating a weak to moderate effect) and classmates' implicitly measured ethnic attitudes were significantly and positively related to students' likability ($b^* = 0.24$, $p = .003$, indicating a weak to moderate effect). This means that students were nominated more often as popular when their classmates explicitly evaluated their group more positively and that students were nominated more often as likable when their classmates were biased

Table 1 Descriptive statistics of the raw data of the present study for the total sample and the ethnic groups separately

	Total Group		ICC1	ICC2	Ethnic Dutch Students (n=139)		Turkish/Moroccan Students (n=105)		Surinamese/Antillean students (n=139)	
	N	M			SD	M	SD	M	SD	M
Explicit attitude toward Ethnic Dutch people	535	3.71	0.81	0.03	0.39	3.994,04 _a	3.70 _b	0.80	3.60 _b	0.74
Explicit attitude toward Turkish people	535	3.56	0.88	.07	0.61	3.57 _a	3.90 _b	0.87	3.50 _a	0.78
Explicit attitude toward Moroccan people	535	3.59	0.97	0.10	0.70	3.47 _a	4.18 _b	0.90	3.46 _a	0.89
Explicit attitude toward Surinamese people	535	3.86	0.88	0.07	0.61	3.73 _a	3.71 _a	0.87	4.34 _b	0.76
Explicit attitude toward Antillean people	535	3.72	0.88	0.03	0.39	3.78 _a	3.54 _a	0.93	4.06 _b	0.78
Explicit attitude toward people from your own vs. other background	535	3.76	0.90	0.03	0.39	3.58 _a	3.80 _b	0.92	3.78 _a	0.94
Race IAT ^a	528	-0.23	0.51	0.04	0.47	-0.41 _a	-0.20 _b	0.50	-0.10 _b	0.53
Ethnic Background IAT ^a	527	-0.28	0.51	0.08	0.65	-0.55 _a	0.05 _b	0.46	-0.37 _c	0.43
Classroom conflict	535	2.89	0.75	0.14	0.77	2.94 _a	2.75 _a	0.73	2.86 _a	0.74
Classroom cohesion	535	3.12	0.71	0.12	0.74	2.92 _a	3.23 _b	0.62	3.217 _b	0.79
Classroom belonging	535	3.87	0.70	0.11	0.72	3.77 _a	3.88 _{ab}	0.72	3.99 _b	0.55
Popularity ^b	528	3.65	4.07	0.02	0.30	2.07 _a	3.79 _b	3.69	5.42 _b	4.95
Likability ^b	528	4.14	2.41	0.03	0.39	4.30 _{ab}	3.81 _a	2.31	4.86 _b	2.36
Age ^c	533	13.28	0.80	0.46	0.95	13.39 ^a	13.56 ^a	0.93	12.96 ^b	0.73
Gender ^c	535	0.57	0.50	0.03	0.39	0.62 ^a	0.56 ^a	0.50	0.58 ^a	0.50
SES ^c	439	-0.67	1.70	0.09	0.68	0.27 ^a	-1.08 ^b	1.44	-1.07 ^b	1.84

Note. Means are significantly different if they have different subscripts

^aThe descriptives for the race and ethnic background IATs reflect the original *d* scores before transforming them to reflect *classmates'* implicitly measured attitudes

^bThe descriptives for popularity and likability reflect the scores before correcting for classroom size and total number of given nominations in the classroom

^cThe descriptives for age, gender, and SES for the total sample differ from the descriptives of the final sample as reported in the methods and Table 2

Table 2 Correlations between variables of the present study

	1	2	3	4	5	6	7	8	9	10	11	N	M	SD
<i>Correlations at the within class level</i>														
1	Proportion in-group students	–										389	0.32	0.17
2	Classmates' explicit attitudes	.58	–									389	3.84	0.25
3	Classmates' implicit attitudes	.15	-.25	–								384	-0.01	0.26
4	Classroom conflict (unique)	.03	-.07	.13	–							389	2.86	0.75
5	Classroom cohesion (unique)	.00	.08	-.19	-.47	–						389	3.12	0.72
6	Classroom belonging (unique)	-.09	-.01	-.14	-.49	.52	–					389	3.87	0.70
7	Popularity	-.04	.19	-.23	-.11	.05	.15	–				384	0.05	0.05
8	Likability	.00	.05	.07	-.06	.06	.24	.35	–			384	0.05	0.03
9	Ethnic Dutch vs. Non-Ethnic Dutch	-.05	-.38	.78	.16	-.20	-.18	.00	.00	–		389	0.36	0.48
10	Age	-.22	-.09	-.02	-.03	-.04	.17	.01	.03	–		388	13.31	0.79
11	Gender	.02	.03	-.03	.02	-.02	-.09	.17	.00	-.03	–	389	0.58	0.49
12	SES	-.04	-.13	.30	.05	-.12	-.02	.04	.31	.06	-.05	314	-0.56	1.64
<i>Correlations at the between class level</i>														
1	Classroom diversity	–										25	0.74	0.10
2	Explicit classroom in-group bias	.32	–									25	3.77	0.23
3	Classroom conflict (shared)	.21	.22	–								25	2.88	0.32
4	Classroom cohesion (shared)	.11	-.02	-.58	–							25	3.13	0.28
5	Classroom belonging (shared)	.09	-.34	-.59	.73	–						25	3.86	0.27

Note. Significant correlations are bold ($p < .05$, 2-tailed)

Table 3 Standardized estimates for the models predicting peer relations

	Peer Relations													
	Classroom belonging						Popularity						Likability	
	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2		Model 1	Model 2
	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>
<i>Level 2 predictors and interactions</i>														
Classroom diversity	.22	0.14	.47 [†]	0.26	-	-	-	-	-	-	-	-	-	-
Explicit classroom in-group bias	-.41 [†]	0.21	-.46**	0.17	-	-	-	-	-	-	-	-	-	-
Classroom diversity * Explicit classroom in-group bias	-	-	.40 [†]	0.24	-	-	-	-	-	-	-	-	-	-
<i>Level 1 predictors and interactions</i>														
Proportion in-group students	-.09	0.07	-.06	0.08	-.17*	0.07	-.19**	0.07	-.11*	0.06	-.10	0.06	-.10	0.06
Classmates' explicit attitudes	-.02	0.06	-.03	0.05	.24***	0.06	.20**	0.06	.12	0.07	.11	0.07	.11	0.08
Proportion in-group students * Classmates' explicit attitudes	-	-	-.02	0.04	-	-	-.08	0.08	-	-	-.01	0.05	-.01	0.05
Classmates' implicit attitudes	.04	0.07	.12 [†]	0.06	-.07	0.07	-.12	0.11	.24**	0.08	.29**	0.08	.29**	0.10
Proportion in-group students * Classmates' implicit attitudes	-	-	.10*	0.04	-	-	-.05	0.05	-	-	.07	0.05	.07	0.05
<i>Level 1 control variables</i>														
Ethnic Dutch vs. Non-Ethnic Dutch	-.20**	0.08	-.26***	0.07	-.10	0.09	-.09	0.11	-.14*	0.06	-.18**	0.06	-.18**	0.06
Gender	-	-	-	-	-	-	-	-	.18***	0.04	.18***	0.04	.18***	0.04
<i>Variance components</i>														
Level 2	0.84***	0.15	0.74***	0.20	-	-	-	-	-	-	-	-	-	-
Level 1	0.96***	0.02	0.96***	0.02	0.91***	0.03	0.90***	0.03	0.95***	0.02	0.95***	0.02	0.95***	0.02
<i>R</i> ²														
Level 2	0.16	0.15	0.26	0.20	-	-	-	-	-	-	-	-	-	-
Level 1	0.04	0.02	0.04*	0.02	0.09**	0.03	0.10**	0.03	0.05*	0.02	0.06*	0.02	0.06*	0.02

[†]*p* < .10 (2-tailed); **p* < .05 level (2-tailed); ***p* < .01 level (2-tailed); ****p* < .001 level (2-tailed)

Table 4 Standardized estimates for the models predicting classroom climate

	Classroom climate							
	Classroom cohesion			Classroom conflict				
	Model 1		Model 2	Model 1		Model 2		
	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>	<i>b</i> *	<i>SE</i>		
<i>Level 2 predictors and interactions</i>								
Classroom diversity	.12	0.18	.234	0.25	.15	0.22	-.06	0.28
Explicit classroom in-group bias	-.06	0.21	-.08	0.19	.18	0.23	.22	0.21
Classroom composition * Explicit classroom in-group bias	–	–	.17	0.21	–	–	-.33	0.24
<i>Level 1 predictors and interactions</i>								
Proportion in-group students	.02	0.08	.02	0.08	.08	0.08	.06	0.07
Classmates' explicit attitudes	-.00	0.08	.01	0.09	-.08	0.08	-.10	0.08
Proportion in-group students * Classmates' explicit attitudes	–	–	.02	0.03	–	–	-.04	0.04
Classmates' implicit attitudes	-.09	0.07	-.07	0.08	-.02	0.07	-.09	0.08
Proportion in-group students * Classmates' implicit attitudes	–	–	.02	0.04	–	–	-.08 [†]	0.05
<i>Level 1 control variables</i>								
Ethnic Dutch vs. Non-Ethnic Dutch	-.12**	0.05	-.13**	0.05	.13*	0.06	.17*	0.07
Age	–	–	–	–	-.03	0.04	-.03	0.04
<i>Variance components</i>								
Level 2	0.99***	0.04	0.97***	0.07	0.93***	0.08	0.86***	0.15
Level 1	0.96***	0.02	0.96***	0.02	0.97***	0.02	0.97***	0.02
<i>R</i> ²								
Level 2	0.01	0.04	0.03	0.07	0.08	0.08	0.14	0.15
Level 1	0.04*	0.02	0.04*	0.03	0.02	0.02	0.03	0.02

[†]*p* < .10 (2-tailed); **p* < .05 level (2-tailed); ***p* < .01 level (2-tailed); ****p* < .001 level (2-tailed)

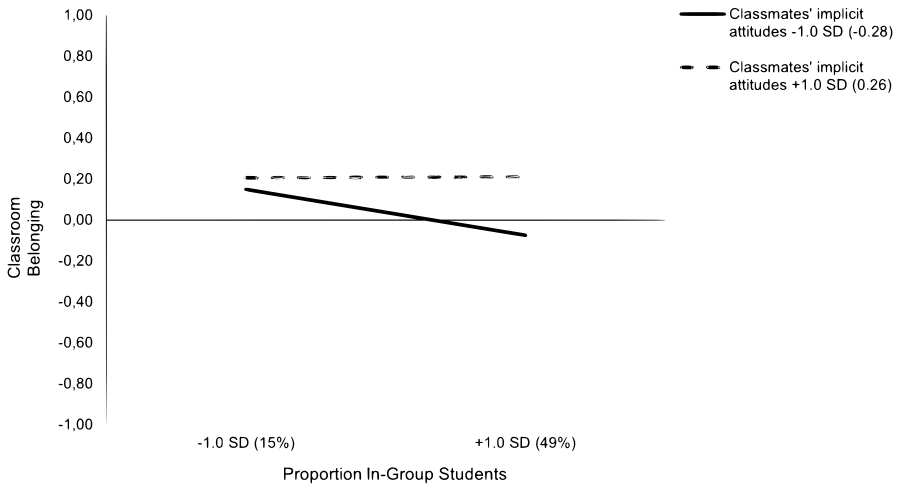


Fig. 2 Interaction effect of the proportion of in-group students and classmates' implicitly measured ethnic attitudes on students' sense of classroom belonging ($N=384$)

against the students' in-group, whereas, unexpectedly, the other relations between classmates' attitudes and students' CSE were not significant.

We also hypothesized that the effect of the presence of in-group classmates on students' unique perception of their CSE experiences would be stronger when their classmates had more negative ethnic attitudes toward the students' in-group (Hypothesis 6). Table 3 (Model 2) and Table 4 (Model 2) show that, unlike we hypothesized, there was a positive interaction between the proportion of in-group students and classmates' implicitly measured ethnic attitudes on students' sense of classroom belonging ($b^*=0.10$, $p=.013$, indicating a weak effect). That is, the effect of the proportion of in-group peers on students' sense of classroom belonging was less (rather than more) positive when the students' classmates had a more negative as compared to more positive implicitly measured attitude toward the students' ethnic in-group. This interaction is shown in Fig. 2. It appeared that students with more same-ethnic peers in the classroom experienced a lower sense of belonging, but only when those peers were biased against the students' in-group. In addition, the interaction effect between the proportion in-group students and classmates' implicitly measured ethnic attitudes on classroom conflict was in the expected direction, although it failed to reach significance ($b^*=-0.08$, $p=.080$).

3.2.2 Between classroom level

The standardized results at the between classroom level for classroom belonging can be found in Table 3 (Model 1) and for classroom cohesion and conflict in Table 4 (Model 1). Our third hypothesis, that students in more heterogeneous classes would experience a more positive CSE, was not supported. Furthermore, it was hypothesized that a less strong average classroom in-group bias would be associated with a more positive CSE (Hypothesis 7). In line with this hypothesis, although borderline

significant, average classroom in-group bias was negatively associated with students' shared perception of their classroom belonging ($b^* = -0.41$, $p = .051$, indicating a strong effect). This indicates that students reported less belonging when the average bias in the classroom was stronger. Finally, it was hypothesized that the anticipated effects of classroom diversity were stronger in classrooms with a strong average degree of in-group bias (Hypothesis 8). Results regarding this hypothesis can be found under the Models 2 in Tables 3 and 4. Although the findings were in the expected direction, the interaction between classroom diversity and classroom explicit in-group bias on classroom belonging just failed to reach significance ($b^* = 0.40$, $p = .086$).

3.2.3 Differences between non-ethnic dutch and ethnic dutch students

To explore if the proposed model differed for non-ethnic Dutch versus ethnic Dutch students, the full *within class level* model (i.e., Model 2) was estimated separately for both groups. As this model did not include predictors at the between level, we did not estimate a two-level model but took the hierarchical structure of the data (i.e., students nested in teachers) into account by using cluster-robust standard errors (i.e., including “type=complex” in the Mplus syntaxes; (McNeish et al., 2017). To examine the differences between non-ethnic Dutch and ethnic Dutch students, z -scores were calculated for the differences between the unstandardized regression coefficients of the predictor pairs using the following equation (Paternoster et al., 1998). A significant difference between the regression coefficients was indicated by a z -score above 1.96.

$$z = \frac{b_{\text{predictor non-ethnic students}} - b_{\text{predictor ethnic-Dutch students}}}{\sqrt{(SE_{\text{predictor non-ethnic students}})^2 + (SE_{\text{predictor ethnic-Dutch students}})^2}}$$

Results showed that for classroom climate the results did not differ between ethnic Dutch and non-ethnic Dutch students (all p -values $> .05$). In terms of peer relations, some effects differed between the ethnic groups. First, the proportion of in-group classmates was only significantly associated with non-ethnic Dutch students' popularity ($b^* = -0.19$, $p = .003$, indicating a weak effect) and likability ($b^* = -0.17$, $p < .001$, indicating a weak effect). These effects significantly differed between the groups (popularity: $z = -3.18$, $p = .001$; likability: $z = -1.96$, $p = .050$). Second, classmates' implicitly measured ethnic attitudes were only significantly associated with non-ethnic Dutch students' likability ($b^* = 0.32$, $p < .001$, indicating a moderate effect). This effect was significantly different from the effect of implicitly measured attitudes on ethnic Dutch students' likability ($z = 2.89$, $p = .004$). Finally, classmates' implicitly measured ethnic attitudes were negatively associated with the classroom belonging of non-ethnic Dutch students ($b^* = -0.23$, $p = .019$, indicating a weak effect) but positively with the classroom belonging of their ethnic Dutch peers ($b^* = 0.35$, $p = .002$, indicating a moderate effect). This difference was significant ($z = -4.03$, $p < .001$). All other effects in terms of peer relations did not differ between the groups (p -values $> .05$).

3.3 Robustness checks

A robustness check was performed in order to examine if the significant findings for proportion in-group and its interactions would change when the proportion of in-group students would be calculated differently. In the main analyses, the proportion of in-group students was determined based on the students' ethnic group membership. For example, for each Turkish student, their proportion in-group represented the proportion of Turkish students in their class. However, in the implicit attitude measures, people with a Turkish and Moroccan background were grouped together, and in the race IAT all people with a dark skin color were grouped together. Thus, for the IAT measures (contrary to the explicit attitude measures), the proportion of in-group measure did not entirely match the groups in the IAT. Therefore, the robustness of the findings for classmates' implicitly measured ethnic attitudes, and its interaction with proportion in-group were checked by recalculating the proportion in-group by combining the Turkish and Moroccan students in one group and by combining the Surinamese and Antillean students in one group. The results of these analyses were mostly similar as the results of the main analyses and the differences were all very small. This shows that the findings of the present study are robust. A second robustness check involved our selection of ethnic Dutch students. It could be that we classified students as ethnic Dutch who actually belonged to other minority groups. The reason for this is that the term Dutch can be used in an ethnic, heritage-based sense of who is a national but also in a civic sense (i.e., who holds a passport) to indicate who is a Dutch citizen (see van Vemde et al., 2021). We could check this possibility by inspecting students' reports of how others would identify them, which were available in the dataset as well. Based on their answers we excluded 10 students who we classified as Dutch but indicated that others saw them as Moroccan, Surinamese, American, Malaysian, Nicaraguan, or Spanish or who indicated that they did not know how others would see them. Excluding these students did not alter our results, which indicates that our findings are robust. Detailed results of the robustness checks are available upon request from the first author.

4 Discussion

The goal of the present study was to examine whether and how ethnic classroom composition and students' explicitly and implicitly measured ethnic attitudes were associated with students' experiences of the classroom social environment (CSE). In contrast to earlier research, the present study examined CSE experiences at both the *within* and *between* classroom level and distinguished between students' unique and shared perceptions of both individual (i.e., peer relations) and collective (i.e., classroom climate) aspects. In general, our findings show that both ethnic classroom composition and ethnic attitudes are associated with different aspects of the CSE, mainly those involving peer relations. Notably, students had a more positive status in the classroom when classmates had more positive attitudes toward their ethnic group. For the other CSE aspects (classroom belonging and classroom climate) we found mixed results.

4.1 Within or between classroom CSE experiences?

In contrast to earlier studies, the present study examined students' CSE experiences at the *within* and *between* classroom level. In doing so, we were able to show that most effects occurred at the *within* classroom level, while most effects of classroom diversity and ethnic attitudes at the *between* classroom level were not significant. The absence of effects at the classroom level may be due to a lack of classroom level variance in the CSE. That is, the relatively low ICC1s at the classroom level (see Table 1), suggest that students within the same class experience the social environment of the class very differently. This was even the case for the classroom climate measures (cohesion and conflict) even though the measures we used targeted students' experience of the class as a whole, as recommended by (Marsh et al., 2012). That is, prior studies often asked about students' individual experience of classroom cohesion and conflict (e.g., "I experience a lot of conflict.") and aggregate these to the classroom level. In this study, we used measures specifically targeting the class (e.g., "Students in this class often have conflict with each other."). Regardless, students within the same class also seemed to experience the classroom climate very differently. Hence, it appears that the experience of the social environment is not so much a shared experience, but predominantly an unique experience. This raises the question to what extent the classroom is the most meaningful unit of analysis when it comes to students' social experiences at school. For future research it could be interesting to focus on smaller units, such as friendship groups or friendship dyads within a class, to investigate to what extent students within these smaller units have stronger shared perceptions of the social environment.

Moreover, with the exception of classroom belonging—which, in line with our hypothesis, was lower when students in a class on average had more bias in favor of their own in-group—other effects of class composition or classmates' attitudes on the CSE at the between level were not significant. For classroom composition, the lack of effect could be due to a restriction of range as the sample did not include fully segregated classes (i.e., classes with only majority or minority students). Hence, the effects of classroom diversity on students' shared CSE experiences might have been more pronounced if the sample had also included more segregated classes. With regards to classmates' explicitly and implicitly measured attitudes, the ICC1 indicated that, similar to the CSE, attitudes differed greatly between students within the same class. This might explain why ethnic bias at the classroom level did not have an impact on classroom climate. Rather, the effects of classroom composition and classmates' attitudes seem to reflect more unique experiences for students.

4.2 Effects of classroom composition on students' CSE

In line with earlier studies using proportion in-group peers (e.g., Hornstra et al., 2015; Rock et al., 2011) and consistent with the notion that it can be easier to connect with others in more homogenous groups (see belongingness perspective; Baumeister & Leary, 1995; Rjosk et al., 2017), it was hypothesized that when students had more

in-group classmates this would positively affect students' CSE. Moreover, based on the *imbalance of power thesis* (Graham, 2006) it was hypothesized that in more diverse classrooms, students collectively experienced a more positive CSE. Contrary to these hypotheses, however, it was found that a larger proportion of in-group peers was related to students being nominated less often as popular and likable. Our additional analyses revealed that this was only the case for non-ethnic Dutch students, which is consistent with earlier research in Dutch secondary schools (Stevens et al., 2020). An explanation for this could be that during adolescence youth begins to prioritize social status as an indicator of who is more or less popular (Mali et al., 2019). In the Netherlands, as well as in other countries, ethnic minority groups (e.g., Turks, Moroccans, Surinamese, or Antilleans) tend to have a lower social standing, indicating that, compared to the ethnic majority group, they are perceived more negatively and sometimes even treated with hostility (Stevens et al., 2020; Zick et al., 2008). The lower social standing of minority group students could have caused classmates to nominate members of these groups less often as popular or likable, even when the students are from the same ethnic group.

4.3 Effects of classmates' attitudes on students' CSE

The current study also examined the effects of classmates' ethnic attitudes on students' unique and shared perceptions of their CSE. In line with previous literature (e.g., Thijs et al., 2014) and our seventh hypothesis, a higher degree of average classroom in-group bias was associated with a lower shared sense of classroom belonging. In line with previous research (Bayram Özdemir et al., 2018; Jugert et al., 2011), it was also hypothesized (Hypothesis six) that positive ethnic attitudes toward the students' in-group would be associated with more positive unique CSE experiences. Interestingly, classmates' implicitly measured ethnic attitudes were associated with more likability but not with popularity, while classmates' explicitly measured ethnic attitudes were associated with more popularity and not with likability. An explanation for this could be that when students nominate someone as likable, this is based on more intuitive or "emotional" judgements indicating someone's private sentiments of attraction or repulsion toward another (Cillessen & Marks, 2011; Moreno, 1934). Emotional judgments are most likely more difficult to control and therefore it could be that implicit attitude measures are more strongly related to student likability than popularity. Students' popularity, on the other hand, reflects status and reputation (Cillessen & Marks, 2011) and therefore nominating someone as popular may not so much be an "emotional" judgement but rather a deliberate judgment. Hence, this could cause popularity nominations to be more strongly affected by explicit attitude measures. Future research could test this explanation and further disentangle the effects of different attitude measures on students' likability and popularity.

4.4 Moderation effects of ethnic attitudes

We also examined if classmates' implicitly and explicitly measured ethnic attitudes moderated the effect of classroom composition on students CSE experiences. Based

on previous literature, we hypothesized that the presence of in-group classmates was especially important when classmates' attitudes toward the students' in-group were less positive and that classroom diversity in general was especially important when there was a strong average degree of in-group bias (Bayram Özdemir et al., 2018). Only the interaction effect of proportion of in-group classmates and classmates' implicitly measured ethnic attitudes on students' classroom belonging was significant. As expected, when students experienced positive attitudes toward their group from their classmates, they also experienced a stronger sense of classroom belonging. More interestingly and in contrast with our expectations, students' classroom belonging was lowest when they were in classes with many in-group peers and when classmates were more biased against the students' in-group. We don't have a clear-cut explanation for this effect, but it might have to do with the specific classmates that held the attitudes. If out-group classmates are the ones with negative ethnic attitudes, this could result in students withdrawing and disidentifying with the general group, because they experience less positive attitudes toward their own ethnic-group (cf. Rejection Disidentification Hypothesis; Jasinskaja-Lahti et al., 2009). This disidentification may happen especially in classes where the in-group is large enough, so students can distance themselves from out-group classmates. If in-group classmates have negative attitudes toward their own group, students might internalize those attitudes, which could ultimately result in to not wanting to belong to the classroom peer group at all. The latter options seems plausible, as our descriptive results in Table 1 show that minority students were less positive about their in-group than majority students.

In addition, there were some interaction effects that just failed to reach significance. Still, they indicated an overall trend—in line with our hypotheses—which suggested that at the *within* classroom level, negative attitudes were more problematic for conflict when the ethnic in-group was larger. While at the *between* classroom level a higher degree of in-group bias seemed to be problematic for belonging when the classroom was less diverse. However, future research should examine the interplay between attitudes and classroom composition with a larger number of classes to shed more light on these potential interactions.

4.5 Ethnic differences in CSE quality

Although not the main focus of the present study, results showed that, in mixed classes, students with an ethnic Dutch background experienced a more negative CSE as compared to students with a non-ethnic Dutch background. This is reflected not only in the descriptive statistics and correlations (see Tables 1 and 2), but also by the effects of our control variable for ethnicity that was significantly associated with almost all CSE measures even after including proportion in-group peers. It could be that, for students with a majority background, the presence of more minority group classmates resulted in the experience of out-group threat (Bubritzki et al., 2018), and thus a lower quality of their CSE. The distribution of ethnic- versus non-ethnic Dutch students across the classrooms in our sample indicated that ethnic majority students often formed the numerical minority in their classrooms (in 22 out of the

25 participating classes, ethnic Dutch students formed the numerical minority). Previous research has indicated, also for majority group students, that being the numerical minority in their schools increased the risk of being victimized by their peers (Felix & You, 2011; Graham & Juvonen, 2002). This could explain why the ethnic Dutch students in our sample experienced a lower quality of their CSE.

4.6 Limitations and future directions

In evaluating the present study, some limitations should be addressed. First, the present study included a relative small number of classes (25 classes) while larger sample sizes (i.e., 30–50 classes) are recommended when performing multilevel analyses (Maas & Hox, 2005). It could therefore be that a higher number of classroom units could have increased the power to detect small differences at the classroom level. Nonetheless, other researchers recommend a minimum of 20 clusters for multilevel analyses (Snijder & Bosker, 2012). Moreover, when the dependent variables are continuous, like in the present study, the bias in standard errors, which might be the result of a smaller sample size at the cluster level, seems limited (Maas & Hox, 2005). Nevertheless, future research should try to include a larger number of clusters. Second, the classes participating in the present study were very ethnically diverse and there were no fully segregated classes in our sample. Stronger effects could have been obtained if there was more variance in classroom composition ranging from very homogeneous to very diverse. Third, there was no information available on how long students had been living in the Netherlands when they were born in a different country or on their generation status (i.e., whether students were first, second, or third generation migrants). Regardless, second generation migrants experience more discrimination than first generation migrants in the Netherlands (Andriesen et al., 2020; Dagevos et al., 2022), suggesting that classmates most probably still see second or third generation minority group peers as such. Hence, it seems unlikely the time that a student was living in the Netherlands or their generation status was a factor which was considered by classmates. This is further supported by a study amongst non-Roma and Roma Hungarian students which found that majority students disliked peers whom they perceived as a member of the minority group, even though these students might be majority group members as well (Boda & Néray, 2015). Still, future research could take the time students are living in the “host” country or immigrant generation status into account in order to examine if this makes the effects of classroom composition and classmates’ ethnic attitudes on students’ CSE experiences stronger or weaker. Fourth, both the IAT and self-report attitude measures have their limitations. Both types of measures are widely debated in terms of their psychometric and conceptual quality. For example, there are concerns about the test–retest reliability and the low convergent validity of the IAT (i.e., the IAT is rather weakly correlated to other measures of implicit attitudes; Blanton & Jaccard, 2022; Lundberg & Payne, 2022). The poor test–retest reliability is however not necessarily problematic as this may be due to the high context dependency of the IAT (Gawronski, 2019). Also self-reports have been scrutinized, especially their susceptibility to self-presentation bias which may limit their

validity (Schwarz & Oyserman, 2001; Steffens, 2004; van den Bergh et al., 2010). To account for the limitations of both attitude measures, and because both attitude measures each show unique associations with different outcome measures (Greenwald et al., 2009; Hahn & Gawronski, 2017; Kurdi et al., 2019)—which was also the case for this study—the present study included both of them. Fifth, the present study was cross-sectional. Therefore, it is not possible to establish causal relationships between classroom diversity, attitudes, and students' CSE experiences. However, we expected that it was more likely for classroom composition to affect the CSE rather than the other way around as it is unlikely that CSE experiences determine structural aspects like ethnic classroom composition. For ethnic attitudes, however, it may be that these are reciprocally related with aspects of the CSE (e.g., when there is less conflict, attitudes may become more positive, and vice versa). Future research could examine the directional nature of these relationships using a longitudinal design to shed more light on the direction of effects. Sixth, we used classmates' attitudes measured by the race IAT to predict outcomes for students with a Surinamese/Antillean background and classmates' attitudes measured by the ethnicity IAT to predict outcomes for students with a Turkish/Moroccan background. However, we cannot fully exclude the possibility that students also thought about Turkish/Moroccan people while completing the race IAT. We deem this unlikely, given that we used images of people with a distinct dark or white skin. Nevertheless, it could be possible that the results of the race IAT are ambiguous for Turkish/Moroccan students, hence our decision not to include Turkish/Moroccan students in the analysis of the race IAT. Likewise, we excluded Surinamese/Antillean students from the analyses with the ethnicity IAT, as this measure was focused on Turks/Moroccans versus Dutch. Although, we could have included such an IAT with typical Surinamese/Antillean names, this was not feasible timewise. But even more importantly, in the Netherlands, names of students with a Surinamese or Antillean background are not always distinct from typical ethnic Dutch names. Therefore, we excluded the Surinamese/Antillean students from the analyses with the ethnicity IAT, as the effects of classmates' attitudes on this IAT might be ambiguous for them. Finally, the choices made regarding the categorization of students into ethnic groups and our instruments might have affected our findings. Ethnic background and the categorization of students based on their ethnic background taps into a complex reality that cannot fully be captured when students are categorized in different subgroups. Each possible way of categorizing (e.g., based on self-identification or other-identification, and grouping different minority groups together or not) has its limitations and may not fully do justice to this complex reality. Nevertheless, our findings indicate that the self-identification corresponded strongly with other-identification and our robustness checks showed that redefining the in-groups or leaving out the small number of students whose self-identification and other-identification did not correspond did not alter our findings. In addition, our choice to measure explicit attitudes for specific target groups at the within student level resulted in a substantial percentage of students (26.4%) being excluded from the analysis as attitude targets. However, more general attitude measures like students' general perceptions of immigrants (Bayram Özdemir et al., 2018) might be less predictive of student outcomes, due to the fact that attitudes might differ per group (Akkermans & Kloosterman, 2022). Still, the

attitudes of the excluded students were not completely disregarded as they were still taken into account as holders of the attitudes towards the specific groups and when calculating the between classroom level more general attitude measure of classroom in-group bias.

The results of the present study offer some interesting lines for future research. One of the strengths of the present study is that we examined the explicitly and implicitly measured ethnic attitudes of students' classmates. Thus, contrary to previous studies we examined the effect of group attitudes on students from the groups concerned. In doing so, we assumed that those "targets" correctly perceived the attitudes, but this needs to be tested. Therefore, future research could include students' perceptions of their classmates' attitudes in relation to students' CSE experiences. Second, in the present study we aggregated classmates' attitudes. However, it could be that certain classmates are more influential than others, or for example, that the presence of only a few classmates with a very negative attitude has more effect on students' CSE than the average ethnic attitude of classmates. Future research could use social network analysis to further examine this. Third, future research could examine if there are age differences in how classmates' attitudes affect students' CSE experiences. Children already display explicit prejudice around the age of 3 to 4 years and display an implicit bias toward their in-group by at least 6 years old (Vezzali et al., 2012). However, children's explicit prejudice seems to decline when they get older (Cristol & Gimbert, 2008), whereas there are indications that implicit bias is more stable during children's development (Dunham et al., 2006). It is therefore interesting to examine if attitudes have different effects on students' CSE in different stages of their development.

4.7 Practical implications

The current study also has some practical implications for education and schools. In general, it showed that the ethnic classroom composition or classmates' ethnic attitudes affected some of students' CSE experiences. Our study indicated that students felt more at home in more diverse classrooms, although the effect was marginally significant. This implies that it could be beneficial if schools aim for a more diverse school population, in which different groups are equally represented. Specifically, this means that local or national policies should be implemented aimed at countering school segregation. Based on earlier research in the Netherlands, effective policies could focus on parents, such as limiting their school choices or facilitating parental initiatives to reduce segregation (Peters & Walraven, 2011). However, solely focusing on classroom diversity and the presence of in-group peers, and thus aiming for mixed classes, to increase students' CSE experiencing might not be enough. Also ethnic attitudes of classmates play a role when it comes to the CSE experiences of secondary school students. Hence, it is important for schools and teachers to be aware of how students' attitudes toward ethnic-others can impact diverse classrooms. Practically, this implies that schools and teachers could try to prevent students from developing a strong in-group bias and tackle negative attitudes toward ethnic out-groups. This could, for example, be done by implementing interventions

aimed at increasing students' perspective taking and empathy skills (Beelmann & Heinemann, 2014).

4.8 Conclusion

The present study examined the effects of classroom composition and (the moderating role of) classmates' implicitly and explicitly measured ethnic attitudes on students' peer relations and classroom climate. The findings suggest that a diverse classroom and less same ethnic classmates, as well as positive ethnic attitudes of classmates can affect students' CSE experiences positively. Thus, in order to make sure that students experience a positive CSE, it is important to not only focus on classroom diversity, but also take into account students' ethnic attitudes and that these could have different effects depending on the ethnic classroom composition.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11218-022-09747-x>.

Acknowledgements The authors would like to thank Frank Scholten and Bettina de Jong for their contribution in setting up the study and data collection.

Author's contribution LvV: Conceptualization, Methodology, Formal analysis, Data curation, Writing—original draft; JT: Conceptualization, Methodology, Writing—review & editing, Supervision; LH: Conceptualization, Methodology, Writing—review & editing, Supervision. All authors read and approved the final manuscript.

Funding This work was supported by the Netherlands Initiative for Education Research [NRO/PROO grant 40.5.18300.032], which resides under the Netherlands Organization for Scientific Research (NWO).

Data availability The data for the variables used in the present study will be made available via the Open Science Foundation.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval The study was approved by the ethics committee of the University of Amsterdam in the Netherlands.

Informed consent Passive parental and student consent was obtained for the students.

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

- Aboud, F. E., Mendelson, M. J., & Purdy, K. T. (2003). Cross-race peer relations and friendship quality. *International Journal of Behavioral Development, 27*(2), 165–173. <https://doi.org/10.1080/01650250244000164>
- Agirdag, O., Demanet, J., van Houtte, M., & van Avermaet, P. (2011). Ethnic school composition and peer victimization: A focus on the interethnic school climate. *International Journal of Intercultural Relations, 35*(4), 465–473. <https://doi.org/10.1016/j.ijintrel.2010.09.009>
- Akkermans, M., & Kloosterman, R. (2022). *Gediscrimineerd gevoeld?* [Experienced discrimination?]. <https://www.cbs.nl/nl-nl/longread/statistische-trends/2022/gediscrimineerd-gevoeld>.
- Andriessen, I., Hoegen Dijkhof, J., van der Torre, A., van den Berg, E., Pulles, I., Ledema, J., & de Voogd-Hamelink, M. (2020). *Ervaren discriminatie in Nederland II*. [Experienced discrimination in the Netherlands II]. <https://www.scp.nl/publicaties/publicaties/2020/04/02/ervaren-discriminatie-in-nederland-ii>.
- Azjen, I., Fishbein, M., Lohman, S., & Albarracín, D. (2018). The influence of attitudes on behavior. In D. Albarracín & B. T. Johnson (Eds.), *The handbook of attitudes* (pp. 197–255). Routledge.
- Bakker, J., Denessen, E., Pelzer, B., Veneman, M., & Lageweg, S. (2007). De houding jegens klasgenoten: Etnisch gekleurd? Een onderzoek naar factoren van invloed op de attitude van basisschoolleerlingen jegens klasgenoten van verschillende etnische herkomst [The attitude towards classmates: Ethnically colored? A study into factors that influence the attitude of primary school students towards classmates of different ethnic origins]. *Pedagogiek, 27*(3), 210–219.
- Baumeister, R. F., Leary, Mark, & R. (1995). Need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Bayram Özdemir, S., Özdemir, M., & Stattin, H. (2016). What makes youth harass their immigrant peers? Understanding the risk factors. *Journal of Early Adolescence, 36*(5), 601–624. <https://doi.org/10.1177/0272431615574887>
- Bayram Özdemir, S., Sun, S., Korol, L., Özdemir, M., & Stattin, H. (2018). Adolescents' engagement in ethnic harassment: Prejudiced beliefs in social networks and classroom ethnic diversity. *Journal of Youth and Adolescence, 47*(6), 1151–1163. <https://doi.org/10.1007/s10964-017-0795-0>
- Beelmann, A., & Heinemann, K. S. (2014). Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs. *Journal of Applied Developmental Psychology, 35*(1), 10–24. <https://doi.org/10.1016/j.appdev.2013.11.002>
- Benner, A. D., & Crosnoe, R. (2011). The racial/ethnic composition of elementary schools and young children's academic and socioemotional functioning. *American Educational Research Journal, 48*(3), 621–646. <https://doi.org/10.3102/0002831210384838>
- Benner, A. D., & Graham, S. (2009). The transition to high school as a developmental process among multiethnic urban youth. *Child Development, 80*(2), 356–376. <https://doi.org/10.1111/j.1467-8624.2009.01265.x>
- Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., Maquil, A., Demoulin, S., & Leyens, J. P. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European countries. *Journal of Personality and Social Psychology, 96*(4), 843–856. <https://doi.org/10.1037/a0013470>
- Blanton, H., & Jaccard, J. (2022). Listening to measurement error: Lessons from the IAT. In J. A. Krosnick, T. H. Stark, & A. L. Scott (Eds.), *The Cambridge handbook of implicit bias and racism*. Cambridge.
- Boda, Z., & Néray, B. (2015). Inter-ethnic friendship and negative ties in secondary school. *Social Networks, 43*, 57–72. <https://doi.org/10.1016/j.socnet.2015.03.004>
- Bubritzki, S., van Tubergen, F., Weesie, J., & Smith, S. (2018). Ethnic composition of the school class and interethnic attitudes: A multi-group perspective. *Journal of Ethnic and Migration Studies, 44*(3), 482–502. <https://doi.org/10.1080/1369183X.2017.1322501>
- Cillessen, A. H. N., & Marks, P. E. L. (2011). Conceptualizing and measuring popularity. In A. H. N. Cillessen, D. Schwartz, & L. Mayeux (Eds.), *Popularity in the peer system* (pp. 25–56). Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Cristol, D., & Gimbert, B. (2008). Racial perceptions of young children: A review of literature post-1999. *Early Childhood Education Journal, 36*(2), 201–207. <https://doi.org/10.1007/s10643-008-0251-6>

- Dagevos, J., de Voogd-hamelink, M., & Damen, R. (2022). *Gevestigd, maar niet thuis* [Settled, but not home]. <https://www.scp.nl/publicaties/publicaties/2022/10/11/gevestigd-maar-niet-thuis.-eerste-bevindingen-uit-de-survey-integratie-migranten-sim2020>.
- Dewey, J. (1966). *Democracy and Education*. The Free Press.
- Dovidio, J. F., Brigham, J. C., Johnson, B. T., & Gaertner, S. L. (1996). Stereotyping, prejudice, and discrimination: Another look. In N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and Stereotyping* (pp. 276–319). Guildford.
- Dunham, Y., Baron, A. S., & Banaji, M. R. (2006). From American city to Japanese village: A cross-cultural investigation of implicit race attitudes. *Child Development*, 77(5), 1268–1281. <https://doi.org/10.1111/j.1467-8624.2006.00933.x>
- Durkin, K., Hunter, S., Levin, K. A., Bergin, D., Heim, D., & Howe, C. (2012). Discriminatory peer aggression among children as a function of minority status and group proportion in school context. *European Journal of Social Psychology*, 42(2), 243–251. <https://doi.org/10.1002/ejsp.870>
- Felix, E. D., & You, S. (2011). Peer victimization within the ethnic context of high school. *Journal of Community Psychology*, 39(7), 860–875. <https://doi.org/10.1002/jcop.20465>
- Fisher, D. L., & Fraser, B. J. (1981). Validity and use of My Class Inventory. *Science Education*, 65(2), 14–156.
- Fleiss, J. (1986). *The design and analysis of clinical experiments*. Wiley.
- Fraser, B. J., Anderson, G. J., & Walberg, H. J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory and My Class Inventory*. (3rd ed.). Western Australian Institute of Technology.
- Gawronski, B. (2019). Six lessons for a cogent science of implicit bias and its criticism. *Perspectives on Psychological Science*, 14(4), 574–595. <https://doi.org/10.1177/1745691619826015>
- Gawronski, B., & Creighton, L. A. (2013). Dual process theory. In D. E. Carlston (Ed.), *Moral thinking, fast and slow* (pp. 5–19). Routledge.
- Graham, S. (2006). Peer victimization in school. *Current Directions in Psychological Science*, 15(6), 317–321. <https://doi.org/10.1111/j.1467-8721.2006.00460.x>
- Graham, S., & Echols, L. (2018). Race and ethnicity in peer relations research. In Bukowski, W. M., Laursen, B., & Rubin, K. H. (Eds.), *Handbook of peer interactions, relationships, and groups* (2nd ed., pp. 590–614). Guilford Publications. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2008-19117-022>.
- Graham, S., & Juvonen, J. (2002). Ethnicity, peer harassment, and adjustment in middle school: An exploratory study. *Journal of Early Adolescence*, 22(2), 173–199. <https://doi.org/10.1177/0272431602022002003>
- Greenwald, A. G., Brendl, M., Cai, H., Cvencek, D., Dovidio, J. F., Friese, M., Hahn, A., Hehman, E., Hofmann, W., Hughes, S., Hussey, I., Jordan, C., Kirby, T. A., Lai, C. K., Lang, J. W. B., Lindgren, K. P., Maison, D., Ostafin, B. D., Rae, J. R., & Wiers, R. W. (2022). Best research practices for using the Implicit Association Test. *Behavior Research Methods*, 54(3), 1161–1180. <https://doi.org/10.3758/s13428-021-01624-3>
- Greenwald, A. G., & Lai, C. K. (2020). Implicit social cognition. *Annual Review of Psychology*, 71, 419–445. <https://doi.org/10.1146/annurev-psych-010419-050837>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. <https://doi.org/10.1037/0022-3514.85.2.197>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17–41. <https://doi.org/10.1037/a0015575>
- Hahn, A., & Gawronski, B. (2017). Implicit social cognition. In Wixted, J. T. (Ed.), *The Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., Vol. 4, pp. 395–427). Wiley.
- Hamm, J. V., Bradford Brown, B., & Heck, D. J. (2005). Bridging the ethnic divide: Student and school characteristics in African American, Asian-descent, Latino, and White adolescents' cross-ethnic friend nominations. *Journal of Research on Adolescence*, 15(1), 21–46. <https://doi.org/10.1111/j.1532-7795.2005.00085.x>
- Hirschman, A. O. (1964). The paternity of an index. *The American Economic Review*, 54(5), 761–762.

- Hoglund, W. L. G., & Hosan, N. E. (2013). The context of ethnicity: Peer victimization and adjustment problems in early adolescence. *Journal of Early Adolescence*, 33(5), 585–609. <https://doi.org/10.1177/0272431612451925>
- Hornstra, L., van der Veen, I., Peetsma, T. T. D., & Volman, M. (2015). Does classroom composition make a difference: Effects on developments in motivation, sense of classroom belonging, and achievement in upper primary school. *School Effectiveness and School Improvement*, 26(2), 125–152. <https://doi.org/10.1080/09243453.2014.887024>
- Jackson, M. F., Barth, J. M., Powell, N., & Lochman, J. E. (2006). Classroom contextual effects of race on children's peer nominations. *Child Development*, 77(5), 1325–1337. <https://doi.org/10.1111/j.1467-8624.2006.00937.x>
- Jasinskaja-Lahti, I., Liebkind, K., & Solheim, E. (2009). To identify or not to identify? National disidentification as an alternative reaction to perceived ethnic discrimination. *Applied Psychology*, 58(1), 105–128. <https://doi.org/10.1111/j.1464-0597.2008.00384.x>
- Jugert, P., Noack, P., & Rutland, A. (2011). Friendship preferences among German and Turkish preadolescents. *Child Development*, 82(3), 812–829. <https://doi.org/10.1111/j.1467-8624.2010.01528.x>
- Juvonen, J. (2006). Sense of belonging, social bonds, and school functioning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 655–674). Lawrence Erlbaum Associates Publishers.
- Juvonen, J., Nishina, A., & Graham, S. (2006). Ethnic diversity and perceptions of safety in urban middle schools. *Psychological Science*, 17(5), 393–400. <https://doi.org/10.1111/j.1467-9280.2006.01718.x>
- Kalter, F., Jonsson, J. O., van Tubergen, F., & Heath, A. (2018). *Growing up in diverse societies: The integration of the children of immigrants in England, Germany, the Netherlands, and Sweden*. British Academy. <https://doi.org/10.5871/bacad/9780197266373.001.0001>
- Knol, F., Boelhouwer, J., & Veldheer, V. (2012). *Statusontwikkeling van wijken in Nederland 1998–2010* [Status development of neighborhoods in the Netherlands from 1998–2010].
- König, S., Stang-Rabrig, J., & McElvany, N. (2022). Adolescents' implicit attitudes towards people with immigrant background: Differences and correlates. *Social Psychology of Education*, 25(6), 1381–1409. <https://doi.org/10.1007/s11218-022-09722-6>
- Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2019). Relationship between the implicit association test and intergroup behavior: A meta-analysis. *American Psychologist*, 74(5), 569–586. <https://doi.org/10.1037/amp0000364>
- LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11(4), 815–852. <https://doi.org/10.1177/1094428106296642>
- Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, 34(2), 120–131. <https://doi.org/10.1016/j.cedpsych.2008.12.001>
- Lundberg, K. B., & Payne, B. K. (2022). The bias of crowds: Rethinking implicit bias in social context. In J. A. Krosnick, T. H. Stark, & A. L. Scott (Eds.), *The Cambridge handbook of implicit bias and racism*. Cambridge.
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>
- Madsen, K. R., Damsgaard, M. T., Rubin, M., Jervelund, S. S., Lasgaard, M., Walsh, S., Stevens, G. G. W. J. M., & Holstein, B. E. (2016). Loneliness and ethnic composition of the school class: A nationally random sample of adolescents. *Journal of Youth and Adolescence*, 45(7), 1350–1365. <https://doi.org/10.1007/s10964-016-0432-3>
- Madva, A., & Brownstein, M. (2018). Stereotypes, prejudice, and the taxonomy of the implicit social mind. *Nous*, 52(3), 611–644. <https://doi.org/10.1111/nous.12182>
- Mali, L. V., Schwartz, D., Badaly, D., Luo, T. J., Malamut, S., Ross, A. C., & Duong, M. T. (2019). Unpopularity with same- and cross-ethnicity peers as predictors of depressive symptoms during adolescence. *Journal of Applied Developmental Psychology*, 62, 93–101. <https://doi.org/10.1016/j.appdev.2019.02.001>
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, 47(2), 106–124. <https://doi.org/10.1080/00461520.2012.670488>

- McMahon, S. D., Wernsman, J., & Rose, D. S. (2009). The relation of classroom environment and school belonging to academic self-efficacy among urban fourth- and fifth-grade students. *Elementary School Journal*, *109*(3), 267–281. <https://doi.org/10.1086/592307>
- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. <https://doi.org/10.1037/met0000078>
- Millisecond Software. (2021). *Inquisit Web*. Millisecond Software.
- Moreno, J. L. (1934). *Who shall survive?* Nervous and Mental Disease Publishing Co.
- Muthén, L. K., & Muthén, B. O. (2021). *Mplus* (8.6). Muthén & Muthén.
- Özdemir, M., & Bayram Özdemir, S. (2017). The role of school context in adolescent's attitudes towards immigrants and inter-ethnic friendships. In Lundberg, E. (Ed.), *The mechanisms of tolerance: An anthology*. Stockholm: The Living History Forum.
- Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, *36*(4), 859–866. <https://doi.org/10.1111/j.1745-9125.1998.tb01268.x>
- Patrick, H., Kaplan, A., & Ryan, A. M. (2011). Positive classroom motivational environments: Convergence between mastery goal structure and classroom social climate. *Journal of Educational Psychology*, *103*(2), 367–382. <https://doi.org/10.1037/a0023311>
- Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, *99*(1), 83–98. <https://doi.org/10.1037/0022-0663.99.1.83>
- Peetsma, T., van Der Veen, I., Koopman, P., & van Schooten, E. (2006). Class composition influences on pupils' cognitive development. *School Effectiveness and School Improvement*, *17*(3), 275–302. <https://doi.org/10.1080/13803610500480114>
- Peetsma, T., Wagenaar, E., & de Kat, E. (2001). School motivation, future time perspective and well-being of high school pupils in segregated and integrated schools in the Netherlands and the role of ethnic self-description. In J. K. Koppen, I. Lunt, & C. Wulf (Eds.), *Education in Europe, cultures, values, institutions in transition* (pp. 54–74). Waxman.
- Peters, D., & Walraven, G. (2011). The Netherlands: Interventions to counteract school segregation. In *International perspectives on countering school segregation* (pp. 131–150). Garant.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. <https://doi.org/10.1037/0022-3514.90.5.751>
- Plenty, S., & Jonsson, J. O. (2017). Social exclusion among peers: The role of immigrant status and classroom immigrant density. *Journal of Youth and Adolescence*, *46*(6), 1275–1288. <https://doi.org/10.1007/s10964-016-0564-5>
- Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century. The 2006 Johan Skytte prize lecture. *Scandinavian Political Studies*, *30*(2), 137–174. <https://doi.org/10.1111/j.1467-9477.2007.00176.x>
- Rjosk, C., Richter, D., Lüdtke, O., & Eccles, J. S. (2017). Ethnic composition and heterogeneity in the classroom: Their measurement and relationship with student outcomes. *Journal of Educational Psychology*, *109*(8), 1188–1204. <https://doi.org/10.1037/edu0000185>
- Rock, P. F., Cole, D. J., Houshyar, S., Lythcott, M., & Prinstein, M. J. (2011). Peer status in an ethnic context: Associations with African American adolescents' ethnic identity. *Journal of Applied Developmental Psychology*, *32*(4), 163–169. <https://doi.org/10.1016/j.appdev.2011.03.002>
- Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development*, *76*(2), 451–466. <https://doi.org/10.1111/j.1467-8624.2005.00856.x>
- Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, *38*(2), 437–460. <https://doi.org/10.3102/00028312038002437>
- Sakiz, G., Pape, S. J., & Hoy, A. W. (2012). Does perceived teacher affective support matter for middle school students in mathematics classrooms? *Journal of School Psychology*, *50*(2), 235–255. <https://doi.org/10.1016/j.jsp.2011.10.005>
- Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, *22*(1), 1–15. [https://doi.org/10.1002/\(SICI\)1098-2337\(1996\)22:1%3c1::AID-AB1%3e3.0.CO;2-T](https://doi.org/10.1002/(SICI)1098-2337(1996)22:1%3c1::AID-AB1%3e3.0.CO;2-T)

- Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*, 22(2), 127–160. [https://doi.org/10.1016/S1098-2140\(01\)00133-3](https://doi.org/10.1016/S1098-2140(01)00133-3)
- Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163, 688. <https://doi.org/10.1038/163688a0>
- Sincer, L., Volman, M., van der Veen, I., & Severiens, S. (2021). The relationship between ethnic school composition, school diversity climate and students' competences in dealing with differences. *Journal of Ethnic and Migration Studies*, 47(9), 2039–2064. <https://doi.org/10.1080/1369183X.2020.1846508>
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4), 571–581. <https://doi.org/10.1037/0022-0663.85.4.571>
- Snijder, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.
- Statistics Netherlands. (2016). Jaarrapport integratie 2016 [Yearly report on integration 2016]. In *Sociaal en Cultureel Planbureau*. <https://www.cbs.nl/nl-nl/publicatie/2016/47/jaarrapport-integratie-2016>.
- Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology*, 51(3), 165–179. <https://doi.org/10.1027/1618-3169.51.3.165>
- Stevens, G. W. J. M., Veldkamp, C., Harakeh, Z., & Laninga-Wijnen, L. (2020). Associations between ethnic minority status and popularity in adolescence: The role of ethnic classroom composition and aggression. *Journal of Youth and Adolescence*, 49(3), 605–617. <https://doi.org/10.1007/s10964-020-01200-6>
- Stewart, K. (2016). Examining student and teacher perceptions of the classroom social environment across school context: Effects of individual factors. In *Graduate Theses and Dissertations*. University of South Florida.
- Thijs, J. (2017). Children's evaluations of interethnic exclusion: The effects of ethnic boundaries, respondent ethnicity, and majority in-group bias. *Journal of Experimental Child Psychology*, 158, 46–63. <https://doi.org/10.1016/j.jecp.2017.01.005>
- Thijs, J., & Verkuyten, M. (2016). Ethnic attitudes and social projection in the classroom. *Child Development*, 87(5), 1452–1465. <https://doi.org/10.1111/cdev.12597>
- Thijs, J., Verkuyten, M., & Grundel, M. (2014). Ethnic classroom composition and peer victimization: The moderating role of classroom attitudes. *Journal of Social Issues*, 70(1), 134–150. <https://doi.org/10.1111/josi.12051>
- van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, 47(2), 497–527. <https://doi.org/10.3102/0002831209353594>
- van der Linden, D., Scholte, R. H. J., Cillessen, A. H. N., & Segers, E. (2010). Classroom ratings of likeability and popularity are related to the Big Five and the general factor of personality. *Journal of Research in Personality*, 44(5), 669–672. <https://doi.org/10.1016/j.jrp.2010.08.007>
- van Leest, A., Hornstra, L., van Tartwijk, J., & van de Pol, J. (2021). Test- or judgement-based school track recommendations: Equal opportunities for students with different socio-economic backgrounds? *British Journal of Educational Psychology*, 91(1), 193–216. <https://doi.org/10.1111/bjep.12356>
- van Vemde, L., Hornstra, L., & Thijs, J. (2021). Classroom predictors of national belonging: The role of interethnic contact and teachers' and classmates' diversity norms. *Journal of Youth and Adolescence*, 50(8), 1709–1725. <https://doi.org/10.1007/s10964-021-01430-2>
- Vervoort, M. H. M., Scholte, R. H. J., & Overbeek, G. (2010). Bullying and victimization among adolescents: The role of ethnicity and ethnic composition of school class. *Journal of Youth and Adolescence*, 39(1), 1–11. <https://doi.org/10.1007/s10964-008-9355-y>
- Vezzali, L., Capozza, D., Giovannini, D., & Stathi, S. (2012). Improving implicit and explicit intergroup attitudes using imagined contact: An experimental intervention with elementary school children. *Group Processes and Intergroup Relations*, 15(2), 203–212. <https://doi.org/10.1177/1368430211424920>
- Wagner, U., Christ, O., & Pettigrew, T. F. (2008). Prejudice and group-related behavior in Germany. *Journal of Social Issues*, 64(2), 403–416. <https://doi.org/10.1111/j.1540-4560.2008.00568.x>
- Wang, M., Te, L., Degol, J., Amemiya, J., Parr, A., & Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Developmental Review*, 57, 100912. <https://doi.org/10.1016/j.dr.2020.100912>

- Wilson, T., Karimpour, R., & Rodkin, P. C. (2011). African American and European American students' peer groups during early adolescence: Structure, status, and academic achievement. *Journal of Early Adolescence*, *31*(1), 74–98. <https://doi.org/10.1177/0272431610387143>
- Wilson, T., & Rodkin, P. C. (2011). African American and European American children in diverse elementary classrooms: Social integration, social status, and social behavior. *Child Development*, *82*(5), 1454–1469. <https://doi.org/10.1111/j.1467-8624.2011.01634.x>
- Zick, A., Pettigrew, T. F., & Wagner, U. (2008). Ethnic prejudice and discrimination in Europe. *Journal of Social Issues*, *64*(2), 233–251. <https://doi.org/10.1111/j.1540-4560.2008.00559.x>

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

Lian van Vemde is a PhD Candidate at the department of Interdisciplinary Social Science at ERCOMER/Utrecht University (The Netherlands). Her major research interests include student (ethnic) diversity, intergroup relations, student and teacher diversity attitudes, and teacher-student relationships.

Jochem Thijs is an Associate Professor at the department of Interdisciplinary Social Science at ERCOMER/Utrecht University (The Netherlands). His research interest is in diversity attitudes and intergroup relations, particularly among youth and in educational contexts.

Lisette Hornstra is an Associate Professor at the department of Education at Utrecht University (The Netherlands). Her major research interests include student diversity, educational equality, and student motivation.