Humanwissenschaftliche Fakultät

Miriam Schwarzenthal | Maja K. Schachner | Linda P. Juang
Fons J. R. Van De Vijver

Reaping the benefits of cultural diversity

Classroom cultural diversity climate and students' intercultural competence

RESEARCH ARTICLE

Reaping the benefits of cultural diversity: Classroom cultural diversity climate and students' intercultural competence

Miriam Schwarzenthal[1] | Maja K. Schachner[1] | Linda P. Juang[1] |
Fons J. R. van de Vijver[2,3,4,5]*

[1]University of Potsdam, Potsdam, Germany

[2]Tilburg University, Tilburg, The Netherlands

[3]North-West University, Potchefstroom, South Africa

[4]University of Queensland, St Lucia, Qld, Australia

[5]Higher School of Economics, Moscow, Russia

Correspondence
Miriam Schwarzenthal, University of Potsdam, Potsdam, Germany.
Email: miriam.schwarzenthal@uni-potsdam.de

Abstract

Culturally diverse schools may constitute natural arenas for training crucial intercultural skills. We hypothesized that a classroom cultural diversity climate fostering contact and cooperation and multiculturalism, but not a climate fostering color-evasion, would be positively related to adolescents' intercultural competence. Adolescents in North Rhine-Westphalia (N = 631, M_{age} = 13.69 years, 49% of immigrant background) and Berlin (N = 1,335, M_{age} = 14.69 years, 52% of immigrant background) in Germany reported their perceptions of the classroom cultural diversity climate and completed quantitative and qualitative measures assessing their intercultural competence. Multilevel structural equation models indicate that contact and cooperation, multiculturalism, and, surprisingly, also color-evasion (as in emphasizing a common humanity), were positively related to the intercultural competence of immigrant and non-immigrant background students. We conclude that all three aspects of the classroom climate are uniquely related to aspects of adolescents' intercultural competence and that none of them may be sufficient on their own.

KEYWORDS
color-evasion, contact, intercultural competence, multiculturalism, school climate

1 | INTRODUCTION

Adolescents growing up in culturally diverse societies encounter a multitude of worldviews, perspectives, and lifestyles on a daily basis. In order to ensure that the human rights of individuals from all cultural backgrounds are acknowledged and respected, tolerance and understanding of different worldviews and lifestyles needs to be promoted from an early age (Barrett, Byram, Lázár, Mompoint-Gaillard, & Philippou, 2013). These skills are usually labeled *intercultural competence* (Deardorff, 2004).

Intercultural competence goes beyond intergroup attitudes by also comprising awareness and knowledge of different worldviews,

as well as the behavioral flexibility to deal with these. While plenty of social psychological contact research has examined intergroup attitudes, little research has empirically examined intercultural competence in the context of culturally diverse societies.

Intercultural competence may be acquired by engaging in prolonged intercultural contact and by actively reflecting on different worldviews (e.g., Allport, 1954; Gurin, Dey, Hurtado, & Gurin, 2002). While the positive social learning outcomes of these experiences have been well documented among students studying abroad, attending intercultural trainings or international colleges (American Field Service [AFS], 2012; Busse, Riedesel, & Krause, 2017; Gurin et al., 2002), the potential of culturally diverse schools to foster

---

*Deceased June 1, 2019

intercultural competence has largely been neglected by the public debate and by psychological research. Instead, in countries like Germany, media and research often focus on the challenges that go along with cultural diversity in schools (Erdmann, 2017; Stanat, 2006).

However, culturally diverse schools may constitute a natural arena for training intercultural skills. Adolescence is a crucial developmental period for acquiring intergroup attitudes and intercultural skills (Quintana, 1998; Raabe & Beelmann, 2011) and schools are a major context for cross-cultural encounters (Sam & Oppedal, 2003). How schools approach cultural diversity may determine whether it mainly poses a threat and a challenge or an opportunity for exchange and personal growth. Therefore, we will investigate how three types of classroom cultural diversity climate (fostering contact and cooperation, color-evasion, and multiculturalism) are related to the intercultural competence of students attending culturally diverse secondary schools in Germany.

1.1 | What is intercultural competence?

Defining intercultural competence requires clarifying the underlying term "culture" first. Culture consists of different layers, for example, of material culture (e.g., food or dress), social culture (e.g., social rules or language), and subjective culture (e.g., attitudes or values) (Chiu & Hong, 2013). Culture is shared by a collective of people, such as a nation, a generation, or a social class. People usually only adopt their collective's culture to a certain degree, and often adopt elements of more than one culture. An individual's compilation of cultural influences determines what is perceived as normative and thus influences how that person interprets, judges, and reacts to others' thinking and behavior (Barrett et al., 2013; Spencer-Oatey, 2008).

Hundreds of characteristics have been brought forward to describe intercultural competence (Deardorff, 2004). The concept of Cultural Intelligence (CQ) (Ang et al., 2007; Van Dyne et al., 2012) was chosen for the present research because it focuses on characteristics that can be learned, rather than stable personality traits, and its measurement has been deemed one of the most reliable and valid to date (Matsumoto & Hwang, 2013). Individuals possessing high CQ enjoy interacting with others who have different cultural affiliations (motivational CQ), know about norms, values, and behaviors in different cultures (cognitive CQ), are aware that culture may influence their own and others' behavior and thinking (metacognitive CQ), and use appropriate behavior in culturally diverse situations (behavioral CQ).

Even before entering school, children are aware of social group identities and become increasingly aware of which attitudes or behaviors are considered to be normative in a group (Rutland & Killen, 2015). Adolescence is a crucial time for the development of intercultural skills as intergroup attitudes are strongly influenced by the social context and adolescents gradually learn to take other's perspectives (Van der Graaff et al., 2014), and to understand more subtle cultural influences on behavior (Quintana, 1998; Schwarzenthal, Juang, Schachner, van de Vijver, & Handrick, 2017).

However, awareness of cultural influences does not mean that all behavior should be explained with cultural factors. Preconceived attributions to culture can also foster stereotypes and lead to essentialism (Barrett et al., 2013). Therefore, in order to develop high metacognitive CQ, students should also learn to ask questions, suspend judgments, and consider alternative explanations for behavior (Sieck, Smith, & Rasmussen, 2013). Completely neglecting culture, on the other hand, may result in a "color-evasive" perspective where important information that might aid interpretation of behavior is neglected (Schofield, 2010).

Intercultural competence is often assessed using self-report measures, such as the cultural intelligence scale (CQS) (Van Dyne et al., 2012). However, these measures are not tailored to adolescents and have been criticized because people might not be able to accurately report their own competences (Klafehn, Li, & Chiu, 2013). In the face of calls for a multimodal assessment of intercultural competence (Deardorff, 2011), situational judgment tests (SJTs) have been used as an alternative assessment method. These are based on a short description of an intercultural situation followed by questions assessing the participants' response to or interpretation of the situation (Rockstuhl, Ang, Ng, Lievens, & Van Dyne, 2015). In our study we use a version of the CQS that was adapted for adolescents, and SJTs set in the school and peer context.

1.2 | How may intercultural competence be fostered in schools?

The presence of a high proportion of outgroup members in the environment alone does not automatically promote intercultural learning. A large proportion of outgroup members may induce feelings of threat and thus lead to more negative intergroup attitudes (Schmid, Al Ramiah, & Hewstone, 2014; Taylor, 1998) or may increase opportunities for intergroup contact (Blau, 1977) and thus lead to more positive intergroup attitudes (Allport, 1954; Pettigrew & Tropp, 2008) and intercultural learning (Gurin et al., 2002). Studies examining relations between the ethnic composition of college or high school classrooms and intergroup attitudes and intercultural learning have come to mixed conclusions—level of classroom diversity was sometimes unrelated, and sometimes positively related to intercultural learning outcomes (e.g., Denson & Chang, 2008; Rothman, 2003; Schwarzenthal, Schachner, Van de Vijver, & Juang, 2018; Terenzini, Cabrera, Colbeck, Bjorklund, & Parente, 2001). It has been argued that the way schools or classrooms approach cultural diversity is crucial in determining intergroup attitudes and intercultural skills (Schwarzenthal et al., 2018; Thijs & Verkuyten, 2014).

Three types of approaches to cultural diversity in schools may be especially relevant for students' intercultural learning: fostering positive interactions between students of diverse cultural affiliations (i.e., fostering an approach of contact and cooperation), promoting the idea that cultural differences are not important (i.e., fostering an approach of color-evasion), and/or actively promoting engagement with students' heritage cultures and different perspectives (i.e., fostering an approach of multiculturalism).

1.2.1 | Contact and cooperation

Fostering contact and cooperation was proposed in the 1950s as a strategy to foster positive interracial attitudes among students attending recently desegregated schools (Allport, 1954). Since then, a large body of research has provided support for the claim that more intergroup contact goes along with more positive intergroup attitudes, especially if it occurs under certain conditions (the groups share equal status and common goals, they cooperate, and the contact is supported by an authority) (Allport, 1954; Levy & Killen, 2010; Pettigrew & Tropp, 2006). In addition to direct personal contact, social norms that support positive contact are crucial in determining people's attitudes (Christ et al., 2014). Schools may constitute an ideal context for setting positive contact norms, for example when students help each other with homework or work cooperatively in class (Dejaeghere, Hooghe, & Claes, 2012; Schachner, Noack, Van de Vijver, & Eckstein, 2016). Promotion of positive contact and cooperation is a key dimension of most conceptualizations of classroom cultural diversity climate, and is subsumed under broader labels such as "interpersonal interactions" (Byrd, 2017), or "equality and inclusion" (Schachner, Schwarzenthal, Moffitt, Civitillo, & Juang, 2019; Schachner et al., 2016).

In classrooms in which positive contact and cooperation between students of different cultural affiliations is supported, students tend to display more positive outgroup orientations (Schwarzenthal et al., 2018), more intercultural friendships (Schachner, Brenick, Noack, van de Vijver, & Heizmann, 2015), and less prejudice (Molina & Wittig, 2006). However, research examining relations between intercultural contact in culturally diverse schools and intercultural competence is scarce. The few existing studies found that the length of time students had attended an international high school was related to their level of intercultural sensitivity (Straffon, 2003), and that students' individually reported intercultural contact at school was related to their CQ (Schwarzenthal et al., 2017). Research conducted on culturally diverse college campuses lends further support to the claim that intercultural contact is not only related to attitudes, but also to intercultural knowledge, understanding, and interculturally competent behavior (Bowman, 2010; Gurin et al., 2002; Jon, 2013). Based on these research findings, we formulate the following hypothesis:

Hypothesis 1: *A classroom climate promoting contact and cooperation between students of diverse cultural affiliations is related to higher motivational, cognitive, metacognitive, and behavioral CQ among students.*

1.2.2 | *Color-evasion*

Systematic efforts to foster contact and cooperation between students can go along with neglecting cultural variations—and thus pursuing a color-evasive approach (Schachner et al., 2016).[1] Color-

---

[1]The concept of color-evasion is comparable to what is usually captured by the term "colorblindness." However, in order to avoid a deficit view on people with disabilities, and to acknowledge that not seeing race/ethnicity/culture is a conscious choice, we prefer the term "color-evasion" in this research.

evasive approaches were originally based on the assumption that prejudice and discrimination result from an emphasis on group categories and that therefore group membership should be deemphasized and category boundaries eliminated (Rosenthal & Levy, 2010). These approaches not only informed social psychological interventions (Gaertner & Dovidio, 2000), but were also adopted by schools (Schofield, 2010). Schools adopting color-evasive approaches emphasize that racial or cultural categories are irrelevant and treat social interactions as interpersonal and not intergroup processes. These approaches have been strongly criticized because they tend to go along with ignorance of existing inequalities and structural discrimination (i.e., with "power-evasion", Neville, Lilly, Duran, Lee, & Browne, 2000). In conceptualizations of classroom cultural diversity climate, the respective constructs are labeled "colorblind socialization" (Byrd, 2017) or "color-evasion" (Schachner et al., 2019).

While power-evasion has an inherently negative connotation, research on color-evasion (as in emphasizing a common humanity), suggests that it has positive as well as negative effects. Color-evasive approaches predict positive intergroup attitudes and behaviors such as helping (for a review, see Dovidio, Gaertner, & Saguy, 2010), and may reduce intergroup anxiety (Schofield, 2010). However, relinquishing group identities altogether may also compromise people's need for distinctiveness, and its effects on positive intergroup attitudes are less stable than expected (Dovidio et al., 2010). Color-evasion promotes suppression of negative thoughts in the short term, but preexisting prejudice and negative behavior can rebound in the long term (Correll, Park, & Smith, 2008; Sasaki & Vorauer, 2013). Moreover, researchers criticizing this approach have argued that a complete neglect of racial and cultural categories implies that opportunities to learn about cultural diversity are not being used (Schofield, 2010), existing discrimination is not recognized (Apfelbaum, Pauker, Sommers, & Ambady, 2010), and appropriate behavioral strategies for interracial or intercultural interactions are not developed (Sasaki & Vorauer, 2013). After 4-year-long observations at a middle school in the northeastern U.S., Schofield (2010, p. 287) concludes that "such a tendency, while undeniably a low-risk one, failed to take advantage of the diversity of experiences and perspectives […] as a resource for the educational process." The concept of color-evasion bears similarities to the concept of minimization in Bennett's developmental model of intercultural sensitivity (Bennett, 1993). People in the minimization stage believe that all humans are similar, which is part of the lower, or "ethnocentric" stages of intercultural sensitivity, as opposed to the "ethnorelative" stages in which people start to accept cultural variations.

The color-evasion scale used for this research (Schachner et al., 2019) captures an emphasis on a common humanity, but does not directly assess ignorance of discrimination and structural inequalities (i.e., "power-evasion"). Since color-evasion has some potential to reduce prejudice in the short term, but not in the long term, and entails that intercultural learning opportunities are not being used, we formulate the following hypothesis:

Hypothesis 2: *While a color-evasive classroom climate may be related positively to motivational CQ, we expect that it is unrelated to cognitive, metacognitive, and behavioral aspects of CQ among students.*

1.2.3 | Multiculturalism

According to the multicultural approach, race, ethnicity, and culture should be paid attention to, group differences should be valued, and individuals should learn about the perspectives of various groups in society (Sasaki & Vorauer, 2013). Supporters of the multicultural approach argue that in order to reduce prejudice, it is not necessary to remove intergroup categories. Instead, it is assumed that learning about and critically reflecting on diversity can reduce prejudice (Park & Judd, 2005; Rosenthal & Levy, 2010), and foster intercultural competence (Barrett, 2018). In education, a multicultural approach has become increasingly popular since the 1970s and 1980s (Schachner, 2017). For example, Banks (2015), one of the leading scholars in multicultural education, proposes that schools should not only focus on prejudice reduction, but should also include cultural content in the curriculum, and foster understanding of implicit cultural assumptions and perspectives. In conceptualizations of classroom cultural diversity climate, multicultural approaches are often subsumed under labels such as "cultural socialization", "promotion of cultural competence" (Byrd, 2017), or "cultural pluralism" (Schachner et al., 2016, 2019).

Multiculturalism is associated with positive intergroup attitudes and lower ethnocentrism (Rosenthal & Levy, 2010; Schwarzenthal et al., 2018; Verkuyten & Thijs, 2013). However, when a multicultural approach puts too much emphasis on differences (and essentializes these differences), it can also increase stereotyping and reduce perceived similarity and liking (Rosenthal & Levy, 2010; Schofield, 2010). Indeed, multiculturalism in schools is also related to more perceived discrimination (Schwarzenthal et al., 2018).

Even though empirical research assessing the impact of multicultural education at school on intercultural knowledge, understanding, and interculturally competent behavior is scarce, theories and previous research from other areas suggest that multicultural education may positively affect these aspects of intercultural competence. A multicultural ideology encourages active engagement with and learning about diversity (Crisp & Turner, 2011; Sasaki & Vorauer, 2013). Active engagement with diversity does not only increase cultural knowledge, but can also shape the structure of our thoughts, as engagement with new information that does not fit into one's schema drives the learning process and fosters cognitive development (Piaget, 1977). This idea informed researchers in acculturation psychology and higher education who assume that engaging with diverse perspectives results in a higher complexity of cultural representations (Tadmor & Tetlock, 2006), and promotes students' intercultural learning and cognitive growth (Gurin et al., 2002). Indeed, norms around multiculturalism (as compared to color-evasion) are related to increased perspective-taking (Todd & Galinsky, 2012) and more positive other-directed remarks (Vorauer, Gagnon, & Sasaki, 2009). Learning about cultural diversity in college is

related to a range of cognitive, metacognitive, and behavioral intercultural learning outcomes (Bowman, 2010; Gurin et al., 2002; Maddux, Bivolaru, Hafenbrack, Tadmor, & Galinsky, 2013). Thus, we formulate the following hypothesis:

Hypothesis 3: *A multicultural classroom climate is related to higher motivational, cognitive, metacognitive, and behavioral CQ among students.*

1.2.4 | Relations between the three approaches

The diversity approaches described above are not mutually exclusive (Rosenthal & Levy, 2010). It is possible that positive contact is essential for anxiety to be reduced before people can undertake the cognitive effort to engage in perspective-taking and gaining knowledge about the outgroup (i.e., before they can profit from multicultural approaches) (Richeson & Shelton, 2003). A focus on a common humanity (as promoted by color-evasive approaches) may help to lessen risks of stereotyping and "othering", while learning about cultural variations (as in multicultural approaches) can increase outgroup knowledge and perspective-taking. Despite these theoretical considerations, relatively few studies have directly compared the different types of classroom cultural diversity climate in one study (Rosenthal & Levy, 2010), and none have looked at relations with students' intercultural competence.

1.3 | Individually perceived and classroom-aggregated cultural diversity climate

In order to assess classroom climate, most researchers ask students for their perceptions, using the classroom or the teacher as a referent (Lam, Ruzek, Schenke, Conley, & Karabenick, 2015; Wang & Degol, 2016). The individual student ratings are then aggregated at the classroom level to form measures of classroom climate (e.g., see Marsh et al., 2012). However, in many studies students' perceptions of their classroom's climate are quite heterogeneous (Lam et al., 2015; Miller & Murdock, 2007). This may be due to differential treatment that students receive in the same classroom, and to students' idiosyncratic interpretations of experiences (Wang & Degol, 2016). For example, cultural minority members tend to perceive a more negative climate of contact and cooperation in the classroom than cultural majority members (Byrd, 2014; Schwarzenthal et al., 2018). Consequently, effects of individually perceived classroom diversity climate on intergroup and psychological outcomes are usually stronger than the effects of the aggregated diversity climate (Schachner et al., 2016; Schwarzenthal et al., 2018). In this research, we will therefore investigate effects of the individually perceived as well as of the classroom-aggregated climate.

1.4 | Differential effects between students of immigrant and non-immigrant background

The effect of cultural diversity climate on CQ may differ between different groups of students. In Germany, about one-third of the student population is of immigrant background, meaning that at least one of

their parents did not acquire German citizenship at birth (Statistisches Bundesamt, 2018). Thus, these students are in the minority in the society. They are more likely to experience discrimination (Frankenberg, Kupper, Wagner, & Bongard, 2013) and typically have lower socioeconomic status than students of non-immigrant background (Kristen & Granato, 2007). If for them intercultural interactions are marked by status differences and experiences of discrimination, it is less likely that they will lead to the development of CQ, because conflict can hinder positive intercultural learning outcomes (Correll et al., 2008).

Moreover, students of immigrant background encounter cultural diversity on a daily basis. While the family and same-ethnic peer contexts are typically more influenced by their ethnic culture, the school and different-ethnic peer contexts are more influenced by the mainstream culture (Motti-Stefanidi, Berry, Chryssochoou, Sam, & Phinney, 2012). For students of non-immigrant background on the other hand, intercultural experiences may be less typical, implying that these experiences may have a stronger impact on their intergroup attitudes and intercultural skills if they occur (Denson & Zhang, 2010; Loes, Pascarella, & Umbach, 2012; Tropp & Pettigrew, 2005). Thus, we formulate the following hypothesis:

Hypothesis 4: *Relations between contact and cooperation and multiculturalism on the one hand and CQ on the other hand are stronger among students of non-immigrant background than among those of immigrant background.*

1.5 | The present research

In this research, we assess how three types of classroom cultural diversity climate (contact and cooperation, color-evasion, and multiculturalism) are related to CQ among adolescents of immigrant and non-immigrant background attending culturally diverse schools in Germany.[2] In Germany, the notion that cultural diversity can be beneficial was only officially recognized in the 1990s, when new guidelines by the Conference of the Ministers of Education of the German Federal States compelled schools to incorporate intercultural learning into the curriculum for all children (Faas, 2008). However, until today German schools focus more on prejudice reduction and equality than on multiculturalism and intercultural learning (Civitillo et al., 2016; Schachner et al., 2016).

We base our analyses on data collected in two culturally diverse regions in Germany: Study 1 is based on data from 631 6th to 10th graders in 29 classrooms in North Rhine-Westphalia. Using this sample, we investigate relations between two cultural diversity approaches in the classroom (contact and cooperation and multiculturalism) and students' CQ. In Study 2, we try to replicate and expand the results from Study 1 with a sample of 1,335 9th graders in 66 classrooms in Berlin. Since these students were also asked to report their perceptions of color-evasion in the classroom, this

dataset allows disentangling effects of contact and cooperation and color-evasion. Moreover, the larger number of classrooms enables better investigation of relations with the classroom-aggregated cultural diversity climate.

To test our hypotheses, we use a multiple indicators, multiple causes structural equation model (see for example van de Vijver, 2002), which links input (i.e., classroom cultural diversity approaches) and output (i.e., different facets of CQ or SJT) through a latent variable (i.e., CQ or SJT). This approach also allows detecting direct relations between predictor variables and the different CQ or SJT facets through modification indices.

We combine this approach with a multilevel framework which allows taking the clustered nature of the data into account, and investigating effects of the classroom-aggregated cultural diversity climate. Thus, our approach is similar to the doubly latent framework suggested by Marsh et al. (2012) to study school context and climate effects, with the exception that only the outcome variables, not the predictor variables, were modeled as latent. We decided to only model the outcome variables as latent variables because in multilevel models, the sample size at the classroom level is determined by the number of clusters. If all variables are modeled as latent variables, the number of parameters to be estimated at the classroom level quickly exceeds the numbers of clusters, and the models get too complex.

We included gender, immigrant background, grade average (i.e., average of grades in Math, English, and German, recoded so that higher scores reflect better grades), number of books in the household (5-point Likert scale from (1) *none or very few* to (5) *more than 200 books*, Bos et al., 2003), and the proportion of students of immigrant background in a classroom as control variables in both studies, since they are related to intercultural competence among adolescents (Schwarzenthal et al., 2017; Terenzini et al., 2001).[3]

2 | STUDY 1

2.1 | Method

2.1.1 | Participants and procedure

After obtaining permission from school principals, parental approval, and students' assent or consent, we administered a questionnaire in 29 6th, 8th, and 10th grade classrooms in seven schools in North Rhine-Westphalia, Germany in the winter of 2015/2016 as part of a larger cross-national study on inclusive identity. A total of 631 students (48.2% female, M_{age} = 13.69 years, SD_{age} = 1.83,

---

[2]We conducted additional analyses looking at the students' cultural self-identification (with the cultural majority vs. with a cultural minority) as a moderator. However, cultural self-identification did not moderate any of the relations in our models.

[3]Re-running the analyses without any control variables did not substantially alter the results. We used the number of books in the household as a control variable for several reasons: Direct information from the parents on their income or education was not available in our study, and children are often unable to report their parent's educational background or occupation correctly (for a review, see Currie et al., 2008). Even though in Study 2, we had also included the Family Affluence scale (Boyce et al., 2006; German version by Richter & Leppin, 2007), the number of books was consistently more strongly related to CQ than the Family Affluence Scale. Therefore, the number of books was deemed to be the most relevant control variable reflecting cultural capital in the family. However, future studies should take into account that more and more people are reading books online (which was less the case in 2015/2016 when our data were collected).

range$_{age}$ = 11–18 years) completed the survey. Of the participating students, 49% were of immigrant background, meaning that at least one parent had immigrated to Germany. Most of these students (86%) were born in Germany. They represented 63 heritage countries, with the largest group from Turkey. The proportion of students of immigrant background in the classroom ranged between 13% and 89%. All students attended the most comprehensive school type in North Rhine-Westphalia (the *Gesamtschule*) that offers various school leaving certificates.

2.2 | Measures

2.2.1 | Cultural intelligence

We used a CQ measure that was specifically developed for adolescents in culturally diverse societies, and that contains a self-report questionnaire as well as SJTs (Schwarzenthal, Juang, Schachner, & van de Vijver, 2019)

Self-reported CQ

The original self-report CQ scale (Ang et al., 2007; Van Dyne et al., 2012) contains four subscales assessing motivational, cognitive, metacognitive, and behavioral CQ, and was related to a range of positive outcomes, such as cultural adaptation and cross-cultural negotiation effectiveness, in previous research (for reviews, see Ott & Michailova, 2018; Sharma & Hussain, 2017). Maintaining the subscales of the original scale, items were developed that were understandable and relevant for adolescents growing up in multicultural contexts. CQ motivation (e.g., "It's fun for me to interact with people from other cultures", α = .88), CQ cognition (e.g., "I can describe what is expected of men and women in various cultures", α = .88), CQ metacognition (e.g., "If I don't understand the behavior of people from another culture, I try to find out why they might have acted the way they did", α = .86), and CQ behavior (e.g., "When I talk to people from another culture, I am considerate regarding their traditions and ways of living", α = .82) were measured with six items each. The response scale ranged from (1) *no, that's not right* to (5) *yes, that's right*. Items from each subscale were averaged to represent the four CQ facets.

Situational judgment tests

In order to assess intercultural competence a multimodal assessment is desirable (Deardorff, 2011). Therefore, Schwarzenthal et al. (2019) also developed SJTs for adolescents in culturally diverse societies, based on the example set by Rockstuhl et al. (2015). The scores that participants received in the original SJTs predicted task performance and organizational citizenship behavior in culturally diverse teams (Rockstuhl et al., 2015). Two SJTs were developed that were situated in the peer (SJTa) and school (SJTb) context. Brief descriptions of intercultural situations were followed by questions assessing situational judgment ("Why does person A/B behave this way?") as well as response judgment ("What would you do next in this situation?").

Three intercultural researchers developed a coding manual (following Syed & Nelson, 2015), partly based on Rockstuhl et al.

(2015) and a study by Sieck et al. (2013), but also drawing on the data at hand to develop appropriate coding categories (for the full coding manual, see Schwarzenthal et al., 2019). Two independent researchers coded the answers. The students' situational judgment was coded with regard to consideration of cultural influences (3-point-scale ranging from [1] *low* to [3] *high*) and suspension of judgment (3-point-scale ranging from [1] *low* to [3] *high*). Their response judgment was coded with regard to interculturally competent behavior (5-point-scale ranging from [1] *not at all effective* to [5] *very effective*). Please note that there was not only "one" right response, but that instead any response by the students that reflected integration of different cultural interests was rated as interculturally competent (e.g., if the students actively mediated between different parties, or if they proposed a solution that fit both parties' interests). Disagreements between coders were resolved via consensus. Two-way-random intraclass-correlations (Shrout & Fleiss, 1979) confirmed high intercoder reliability (from .86 to .96).

2.2.2 | Cultural diversity climate

The response scale for both subscales of cultural diversity climate ranged from (1) *no, that's not right* to (5) *yes, that's right*. For both subscales, we included students' individual perceptions, as well as the classroom-level aggregates of these perceptions in our analyses. Since students in Germany spend most of their time at school in one classroom, we chose to aggregate at the classroom level, and not at the school level.

Contact and cooperation

Based on Beaton, Dovidio, and Léger (2008), four self-developed items measured positive interactions between students of different cultural backgrounds in the classroom, such as "Students from my cultural group have good relationships with students from other cultural groups" (α = .75). Since this measure was used for the first time in this study, we ran a CFA in MPlus (Muthén & Muthén, 1998–2011) to confirm the one-dimensional factor structure. The CFA revealed an adequate fit of the one-factorial model (with one correlated error because of similar item wording), χ^2/df (N = 591) = 5.22, p = .02, RMSEA = .09, CFI = .99, SRMR = .02.

Multiculturalism

Multiculturalism was assessed with the cultural pluralism scale developed by Schachner et al. (2016), which contains three subscales : (a) the perceived interest of students and teachers in students' cultural background (six items; e.g., "At school students are interested in how people from different cultures and countries live"); (b) learning about intercultural relations (three items; e.g., "At school we learn to respect each other even though we are from different cultures"); and (c) learning about multicultural topics (five items; e.g., "At school we talk about the celebrations and traditions of other cultures"). The response scale ranged from (1) *no, that's not right* to (5) *yes, that's right*. The three subscales were averaged to create an overall score for multiculturalism. Cronbach's α was .92.

2.2.3 | Control variables

Since the students in Study 1 were between 11 and 18 years old, and age is related to CQ (Schwarzenthal et al., 2017), we included age as an additional control variable in Study 1.

2.3 | Results

2.3.1 | Preliminary analyses

We first conducted preliminary analyses. In total, 3% of data were missing from the variables that were used for analyses. We used full information maximum likelihood (FIML) in Mplus 7.3 (Muthén & Muthén, 1998–2011), which is considered best practice for dealing with missing data (Enders, 2010).

2.3.2 | Measurement models at the individual level

As a first step, we specified the measurement models at the individual level. The latent CQ variable was measured by the four means of the self-report CQ motivation, cognition, metacognition, and behavior subscales, and the latent SJT variable by the six ratings that the students received for their responses in the two SJTs. Both measurement models fit well (see Appendix 1). Tests of measurement equivalence between students of immigrant and non-immigrant background confirmed metric invariance (i.e., equivalence of factor loadings) for the self-report CQ factor model and scalar invariance (i.e., equivalence of factor loadings and intercepts) for the SJT factor model (see Appendix 2), implying that for the CQ factor, only relations between constructs, not means, should be compared between students of immigrant and non-immigrant background.

2.3.3 | ICCs and measurement models at the classroom level

As a next step, we estimated intraclass correlations (ICCs) for all study variables (see Appendix 3). The proportion of total variance explained by student membership in different classrooms (i.e., ICC1) of the outcome variables was mostly lower than .05, and the reliability of the observed classroom aggregate (ICC2, as in Marsh et al., 2012) was low. Thus, multilevel modeling procedures may provide few benefits and might run into estimation problems (Dyer, Hanges, & Hall, 2005). Specification of measurement models for the latent self-report CQ and SJT factors at the classroom level resulted in bad model fit, indicating that these constructs could not be reliably modeled at the classroom level (see Appendix 1). Exploratory analyses revealed that none of the classroom level predictors significantly predicted any of the outcome variables. Therefore, we proceeded by only investigating relations at the individual level. However, since even very low ICCs of .05 or .01 can lead to a significant distortion of the results of significance tests in conventional regression analyses (Geiser, 2011; Hox, 2010), we used the two-level option in Mplus 7.3 to take the clustered structure of the data into account even though no predictors were introduced at the classroom level.

2.3.4 | Correlations

As a first step, we estimated correlations between the predictor variables and the self-report CQ subscales and SJT ratings separately for students of immigrant and non-immigrant background (see Table 1). Contact and cooperation and multiculturalism were positively related to the self-reported CQ subscales in both groups, and to some of the SJT ratings among students of non-immigrant background.

2.3.5 | Multilevel SEM

To test our hypotheses, we ran two separate models for the self-report CQ factor and the SJT factor.[4] We built the models step by step. All continuous predictors that were strictly at the individual level were grand-mean centered prior to the analyses. Correlations between predictors were allowed. First, control variables (age, gender, immigrant background, grade average, and number of books in the household) were introduced as predictors at the individual-level. Students of immigrant background (compared to students of non-immigrant background) and students with more books reported higher CQ. Older students, females, and students with higher grade average showed higher scores in the SJTs. The model fit was good for the self-reported CQ model, $\chi^2/df(N = 631) = 2.49$, $p < .001$, RMSEA = .05, CFI = .97, $SRMR_{within} = .03$, AIC = 12,812.95, and adequate for the SJT model, $\chi^2/df(N = 631) = 2.47$, $p < .001$, RMSEA = .05, CFI = .91, $SRMR_{within} = .04$, AIC = 15,187.82.

We then introduced individually perceived contact and cooperation and multiculturalism in the classroom as predictors at the individual level (testing Hypotheses 1 and 3). Both positively predicted self-reported CQ. Only perceived multiculturalism positively predicted SJT scores.

Finally, we introduced interaction terms one by one to test whether the associations between the individually perceived classroom cultural diversity climate and the outcomes were stronger among students of non-immigrant background (testing Hypothesis 4). Multiculturalism was less strongly related to the SJT factor among students of immigrant background ($\beta = -.14$, $p = .03$). None of the other interaction effects were significant. The fit of the final models was good for the self-report CQ outcome and adequate for the SJT outcome. Standardized coefficients and model fit indices of the final models are presented in Figures 1 and 2. Overall, the models explained 38% of variance in the latent self-report CQ variable, and 56% of variance in the latent SJT variable.

2.4 | Discussion

In line with Hypotheses 1 and 3, students who perceived a stronger classroom climate fostering contact and cooperation between students of diverse cultural backgrounds, or a stronger climate fostering multiculturalism, reported higher CQ. However, only a perceived climate of multiculturalism, but not of contact

---

[4]Additional analyses were run with both outcome variables in one model. However, this did not alter the results.
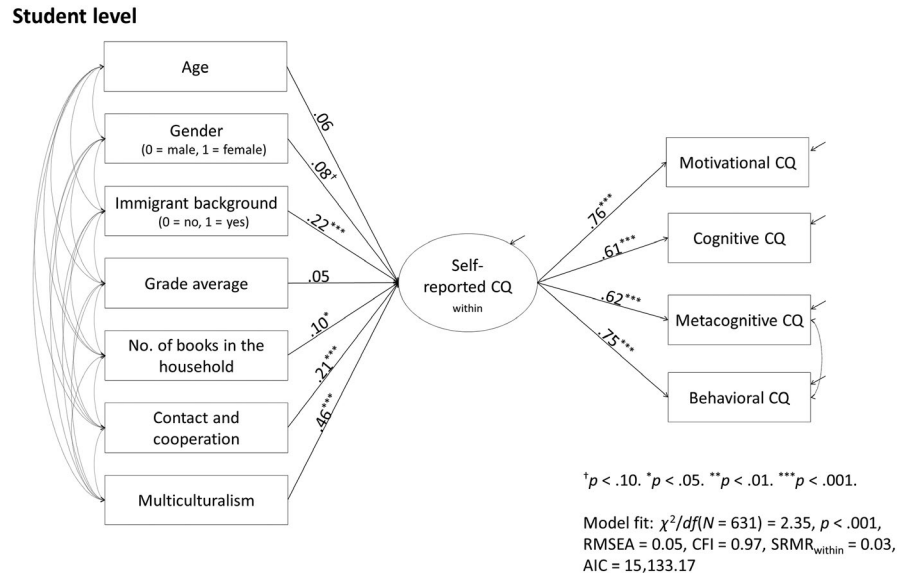
TABLE 1 Study 1 individual level bivariate correlations and descriptives

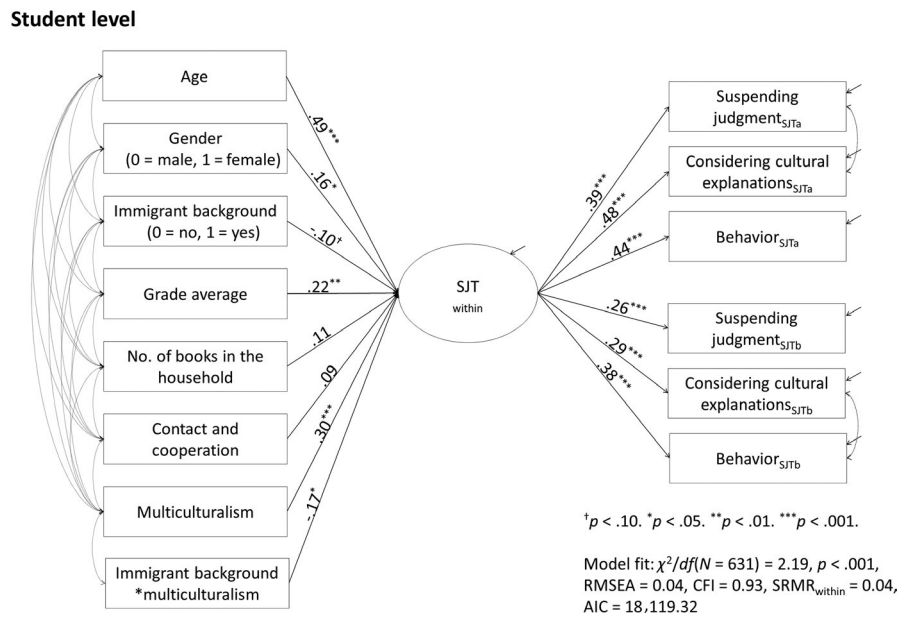|	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1. Age	–	-.08	-.12*	-.13*	-.04	.07	.01	.09†	-.07	-.01	.30***	.24***	.12*	.09	.17**	.03
2. Gender (0 = male, 1 = female)	-.03	–	-.01	-.02	-.04	.08	.11*	-.02	-.00	.01	.08	-.05	.13*	.04	.07	.02
3. Grade average	-.17**	.03	–	.27***	.13*	.08	.09	.07	.06	.13*	.05	.12*	.12*	.02	.02	.20**
4. No. of books in the household	-.11†	.01	.18**	–	.11*	.04	.14*	.10†	.04	.02	.06	.06	-.01	-.01	-.08	.09
5. Contact and cooperation	.01	.02	.21***	.05	–	.24***	.35***	.13†	.17**	.28***	.12*	.05	.12†	.04	-.05	.18**
6. Multiculturalism	.03	.07	.00	-.02	.25***	–	.44***	.36***	.30***	.40***	.12*	.15**	.20**	.11†	.12*	.24***
7. Self-reported CQ motivation	.09†	.17**	.08	.07	.27***	.34***	–	.36***	.49***	.59***	.23***	.11*	.29***	.07	-.06	.22***
8. Self-reported CQ cognition	.07	.07	-.03	.10	.14*	.31***	.48***	–	.42***	.40***	.18**	.14*	.06	.06	.06	.18**
9. Self-reported CQ metacogn.	-.00	.10†	.08	.07	.16*	.36***	.41***	.44***	–	.70***	.12*	.11	.27***	.06	-.03	.23***
10. Self-reported CQ behavior	.05	.11*	.11†	.11†	.23***	.37***	.54***	.50***	.69***	–	.12*	.11	.28***	.07	.01	.27***
11. SJTa suspending judgment	.19**	-.02	.07	-.07	.15*	.01	.02	.05	-.05	-.03	–	.48***	.26***	.09	.03	.10
12. SJTa considering cult. infl.	.35***	.01	.03	.11†	.06	-.03	.03	.06	.00	-.00	.48***	–	.28***	.02	.05	.21***
13. SJTa behavior	.12*	.10†	.03	.07	.08	.09	.16**	.06	.12†	.19**	.12*	.19***	–	.07	.12*	.26***
14. SJTb suspending judgment	.17**	.05	.14*	.09	-.01	.07	.02	.05	-.01	.05	.10	.06	.10†	–	.02	.15**
15. SJTb considering cult. infl.	.24***	.13*	.05	.01	-.05	.04	.13*	.07	-.02	.03	.01	.16**	.21***	.13†	–	.11†
16. SJTb behavior	.07	.14*	.11	.12†	.07	.07	.12*	.10	.11†	.19**	.09	.13*	.20***	.18**	.44***	–
Students of non-immigrant background																
M (SD)	13.53 (1.82)	0.48 (0.5)	4.40 (0.77)	3.44 (1.21)	4.02 (0.71)	3.51 (0.72)	3.73 (0.83)	2.86 (0.86)	3.27 (0.9)	3.51 (0.84)	2.28 (0.75)	1.90 (0.73)	3.41 (0.9)	2.09 (0.47)	1.68 (0.77)	3.03 (1.17)
Students of immigrant background																
M (SD)	13.86 (1.83)	0.50 (0.5)	4.18 (0.76)	2.75 (1.27)	3.96 (0.67)	3.60 (0.70)	3.95 (0.73)	3.24 (0.81)	3.46 (0.85)	3.72 (0.73)	2.24 (0.76)	1.87 (0.74)	3.22 (0.93)	2.05 (0.5)	1.70 (0.81)	2.99 (1.22)

Note: $N = 324$ students of non-immigrant background and $N = 307$ students of immigrant background. Correlations for students of non-immigrant background above the diagonal, correlations for students of immigrant background below the diagonal.
†$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.

FIGURE 1 Study 1: Model with CQ as outcome variable, standardized coefficients

Student level

$^\dagger p < .10.$ $^* p < .05.$ $^{**} p < .01.$ $^{***} p < .001.$

Model fit: $\chi^2/df(N = 631) = 2.35$, $p < .001$, RMSEA = 0.05, CFI = 0.97, SRMR$_{within}$ = 0.03, AIC = 15,133.17

FIGURE 2 Study 1: Model with SJT as outcome variable, standardized coefficients

Student level

$^\dagger p < .10.$ $^* p < .05.$ $^{**} p < .01.$ $^{***} p < .001.$

Model fit: $\chi^2/df(N = 631) = 2.19$, $p < .001$, RMSEA = 0.04, CFI = 0.93, SRMR$_{within}$ = 0.04, AIC = 18,119.32

and cooperation, was related to students' scores in the SJTs. Promoting contact and cooperation can sometimes go along with pursuing a color-evasive approach (Schachner et al., 2016) and neglecting opportunities to learn about cultural variations. In order to explore this possibility further, we explicitly distinguished between a classroom climate fostering contact and cooperation, and a classroom climate fostering color-evasion in Study 2.

Not supporting Hypothesis 4, relations between classroom cultural diversity climate and students' CQ in Study 1 were largely similar between students of immigrant and non-immigrant background. Only the positive link between multiculturalism and performance in the SJTs was stronger among students of non-immigrant background.

Due to low variance at the classroom level, the outcome variables in this study could not be reliably modeled at this level and were unrelated to the classroom-aggregated climate. This may be due to

the fact that students in one classroom usually perceive classroom climate quite differently (e.g., Schenke, Ruzek, Lam, Karabenick, & Eccles, 2017), and that these idiosyncratic experiences are more closely linked to learning outcomes than abstract classroom-level aggregates. However, it may also be a result of the rather small number of classrooms analyzed for this study (N = 29) and the fact that these were all from a single school type (the *Gesamtschule*), and thus may not vary much with regard to their cultural diversity climate and the students' CQ. We explored this possibility by analyzing a larger number of classrooms from different school types in Study 2.

Conceptions of intercultural competence are sometimes criticized for encouraging preconceived attributions to culture, which can also foster stereotypes (Barrett et al., 2013). In order to address this critique, we introduce an additional coding category for the SJTs in Study 2. Previous research has found that individuals with more intercultural

experience are not only more likely to consider cultural influences as explanations for behavior, but also to consider multiple alternative explanations for behavior, thereby reducing the likelihood of stereotyped interpretations of cultural differences (Sieck et al., 2013). Therefore, in Study 2, we will also code whether students consider alternative explanations when interpreting the behavior in the SJTs.

3 | STUDY 2

3.1 | Method

3.1.1 | Participants and procedure

We obtained permission from the Berlin Senate Committee for Education, Youth, and Science and from 17 principals to conduct the study in their schools. Following the guidelines of the Berlin Senate, parental consent was only required for questions asking about the participants' parents. After obtaining parental consent, and students' assent or consent, we administered a questionnaire in 66 9th grade classrooms in 2016. The final sample comprised 1,335 adolescents (48% female, M_{age} = 14.69 years, SD_{age} = 0.74). Of these students, 52% were of immigrant background, 37% of non-immigrant background, and 11% did not receive permission from their parents to answer questions related to their parents' place of birth. Most of the students of immigrant background (74%) were born in Germany. They represented 80 different heritage countries, with the largest group coming from Turkey. The proportion of students of immigrant background in a classroom ranged between 9% and 100%. The students either attended a *Gymnasium* (32% of the sample), which is the academic track, or an Integrated Secondary School (68% of the sample), which combines the former vocational and comprehensive school types, and offers all school leaving certificates.[5]

3.2 | Measures

3.2.1 | Cultural intelligence

As in Study 1, CQ was assessed with an adapted version of the self-report CQ questionnaire (Van Dyne et al., 2012), as well as with two SJTs (for a more detailed description of the measures, see Schwarzenthal et al., 2019). In order to ensure that the results from Study 1 are not limited to the specific situations depicted in the two SJTs, one SJT used for Study 2 was the same as in Study 1 (SJTb), and one was different (SJTc)—the new SJT focused specifically on the arrival of a refugee adolescent in school (this situation was chosen to acknowledge the high number of refugee students, largely from Syria, who recently entered the German school system). Cronbach's α of the CQ subscales ranged between .86 and .90. For the SJTs, the same coding procedure as in Study 1 was followed to obtain

scores for consideration of cultural influences, suspension of judgment, and behavior. Two-way-random intraclass-correlations (Shrout & Fleiss, 1979) confirmed high intercoder reliability (from .76 to .93). Moreover, it was coded whether students considered multiple possible explanations when trying to make sense of the behavior in the SJTs, in order to avoid stereotyped explanations of cultural differences. The scale ranged from (1) *student only considered one possible explanation for the behavior of the parties in the situation* to (3) *student considered alternative explanations for the behavior of both parties in the situation*. Intercoder reliability was high (.82 and .83, respectively).

3.2.2 | Cultural diversity climate

In order to assess cultural diversity approaches, we used subscales from the revised version of the Cultural Diversity Climate Scale (Schachner et al., 2019). The response scales ranged from (1) *no, that's not right* to (5) *yes, that's right*. Validity and measurement equivalence between students of immigrant and non-immigrant background were confirmed by Schachner et al. (2019) with the same sample that was used in the present research. We included both individual perceptions of these approaches as well as classroom-level aggregates in our analyses.

Contact and cooperation
Contact and cooperation between students of diverse cultural backgrounds was assessed with three items (e.g., "Students in my class from different heritage cultures get along well with one another", α = .91).

Color-evasion
Norms fostering a color-evasive perspective were assessed with five items (e.g., "In class we learn that people of different backgrounds are all the same at heart"; α = .90).

Multiculturalism
For multiculturalism, we used the "heritage and intercultural learning" learning subscale, containing seven items (e.g., "During class we learn about the heritage cultures of fellow students"; α = .90).

3.2.3 | Control variables

As the students in Study 2 attended two different school tracks, and school track is related to intergroup outcomes (Schwarzenthal et al., 2018), we included school track (0 = *Integrated Secondary School*, 1 = *Gymnasium*) as an additional control variable in Study 2.

3.3 | Results

3.3.1 | Preliminary analyses

In total, 4% of data were missing on the variables that were used for analyses. Two students were excluded from analyses because cluster membership was unknown or because the student had missings on all outcome variables. As in Study 1, we employed multilevel SEM with FIML to test our hypotheses.

---

[5]Each federal state in Germany has its own school system and its own school types. The "Integrated Secondary School" in Berlin is not equivalent to the "Gesamtschule" in NRW. The "Integrated Secondary School" combines the former school types of Haupt-, Real- and Gesamtschule, while the "Gesamtschule" is an inclusive school form that exists parallel to the other school forms of Haupt-, Real-, and Gesamtschule.

3.3.2 | Measurement models at the individual level

The latent self-report CQ variable was measured as in Study 1. Due to the introduction of an additional coding category, the latent SJT variable was now measured by eight ratings that the students received based on their responses to the SJTs. Both measurement models fit well at the individual level (see Appendix 1). Tests of measurement equivalence between students of immigrant and non-immigrant background confirmed metric invariance for both models. The model fit decreased slightly for the scalar invariance models where intercepts were set to be equal across groups, implying that comparisons of means between the two groups should be treated with caution (see Appendix 2).

3.3.3 | ICCs and measurement models at the classroom level

ICC1s were all ≥.05 and thus higher than in Study 1. The reliabilities of the classroom aggregates (ICC2) were all higher than .51 (see Appendix 3). When the measurement models of the CQ and SJT variables were specified at the classroom level, the model fit was good (see Appendix 1). Thus, we conducted multilevel CFAs and tested whether factor loadings were invariant at the individual and classroom level in order to verify if the structure of the constructs was the same across levels (following Marsh et al., 2012). Since this did not lead to a substantial decrease in model fit (see Appendix 1), factor loadings were constrained to be invariant across levels in all following models. Standardized factor loadings were all significant, and, as is typically the case, were high at the classroom level (Marsh et al., 2012).

3.3.4 | Correlations

Since the measurement models fit well at both levels, we proceeded by calculating correlations at the individual and classroom level (see Tables 2 and 3). At the individual level, perceptions of all three types of classroom cultural diversity climate were positively related to both the self-report CQ subscales as well as to several of the SJT ratings. At the classroom level, the three types of classroom diversity climate were positively related to the self-report CQ subscales. Moreover, contact and cooperation was positively related to the SJT ratings.

3.3.5 | Multilevel SEM

Multilevel random-intercept models using the MLR estimator in Mplus 7.3 (Muthén & Muthén, 1998–2011) were run to test our hypotheses. All continuous predictors that were strictly at the individual level were grand-mean centered prior to the analyses. Correlations between predictors were allowed. As in Study 1, we ran separate analyses with the latent self-report CQ factor as outcome variable and with the latent SJT factor as outcome variable. After specifying the measurement models at the individual and classroom

level, we introduced predictors step by step. First, control variables (gender, immigrant background, grade average, and number of books) were introduced as predictors at the individual level. Females (compared to males), students of immigrant background (compared to students of non-immigrant background), students with higher grades, and students with more books in the household reported higher CQ. Female gender, grade average, and number of books in the household were related to higher scores in the SJTs, while immigrant background was related to lower scores (however, due to lack of scalar measurement invariance, this mean difference should be treated with caution). Model fit was good for the self-report CQ model, $\chi^2/df(N = 1,333) = 2.68$, $p < .001$, RMSEA = .04, CFI = .98, TLI = .96, $SRMR_{within}$ = .02, $SRMR_{betweenv}$ = .15, AIC = 22,829.54, as well as for the SJT model, $\chi^2/df(N = 1,333) = 1.66$, $p < .001$, RMSEA = .02, CFI = .96, TLI = 0.95, $SRMR_{within}$ = .03, $SRMR_{between}$ = .19, AIC = 29,579.95.

As a next step, we introduced the subscales of perceived classroom cultural diversity climate as predictors at the individual level (testing Hypotheses 1–3). Contact and cooperation as well as color-evasion positively predicted both self-reported CQ and performance in the SJTs, while multiculturalism only positively predicted self-reported CQ. Since we were also interested in differential relations between the predictors and the CQ subscales, we inspected modification indices. These indicated additional direct positive relations between contact and cooperation and CQ motivation and between multiculturalism and CQ metacognition, as well as a direct negative relation between multiculturalism and CQ motivation. Adding these relations improved model fit of the self-report CQ model (before: $\chi^2/df(N = 1,333) = 4.57$, $p < .001$, RMSEA = .05, CFI = .95, TLI = .90, $SRMR_{within}$ = .03, $SRMR_{between}$ = .17, AIC = 32,530.82, after: $\chi^2/df(N = 1,333) = 1.90$, $p = .002$, RMSEA = .03, CFI = .99, TLI = .98, $SRMR_{within}$ = .02, $SRMR_{between}$ = .17, AIC = 32,451.51).

We then tested one by one whether immigrant background moderated the associations between individually perceived classroom cultural diversity climate and the outcomes (testing Hypothesis 4). A perceived climate of contact and cooperation was more strongly related to CQ among students of non-immigrant background ($\beta = -.11$, $p = .04$). None of the other interaction effects were significant. Model fit was good for the model with the self-report CQ factor as outcome variable, $\chi^2/df(N = 1,333) = 4.92$, $p < .001$, RMSEA = .05, CFI = .95, TLI = .90, $SRMR_{within}$ = .06, $SRMR_{within}$ = .17, AIC = 34,046.33, as well as for the model with the SJT factor as outcome variable, $\chi^2/df(N = 1,333) = 1.60$, $p < .001$, RMSEA = .02, CFI = .97, TLI = .96, $SRMR_{within}$ = .03, $SRMR_{between}$ = .22, AIC = 39,416.77. Overall, the models explained 30% of variance in the latent self-report CQ factor, and 31% of the variance in the latent SJT factor at the individual level.

We proceeded by adding school track and proportion of students of immigrant background in a class as control variables at the classroom level. Students attending *Gymnasium* (as compared to students attending the Integrated Secondary School) reported marginally higher CQ and scored higher in the SJTs. The proportion of

TABLE 2 Study 2 individual level bivariate correlations and descriptives

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.
1. Gender (0 = male, 1 = female)	–	.24***	.06	.20***	.10*	.03	.24****	.14**	.11*	.17***	.19***	.21***	.22***	.28***	.18***	.21***	.19***	.22***
2. Grade average	.05	–	.33***	.26***	.20***	.12*	.23***	.23***	.18***	.19***	.13*	.19***	.24***	.19**	.04	.12*	.17***	.18***
3. No. of books in the household	-.03	.32***	–	.21***	.12*	.03	.30***	.28***	.19***	.22***	.08†	.02	.11*	.16***	.03	.08†	.14*	.20***
4. Contact and cooperation	.13**	.20***	.19***	–	.40***	.24***	.40***	.21***	.27***	.34***	.13*	.16**	.15**	.22***	.09*	.11*	.21***	.25***
5. Color-evasion	.01	.14***	.19***	.38***	–	.50***	.26***	.17**	.22***	.29***	.03	.05	.11*	.12*	.02	.05	.13***	.16**
6. Multiculturalism	.04	.10*	.12*	.28***	.52***	–	.16*	.23***	.32***	.29***	.01	.01	.08†	.04	.02	.03	.11**	.04
7. Self-reported CQ motivation	.15**	.16***	.21***	.33***	.32***	.20***	–	.42***	.34***	.49***	.15*	.12*	.20***	.33***	.12*	.12*	.20***	.43***
8. Self-reported CQ cognition	.10**	.16***	.16***	.17***	.23***	.24***	.41***	–	.37***	.41***	.07	.07†	.14***	.15**	.01	.11**	.08†	.24***
9. Self-reported CQ metacogn.	.08†	.09*	.07†	.16**	.26***	.35***	.34***	.42***	–	.65***	.06	.08†	.15***	.23***	.06	.13**	.15**	.22***
10. Self-reported CQ behavior	.13**	.14***	.09*	.17***	.26***	.27***	.46***	.47***	.65***	–	.15***	.06	.16***	.28***	.09*	.13**	.17***	.31***
11. SJTb suspending judgment	.10**	.10*	.06	.09†	.07	.02	.13**	.06	.09*	.11*	–	.31***	.27***	.25***	.02	.06	.17***	.28***
12. SJTb considering cult. infl.	.06	.05	.16***	.12*	.11*	-.01	.18***	.06	.02	.06	.29***	–	.21***	.25***	.07	.14***	.09*	.16***
13. SJTb considering altern. expl.	.06	.03	.11*	.06	.11*	-.01	.04	.04	.01	-.00	.23***	.31***	–	.25***	.11	.23***	.32***	.25***
14. SJTb behavior	.16**	.21***	.20***	.15***	.19***	.12*	.24***	.14**	.14**	.22***	.25***	.24***	.18***	–	.14***	.26***	.21***	.31***
15. SJTc suspending judgment	.02	-.03	.08†	.10*	.07†	.02	.05	.05	.03	-.00	.02	.08	-.02	.05	–	.34***	.10*	.11*
16. SJTc considering cult. infl.	.02	.10*	.11**	.09*	.02	-.01	.09†	.06	.00	.07	.10*	.19***	.06	.14**	.23***	–	.46***	.18***
17. SJTc considering altern. expl.	.07†	.10*	.11*	.10*	.10*	.09†	.06	.02	-.03	.04	.07	.16**	.23***	.14**	.11***	.30***	–	.21***

(Continues)

TABLE 2 (Continued)

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.
18. SJTc behavior	.14***	.16**	.18***	.23***	.24***	.20***	.29***	.22***	.21***	.26***	.11*	.19***	.10*	.31***	.09*	.16***	.13**	–
Students of non-immigrant background																		
M (SD)	0.48 (0.50)	3.26 (0.75)	3.81 (1.27)	4.12 (0.92)	3.51 (0.99)	2.66 (0.87)	3.89 (0.86)	2.90 (0.86)	2.79 (0.91)	3.18 (0.84)	2.42 (0.53)	2.26 (0.75)	1.38 (0.59)	3.22 (1.13)	1.88 (0.66)	2.08 (0.73)	1.90 (0.74)	3.31 (0.91)
Students of immigrant background																		
M (SD)	0.50 (.50)	3.18 (0.79)	3.16 (1.33)	4.08 (0.93)	3.59 (1.02)	2.86 (0.90)	4.00 (0.82)	3.12 (0.88)	2.92 (0.97)	3.28 (0.88)	2.36 (0.53)	2.04 (0.82)	1.27 (0.53)	3.04 (1.22)	1.76 (0.64)	1.94 (0.7)	1.68 (0.64)	3.22 (0.83)

Note: $N = 501$ students of non-immigrant background and $N = 689$ students of immigrant background, $N = 145$ students were excluded from this analysis because immigrant background was unknown. Correlations for students of non-immigrant background above the diagonal, correlations for students of immigrant background below the diagonal.

$^\dagger p < .10$, $^* p < .05$, $^{**} p < .01$, $^{***} p < .001$.

students of immigrant background in a class was unrelated to the outcome variables. Model fit was good for the self-report CQ model, $\chi^2/df(N = 1,333) = 4.86$, $p < .001$, RMSEA = .05, CFI = .95, TLI = .90, $SRMR_{within} = .06$, $SRMR_{between} = .17$, AIC = 34,117.93, as well as for the SJT model, $\chi^2/df(N = 1,333) = 1.63$, $p < .001$, RMSEA = .02, CFI = .97, TLI = .96, $SRMR_{within} = .03$, $SRMR_{between} = .22$, AIC = 39,476.67.

When all three subscales of classroom cultural diversity climate were introduced as predictors at the classroom level (testing Hypotheses 1–3 at the classroom level), none of them were related to the outcome variables. Since the subscales of classroom cultural diversity climate were highly correlated at this level, the results may have been affected by multicollinearity. Therefore, we ran separate analyses in which only one subscale of classroom cultural diversity climate was introduced as a predictor at a time. In these analyses, color-evasion and multiculturalism positively predicted self-reported CQ, and contact and cooperation positively predicted SJT. Color-evasion and multiculturalism marginally positively predicted SJT. In order to test whether there was an interaction between the aggregated diversity climate and immigrant background (testing Hypothesis 4 at the classroom level), we introduced random slopes between immigrant background and the outcome variables. However, the variance of the random slopes was close to zero and non-significant, which pre-empted a test of the interaction. Therefore, we did not include the random slopes in the final models. Standardized coefficients and model fit indices of the final models are depicted in Figures 3–6. Overall, the final models explained between 35 and 52% of variance in the self-report CQ factor, as well as between 55% and 98% of variance in the SJT factor at the classroom level.

Hypothesis 2 assumed that color-evasion would be unrelated to cognitive, metacognitive, and behavioral CQ. Since conventional significance tests technically cannot test for the absence of an effect, we used equivalence testing with one-sided t-tests to reject the presence of a smallest effect size of interest (SESOI) (Lakens, Scheel, & Isager, 2018). As these procedures are not yet available for multilevel regression, we only tested equivalence for the bivariate correlations between color-evasion with self-reported cognitive, metacognitive, and behavioral CQ as well as with the SJT score. As a SESOI we used the smallest effect size that we could detect with our sample of $N = 1,335$ students at an α-level of .05, with a power of .95, which was $r = .09$. Using the TOSTER module in the software jamovi (jamovi project, 2017), we found that equivalence could be rejected for the correlation between color-evasion with self-report cognitive CQ, $r = .20$, $p = 1.00$, self-report metacognitive CQ, $r = .25$, $p = 1.00$, self-report behavioral CQ, $r = .26$, $p = 1.00$, as well with the SJT score, $r = .17$, $p = .99$, confirming that the hypothesis that color-evasion was unrelated to CQ could not be supported.

3.4 | Discussion

In line with Hypothesis 1, students who perceived a more positive contact and cooperation climate in the classroom reported higher CQ, especially higher motivational CQ, and, unlike in Study 1, also showed marginally better performance in the SJTs. One reason for

**TABLE 3** Study 2 classroom level bivariate correlations

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1. School track (0 = Integrated Secondary School, 1 = Gymnasium)	–	$-.56^{***}$	$.50^{***}$	$.24^{†}$	$-.27^{*}$	$.39^{***}$	$.34^{*}$	.10	.19	.54	.66	$.61^{†}$	$.61^{†}$	$.51^{***}$	$.66^{*}$	$.64^{***}$	.53
2. Proportion of students of imm. background		–	$-.16$	.07	$.50^{***}$	.06	.22	.34	.13	$-.41$	$-.64$	$-.33^{*}$	$-.18$	$.39^{†}$	$-.51$	$.58^{**}$	$-.15$
3. Contact and cooperation			–	$.74^{**}$	$.28^{†}$	$.84^{***}$	$.61^{*}$	$.48^{*}$	$.61^{*}$	.77	$.75^{†}$	$.69^{*}$	$.84^{†}$	$.39^{*}$	.83	$.85^{***}$	$.91^{*}$
4. Color-evasion				–	$.81^{**}$	$.72^{**}$	$.71^{***}$	.76	$.76^{*}$	.59	.35	$.51^{*}$	$.81^{†}$	.10	.48	.50	.80
5. Multiculturalism					–	$.41^{**}$	$.55^{*}$	$.74^{**}$	$.62^{**}$	.06	.26	.07	.31	$-.09$	.01	.02	.34
6. Self-reported CQ motivation						–	.92	$.84^{***}$	$.91^{***}$	.72	.48	.65	.80	.50	.50	.57	.88
7. Self-reported CQ cognition							–	.95	.98	.59	.21	.49	.70	.28	.24	.25	.74
8. Self-reported CQ metacogn.								–	$.97^{***}$	.40	$-.02$	.27	.55	.22	.12	.12	.59
9. Self-reported CQ behavior									–	.55	.15	.40	.69	.21	.21	.11	.62
10. SJTb suspending judgment										–	.87	.91	.85	.17	.78	.23	.87
11. SJTb considering cult. infl.											–	.81	.71	.46	.81	$.89^{***}$	.77
12. SJTbconsidering altern. expl.												–	.78	.40	$.77^{*}$	$.77^{*}$	.86
13. SJTb behavior													–	.16	.64	.67	.94
14. SJTc suspending judgment														–	.56	.65	.36
15. SJTc considering cult. infl.															–	$.94^{†}$	.74
16. SJTc considering altern.expl.																–	$.79^{†}$
17. SJTc behavior																	–

Note: N = 66 classrooms.

$^{†}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

FIGURE 3 Study 2: Model with CQ as outcome variable, standardized coefficients
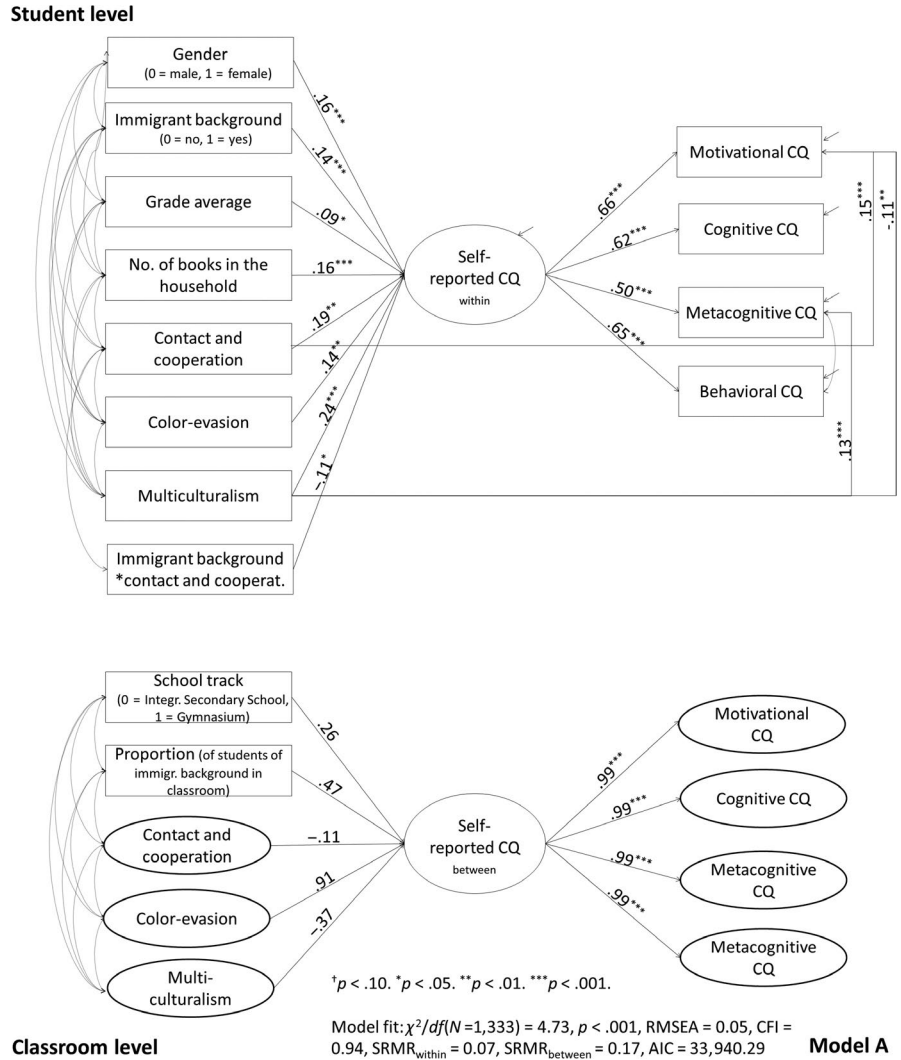
this finding may be that color-evasion was included as an additional predictor in Study 2, thus, the effects of contact and color-evasion can be distinguished more clearly.

Not supporting Hypothesis 2, if students were taught in school that cultural differences are not important (reflecting a color-evasive approach), they reported higher CQ and performed better in the SJTs. However, our results also support the assumption that active discussion of cultural variations at school is important to make use of the learning potential inherent in cultural diversity. Students' perceptions of multiculturalism in the classroom were the strongest predictor of self-reported CQ, and were especially important for students' awareness of cultural variations, i.e., for metacognitive CQ. However, multiculturalism only showed weak relations with students' self-reported motivational CQ, and was unrelated to students' performance in the SJTs (partly supporting Hypothesis 3).

Contrary to Study 1, the outcome variables in Study 2 could be reliably modeled at the classroom level, and we found relations between all three types of classroom cultural diversity climate and the outcome variables at this level. As in Study 1, we found few indications that associations differed between students of immigrant and non-immigrant background (not supporting Hypothesis 4). Only the

individually perceived contact and cooperation climate was more positively related to CQ among students of non-immigrant background.

4 | GENERAL DISCUSSION

We investigated relations between three classroom cultural diversity approaches and students' CQ based on data collected in classrooms in two culturally diverse regions in Germany. Our results show that each of the classroom diversity approaches is positively related to adolescents' intercultural competence.

### 4.1 | Contact and cooperation and students' CQ

In line with Hypothesis 1, a classroom climate supporting contact and cooperation between students of diverse cultural affiliations was related to all aspects of students' CQ, with slightly different associations in the two datasets. This finding is consistent with previous research showing that positive contact norms in a classroom are positively related to students' intergroup attitudes (Molina & Wittig, 2006; Schwarzenthal et al., 2018) and that intercultural contact is
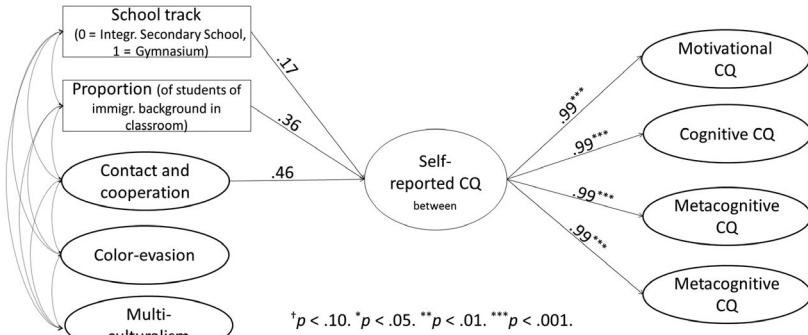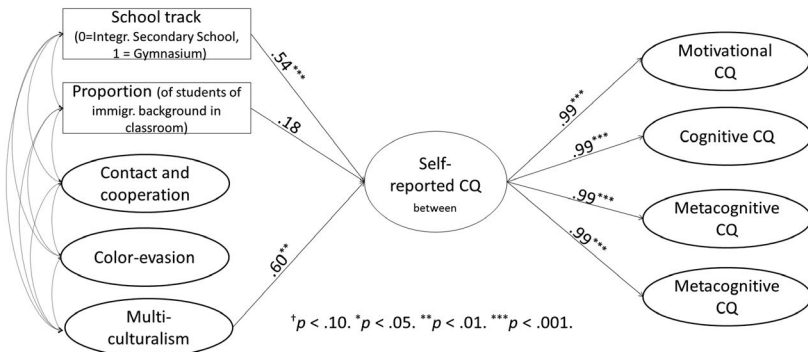
FIGURE 4 Study 2: Alternative classroom-level models with CQ as outcome variable, only one aspect of classroom cultural diversity climate introduced as predictor at a time, standardized coefficients
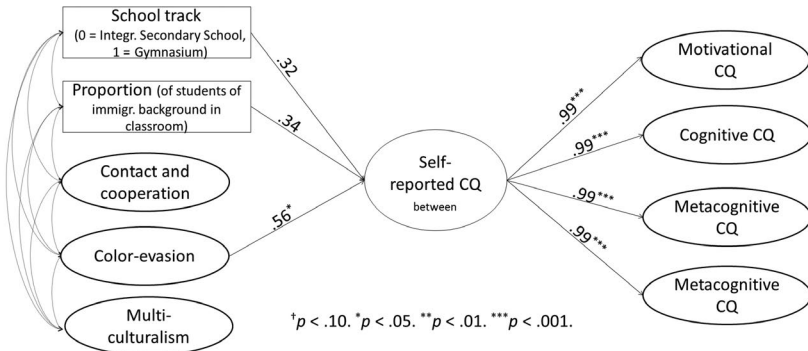
related to CQ (Schwarzenthal et al., 2017). Thus, if schools promote positive contact between students of diverse cultural backgrounds, students do not only display more positive attitudes, but also higher intercultural knowledge, awareness, and more effective behavioral strategies in intercultural situations.

Students in classrooms that support contact and cooperation are more likely to engage in prolonged and personal intercultural experiences such as intercultural friendships (Schachner et al., 2015). Students with intercultural friendships experience cultural variation when they are invited home to a friend's house, listen to them speak their heritage language with their parents, and see them celebrate different traditions. These experiences can stimulate experiential learning processes if the students actively reflect on them (Kolb, 1984). Future research is needed to investigate to what degree students perceive cultural variation in intercultural friendships, to what
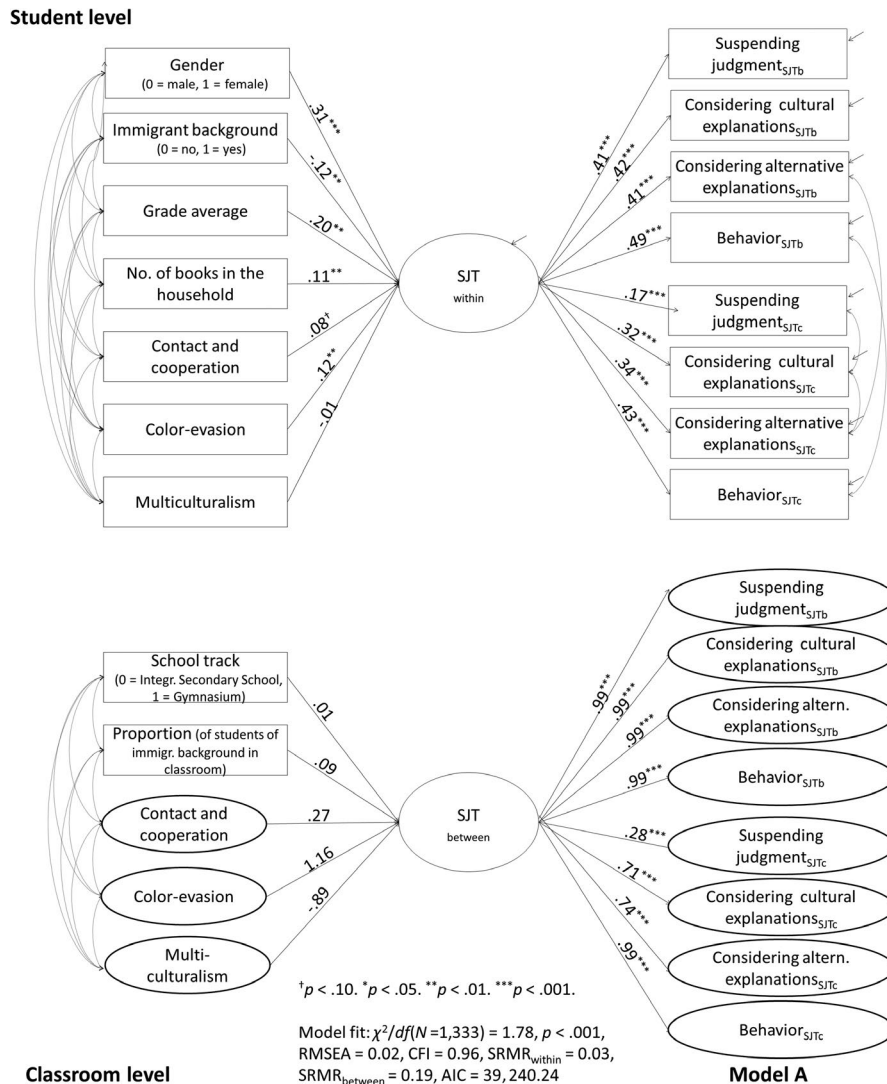
degree they actively reflect on these experiences, and whether this is related to their intercultural learning.

4.2 | Color-evasion and students' CQ

Students in Study 2 also completed a measure assessing their perception of a classroom climate of color-evasion (as in emphasizing a common humanity). We had expected that color-evasion would be unrelated to cognitive, metacognitive, and behavioral CQ because it suppresses perceptions of cultural variations (Hypothesis 2). However, color-evasion was related to higher self-reported CQ as well as to better performance in the SJTs.

There are several possible explanations for these findings: The generic terms "colorblindness/color-evasion" have a rather negative connotation, but they are used to refer to a range of different constructs

FIGURE 5 Study 2: Model with SJT as outcome variable, standardized coefficients


Student level

Classroom level

Model A

$\dagger p < .10.$ $^* p < .05.$ $^{**} p < .01.$ $^{***} p < .001.$

Model fit: $\chi^2/df(N = 1{,}333) = 1.78$, $p < .001$, RMSEA = 0.02, CFI = 0.96, SRMR$_{within}$ = 0.03, SRMR$_{between}$ = 0.19, AIC = 39,240.24

whose effects may differ. In our study, we only included a measure of "color-evasion" (i.e., emphasizing a common humanity), but not of "power-evasion" (i.e., ignoring the existence of racism and discrimination) (Neville, Awad, Brooks, Flores, & Bluemel, 2013). While we expect that ignorance of racism and discrimination is clearly detrimental to students' intercultural learning, fostering a common group identity and emphasizing similarities may also promote positive intergroup attitudes and increase students' self-efficacy to manage intercultural interactions successfully. Moreover, it may stimulate students to see others as humans with multiple possible motives for behavior, and not only as members of a cultural group. An emphasis on similarities between people of diverse cultural affiliations can go along with ignoring differences, but does not necessarily have to (Dovidio et al., 2010). Indeed, perceptions of color-evasion in the classroom were positively related to perceptions of multiculturalism ($r = .50$ among students of non-immigrant background and $r = .52$ among students of immigrant background) in our study, suggesting that both approaches can coexist and may be compatible.

Moreover, the sociohistorical context in Germany may partly explain why color-evasion in the classroom was related to positive outcomes in this study. Dealing with memories of what is widely perceived

as a shameful historical period is central to Germany's identity and Holocaust education is part of every school curriculum (Eckmann & Österberg, 2017; MacGregor, 2014). Holocaust education aims to foster reflection on structures promoting hatred against outgroups, and, as German philosopher Theodor Adorno famously stated, to not let Auschwitz happen again (Adorno, 1977, cited in Eckmann & Österberg, 2017). Students in German schools learn early to not use the term "race" due to associations with genocide. As a consequence, many schools promote the idea that all humans are equal, and focus less on discussing diversity and differences (Civitillo et al., 2016). Thus, ideologies that promote treating everyone the same and ignoring differences may have a more positive connotation in Germany than in the U.S.

4.3 | Multiculturalism and students' CQ

Supporting Hypothesis 3, a classroom climate of multiculturalism was positively related to students' CQ. This is in line with previous theory and research suggesting that learning about diversity is related to positive intergroup attitudes (Schwarzenthal et al., 2018; Verkuyten & Thijs, 2013), as well as to intercultural knowledge, understanding, and
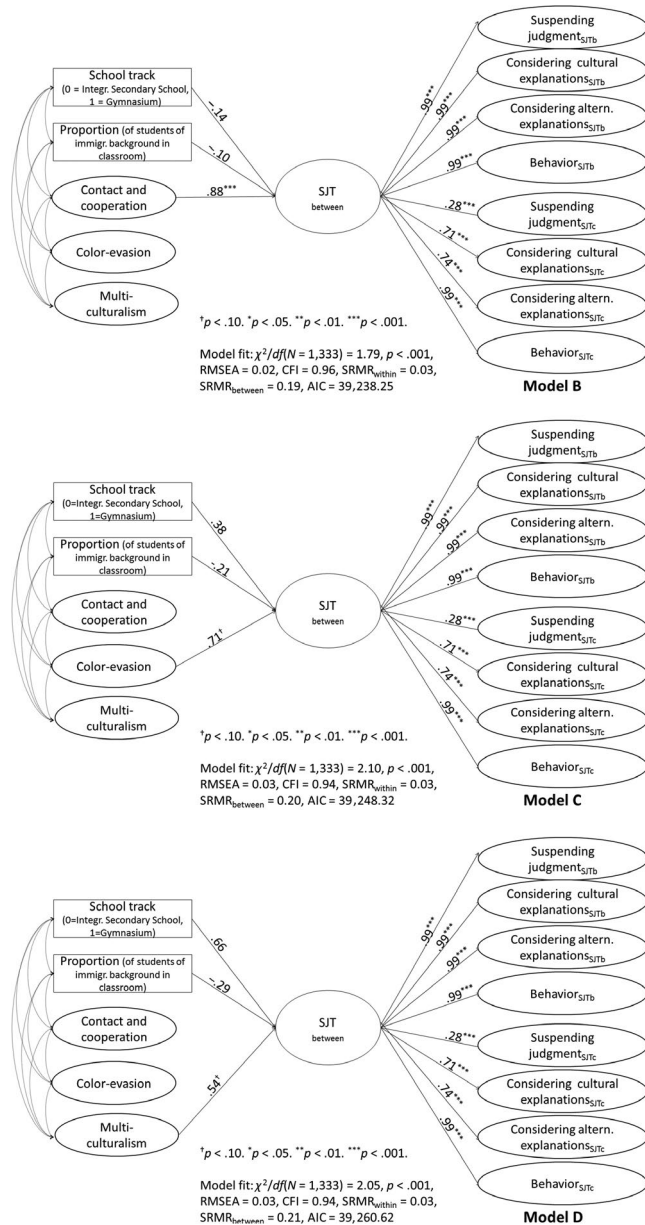
FIGURE 6 Study 2: Alternative classroom-level models with SJT as outcome variable, only one aspect of classroom cultural diversity climate introduced as predictor at a time, standardized coefficients

interculturally competent behavior (e.g., Gurin et al., 2002). However, in Study 2 the effects were smaller than expected since multiculturalism was neither related to students' motivational CQ, nor to their performance in the SJTs.

Multiculturalism as operationalized in this study mainly captures what teachers teach in class, and not whether students encounter diverse perspectives in more informal interactions. Previous research conducted on college campuses showed that diversity coursework and workshops are less strongly related to intercultural skills than interpersonal interactions (Bowman, 2010). Thus, multicultural education alone may not be sufficient to foster all aspects of CQ if it is not combined with personal experiences. Moreover, multicultural education can also promote stereotypes and reduce liking, especially if

it is implemented in a rather superficial way, which is often the case in German schools (Civitillo et al., 2016; Rosenthal & Levy, 2010). A strong and superficial focus on differences may promote "othering" and evoke the impression that intercultural interactions are difficult and challenging, and thus, should rather be avoided.

4.4 | Individually perceived climate and aggregated climate

Our results suggest that both students' individual perceptions of classroom cultural diversity climate, as well as the classroom-aggregated climate predict students' CQ. Thus, even though there is variability in students' perceptions of diversity climate within a classroom, to some extent students also agree in their perceptions, and this shared perception is related to their intercultural learning outcomes. The subscales of classroom cultural diversity climate were highly correlated at the classroom level, hinting at a broad underlying dimension of classroom cultural diversity climate at this level that mainly reflects how much attention is devoted in a classroom to (constructively) deal with cultural diversity.

4.5 | Differential effects among students of immigrant and non-immigrant background

Not supporting Hypothesis 4, relations between classroom cultural diversity climate and students' CQ were largely similar among students of immigrant and non-immigrant background. The few differences that we found were in the expected direction, that is, relations between perceived cultural diversity climate and CQ were stronger among students of non-immigrant background. This is in line with meta-analytic research on the contact hypothesis that found stronger effects of contact on majority members' intergroup attitudes (Tropp & Pettigrew, 2005). The students in our samples all attended very diverse schools. In some classrooms, students of non-immigrant background were clearly in the minority and may also encounter cultural diversity on a daily basis, for example, at school, in their peer group or in their neighborhood. Thus, the two groups may not differ as clearly with regard to their social status and their previous levels of intercultural experience as we had assumed.

4.6 | Strengths, limitations, and suggestions for future research

Our study is one of the first examining intercultural competence among adolescents attending culturally diverse schools. The findings point to the positive potential of culturally diverse schools for the development of these skills, and may contribute to shifting the debate in media and research from a deficit-oriented to a more resource-oriented perspective on these schools.

In order to explore the conditions under which diversity in schools may be a resource, we directly compared three approaches, and could show that each of them may have its unique strengths. Future research may build on our results and investigate how nuances of these approaches (e.g., different nuances of color-evasion

or power-evasion, superficial vs. non-superficial implementation of multiculturalism) relate to students' intercultural learning.

In contrast to the majority of research on intercultural competence, we not only assessed intercultural competence with a self-report measure, but also with SJTs asking students for their interpretation of and reaction to intercultural situations. Despite this strength, our measures still have shortcomings: The format of the study, a large-scale questionnaire study, did not allow the use of more extensive assessments of intercultural competence. In order to capture whether students have a more nuanced and non-essentialist idea of what culture is and how it may or may not impact people's behavior, alternative measurement methods need to be developed, for example interviews, role-plays, or more extensive student essays.

Both of our samples were collected in ethnically diverse regions in Germany, where there is a high degree of intergroup contact, and intergroup attitudes are typically rather positive (Wagner, Van Dick, Peterson, & Christ, 2006). Students growing up in less diverse regions may profit even more from a classroom climate that encourages discussions on cultural variability. Even though their schools may be less diverse with regard to ethnicity, variability with regard to other dimensions, such as SES, gender, or generation, could be discussed. Future research is needed to understand the manifestation and impact of classroom cultural diversity climate in contexts with low ethnic diversity.

An additional limitation of our study is that both datasets were cross-sectional; therefore, it is impossible to draw conclusions about causality. Empirical support for the causal link from contact to attitudes is stronger and more consistent than vice versa (Pettigrew & Tropp, 2011). Therefore, we do expect that the effect of classroom cultural diversity climate on intercultural competence is stronger than vice versa. Future research is needed to test this assumption empirically.

5 | CONCLUSION

We showed that contact and cooperation, color-evasion, and multiculturalism in the classroom are uniquely related to aspects of students' intercultural competence. In line with other research, our results suggest that none of the approaches is sufficient on their own (Rosenthal & Levy, 2010). An approach of contact and cooperation may foster the personal and prolonged intercultural experiences that are central for intercultural competence. Color-evasion, as in emphasizing a common humanity, may encourage students to look beyond group categories and to recognize that humans are still individuals, and may act out of a variety of motives, be they cultural, individual, or interpersonal. A multicultural approach may encourage active reflections on diversity that are important for intercultural knowledge and awareness. Together, the three approaches may contribute to reaping the benefits of cultural diversity in schools.

CONFLICT OF INTEREST

The manuscript has not been previously published, nor is it submitted for publication elsewhere. It is our original work. The authors declare that they do not have any interests that might be interpreted as influencing the research.

ETHICS

In conducting this study, APA ethical standards have been followed. In particular, the content of the study was approved by the Berlin Senate Committee for Education, Youth, and Science. We obtained permission from school principals to conduct the study in their schools, and obtained informed consent from participants and their parents. Participation was voluntary and confidential.

TRANSPARENCY STATEMENT

Since we do not have the permission from the Berlin Senate Comittee for Education, Youth, and Science to share data with third parties, data cannot be shared publicly.

ORCID

Miriam Schwarzenthal (ID) https://orcid.org/0000-0002-6278-6227

REFERENCES

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Perseus Books.

American Field Service (AFS) (2012). *The impact of living abroad study*. Retrieved from https://www.afs.de/publikationen-zum-thema-ikl.html.

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural intelligence: Its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance. *Management and Organization Review*, *3*, 335–371. https://doi.org/10.1111/j.1740-8784.2007.00082.x

Apfelbaum, E. P., Pauker, K., Sommers, S. R., & Ambady, N. (2010). In blind pursuit of racial equality? *Psychological Science*, *21*, 1587–1592. https://doi.org/10.1177/0956797610384741

Banks, J. A. (2015). *Cultural diversity and education* (6th ed.). New York, NY: Routledge.

Barrett, M. (2018). How schools can promote the intercultural competence of young people. *European Psychologist*, *23*, 93–104. https://doi.org/10.1027/1016-9040/a000308

Barrett, M., Byram, M., Lázár, I., Mompoint-Gaillard, P., & Philippou, S. (2013). *Developing intercultural competence through education*. Strasbourg, France: Council of Europe Publishing.

Beaton, A. M., Dovidio, J. F., & Léger, N. (2008). All in this together? Group representations and policy support. *Journal of Experimental Social Psychology*, *44*, 808–817. https://doi.org/10.1016/j.jesp.2007.07.002

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity (revised). In R. M. Paige (Ed.), *Education for the intercultural experience* (pp. 21–71). Yarmouth, ME: Intercultural Press.

Blau, P. M. (1977). A macrosociological theory of social structure. *American Journal of Sociology*, *83*, 26–54. https://doi.org/10.1086/226505

Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G., & Valtin, R. (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich [First results from the IGLU study: Student achievement at the end of fourth grade in international comparison]*. Münster, Germany: Waxmann.

Bowman, N. A. (2010). College diversity experiences and cognitive development: A meta-analysis. *Review of Educational Research*, 80, 4–33. https://doi.org/10.3102/0034654309352495

Boyce, W., Torsheim, T., Currie, C., & Zambon, A. (2006). The family affluence scale as a measure of national wealth: Validation of an adolescent self-report measure. *Social Indicators Research*, 78, 473–487. https://doi.org/10.1007/s11205-005-1607-6

Busse, V., Riedesel, L., & Krause, U.-M. (2017). Anregung von Reflexionsprozessen zur Förderung interkultureller Kompetenz [Stimulation of reflective processes in order to promote intercultural competence]. *Zeitschrift für Pädagogik*, 63, 361–385.

Byrd, C. M. (2014). The associations of intergroup interactions and school racial socialization with academic motivation. *The Journal of Educational Research*, 108, 10–21. https://doi.org/10.1080/00220671.2013.831803

Byrd, C. M. (2017). The complexity of school racial climate: Reliability and validity of a new measure for secondary students. *British Journal of Educational Psychology*, 87, 700–721. https://doi.org/10.1111/bjep.12179

Chiu, C. Y., & Hong, Y.-Y. (2013). *Social psychology of culture*. New York, NY: Routledge. https://doi.org/10.4324/9781315782997

Christ, O., Schmid, K., Lolliot, S., Swart, H., Stolle, D., Tausch, N., & Hewstone, M. (2014). Contextual effect of positive intergroup contact on outgroup prejudice. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 3996–4000. https://doi.org/10.1073/pnas.1320901111

Civitillo, S., Schachner, M., Juang, L., van de Vijver, F. J. R., Handrick, A., & Noack, P. (2016). Towards a better understanding of cultural diversity approaches at school: A multi-informant and mixed-methods study. *Learning, Culture and Social Interaction*, 12, 1–14. https://doi.org/10.1016/j.lcsi.2016.09.002

Correll, J., Park, B., & Smith, J. A. (2008). Colorblind and multicultural prejudice reduction strategies in high-conflict situations. *Group Processes & Intergroup Relations*, 11, 471–491. https://doi.org/10.1177/1368430208095401

Crisp, R. J., & Turner, R. N. (2011). Cognitive adaptation to the experience of social and cultural diversity. *Psychological Bulletin*, 137, 242–266. https://doi.org/10.1037/a0021840

Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: The development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Social Science and Medicine*, 66, 1429–1436. https://doi.org/10.1016/j.socscimed.2007.11.024

Deardorff, D. K. (2004). The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States.), (Doctoral dissertation). Retrieved from the North Carolina State University Library. (http://www.lib.ncsu.edu/resolver/1840.16/5733) North Carolina State University, Raleigh, NC.

Deardorff, D. K. (2011). Assessing intercultural competence. *New Directions for Institutional Research*, 149, 65–79. https://doi.org/10.1002/ir.381

Dejaeghere, Y., Hooghe, M., & Claes, E. (2012). Do ethnically diverse schools reduce ethnocentrism? A two-year panel study among majority group late adolescents in Belgian schools. *International Journal of Intercultural Relations*, 36, 108–117. https://doi.org/10.1016/j.ijintrel.2011.02.010

Denson, N., & Chang, M. J. (2008). Racial diversity matters: The impact of diversity-related student engagement and institutional context. *American Educational Research Journal*, 46, 322–353. https://doi.org/10.3102/0002831208323278

Denson, N., & Zhang, S. (2010). The impact of student experiences with diversity on developing graduate attributes. *Studies in Higher Education*, 35, 529–543. https://doi.org/10.1080/03075070903222658

Dovidio, J. F., Gaertner, S. L., & Saguy, T. (2010). Another view of "we": Majority and minority group perspectives on a common ingroup identity. *European Review of Social Psychology*, 18, 296–330. https://doi.org/10.1080/10463280701726132

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, 16, 149–167. https://doi.org/10.1016/j.leaqua.2004.09.009

Eckmann, M., & Österberg, O. (2017). Chapter 1: Research in German. In M. Eckmann, D. Stevick & J. Ambrosewicz-Jacobs (Eds.), *Research in teaching and learning about the Holocaust: A dialogue beyond borders* (pp. 37–54). Berlin, Germany: Metropol.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY, USA: Guilford Publications.

Erdmann, L. (2017). *Wanka für begrenzten Anteil von Migranten in Klassen [Wanka proposes limited number of migrants in classrooms]*. Spiegel Online. Retrieved from: http://www.spiegel.de/lebenundlernen/schule/johanna-wanka-anteil-von-migranten-in-schulklassen-begrenzen-a-1144387.html.

Faas, D. (2008). From foreigner pedagogy to intercultural education: An analysis of the German responses to diversity and its impact on schools and students. *European Educational Research 7*, 108–123. https:doi.org/10.2304/eerj.2008.7.1.108.

Frankenberg, E., Kupper, K., Wagner, R., & Bongard, S. (2013). Immigrant youth in Germany. *European Psychologist*, 18, 158–168. https://doi.org/10.1027/1016-9040/a000154

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The Common Ingroup Identity Model*. Philadelphia, PA: Psychology Press.

Geiser, C. (2011). *Datenanalyse mit Mplus*. Wiesbaden, Germany: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-93192-0

Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review*, 72, 330–366. https://doi.org/10.17763/haer.72.3.01151786u134n051

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge. https://doi.org/10.4324/9780203852279

jamovi project. (2017). Jamovi (Version 0.8) [Computer software]. Retrieved from https://www.jamovi.org

Jon, J. E. (2013). Realizing internationalization at home in Korean higher education: Promoting domestic students' interaction with international students and intercultural competence. *Journal of Studies in International Education*, 17, 455–470. https://doi.org/10.1177/1028315312468329

Klafehn, J., Li, C., & Chiu, C. Y. (2013). To know or not to know, is that the question? Exploring the role and assessment of metacognition in cross-cultural contexts *Journal of Cross-Cultural Psychology*, 44, 963–991. https://doi.org/10.1177/0022022113492893

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, NJ: Prentice-Hall Inc.

Kristen, C., & Granato, N. (2007). The educational attainment of the second generation in Germany: Social origins and ethnic inequality. *Ethnicities*, 7, 343–366. https://doi.org/10.1177/1468796807080233

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1, 259–269. https://doi.org/10.1177/2515245918770963

Lam, A. C., Ruzek, E. A., Schenke, K., Conley, A. M., & Karabenick, S. A. (2015). Student perceptions of classroom achievement goal structure: Is it appropriate to aggregate? *Journal of Educational Psychology*, 107, 1102–1115. https://doi.org/10.1037/edu0000028

Levy, S. R., & Killen, M. (Eds.) (2010). *Intergroup attitudes and relations in childhood through adulthood*. Oxford, UK: Oxford University Press.

Loes, C., Pascarella, E., & Umbach, P. (2012). Effects of diversity experiences on critical thinking skills: Who benefits? *The Journal of Higher Education*, 83, 1–25. https://doi.org/10.1353/jhe.2012.0001

MacGregor, N. (2014). *Germany: Memories of a nation*. London, UK: Allen Lane.

Maddux, W. W., Bivolaru, E., Hafenbrack, A. C., Tadmor, C. T., & Galinsky, A. D. (2013). Expanding opportunities by opening your mind: Multicultural engagement predicts job market success through longitudinal increase in cognitive complexity. *Social Psychological and Personality Science*, 5, 608–615. https://doi.org/10.1177/19485 50613515005

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, 106–124. https://doi.org/10.1080/00461520.2012.670488

Matsumoto, D., & Hwang, H. C. (2013). Assessing cross-cultural competence: A review of available tests. *Journal of Cross-Cultural Psychology*, 44, 849–873. https://doi.org/10.1177/0022022113492891

Miller, A. D., & Murdock, T. B. (2007). Modeling latent true scores to determine the utility of aggregate student perceptions as classroom indicators in HLM: The case of classroom goal structures. *Contemporary Educational Psychology*, 32, 83–104. https://doi.org/10.1016/j.cedpsych.2006.10.006

Molina, L. E., & Wittig, M. A. (2006). Relative importance of contact conditions in explaining prejudice reduction in a classroom context: Separate and equal? *Journal of Social Issues*, 62, 489–509. https://doi.org/10.1111/j.1540-4560.2006.00470.x

Motti-Stefanidi, F., Berry, J. W., Chryssochoou, X., Sam, D. L., & Phinney, J. S. (2012). Positive immigrant youth adaptation in context: Developmental, acculturation, and social-psychological perspectives. In A. S. Masten, K. Liebkind, & D. J. Hernandez (Eds.), *Realizing the potential of immigrant youth* (pp. 117–158). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9781139094696

Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Neville, H. A., Awad, G. H., Brooks, J. E., Flores, M. P., & Bluemel, J. (2013). Color-blind racial ideology: Theory, training, and measurement implications in psychology. *American Psychologist*, 68, 455–466. https://doi.org/10.1037/a0033282

Neville, H. A., Lilly, R. L., Duran, G., Lee, R. M., & Browne, L. (2000). Construction and initial validation of the Color-Blind Racial Attitudes Scale (CoBRAS). *Journal of Counseling Psychology*, 47, 59–70. https://doi.org/10.1037/0022-0167.47.1.59

Ott, D. L., & Michailova, S. (2018). Cultural intelligence: A review and new research avenues. *International Journal of Management Reviews*, 20, 99–119. https://doi.org/10.1111/ijmr.12118

Park, B., & Judd, C. M. (2005). Rethinking the link between categorization and prejudice within the social cognition perspective. *Personality and Social Psychology Review*, 9, 108–130. https://doi.org/10.1207/s15327957pspr0902_2

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90, 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, 38, 922–934. https://doi.org/10.1002/ejsp.504

Pettigrew, T. F., & Tropp, L. R. (2011). *When groups meet: The dynamics of intergroup contact*. New York, NY: Psychology Press.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. New York, NY: Viking Press.

Quintana, S. M. (1998). Children's developmental understanding of ethnicity and race. *Applied & Preventive Psychology*, 7, 27–45. https://doi.org/10.1016/s0962-1849(98)80020-6

Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development*, 82, 1715–1737. https://doi.org/10.1111/j.146 7-8624, 2011.01668.x

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, 14, 287–290. https://doi.org/10.1111/1467-9280.03437

Richter, M., & Leppin, A. (2007). Trends in socio-economic differences in tobacco smoking among German schoolchildren, 1994–2002. *European Journal of Public Health*, 17, 565–571. https://doi.org/10.1093/eurpub/ckm01

Rockstuhl, T., Ang, S., Ng, K. Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology*, 100, 464–480. https://doi.org/10.1037/a0038098

Rosenthal, L., & Levy, S. R. (2010). The colorblind, multicultural, and polycultural ideological approaches to improving intergroup attitudes and relations. *Social Issues and Policy Review*, 4, 215–246. https://doi.org/10.1111/j.1751-2409.2010.01022.x

Rothman, S. (2003). Does enrollment diversity improve university education? *International Journal of Public Opinion Research*, 15, 8–26. https://doi.org/10.1093/ijpor/15.1.8

Rutland, A., & Killen, M. (2015). A developmental science approach to reducing prejudice and social exclusion: Intergroup processes, social-cognitive development, and moral reasoning. *Social Issues and Policy Review*, 9, 121–154. https://doi.org/10.1111/sipr.12012

Sam, D. L., & Oppedal, B. (2003). Acculturation as a developmental pathway. *Online Readings in Psychology and Culture*, 8, https://doi.org/10.9707/2307-0919.1072

Sasaki, S. J., & Vorauer, J. D. (2013). Ignoring versus exploring differences between groups: Effects of salient color-blindness and multiculturalism on intergroup attitudes and behavior. *Social and Personality Psychology Compass*, 7, 246–259. https://doi.org/10.1111/spc3.12021

Schachner, M. K. (2017). From equality and inclusion to cultural pluralism: Evolution and effects of cultural diversity perspectives in schools. *European Journal of Developmental Psychology*, 16, 1–17. https://doi.org/10.1080/17405629.2017.1326378

Schachner, M. K., Brenick, A., Noack, P., van de Vijver, F. J. R., & Heizmann, B. (2015). Structural and normative conditions for interethnic friendships in multiethnic classrooms. *International Journal of Intercultural Relations*, 47, 1–12. https://doi.org/10.1016/j.ijintrel.2015.02.003

Schachner, M. K., Noack, P., Van de Vijver, F. J., & Eckstein, K. (2016). Cultural diversity climate and psychological adjustment at school-equality and inclusion versus cultural pluralism. *Child Development*, 87, 1175–1191. https://doi.org/10.1111/cdev.12536

Schachner, M. K., Schwarzenthal, M., Moffitt, U., Civitillo, S., & Juang, L. P. (2019). Capturing a nuanced picture of classroom cultural diversity climate: Multigroup and multilevel analyses among secondary school students in Germany. Manuscript submitted for publication.

Schenke, K., Ruzek, E., Lam, A. C., Karabenick, S. A., & Eccles, J. S. (2017). Heterogeneity of student perceptions of the classroom climate: A latent profile approach. *Learning Environments Research*, 20, 289–306. https://doi.org/10.1007/s10984-017-9235-z

Schmid, K., Al Ramiah, A., & Hewstone, M. (2014). Neighborhood ethnic diversity and trust: The role of intergroup contact and perceived threat. *Psychological Science*, 25, 665–674. https://doi.org/10.1177/0956797613508956

Schofield, J. W. (2010). The colorblind perspective in school: Causes and consequences. In J. A. Banks & C. A. M. Banks (Eds.), *Multicultural education: Issues and perspectives* (pp. 271–288). Hoboken, NJ: John Wiley & Sons.

Schwarzenthal, M., Juang, L.P., Schachner, M. K., & van de, Vijver, A. J. R. (2019). A multimodal measure of cultural intelligence for adolescents growing up in culturally diverse societies. *International Journal of Intercultural Relations*, 72, 109–121. https://doi.org/10.1016/j.ijintrel.2019.07.007

Schwarzenthal, M., Juang, L. P., Schachner, M. K., van de Vijver, A. J. R., & Handrick, A. (2017). From tolerance to understanding: Exploring the development of intercultural competence in multiethnic contexts from early to late adolescence. *Journal of Community & Applied Social Psychology*, *27*, 388–399. https://doi.org/10.1002/casp.2317

Schwarzenthal, M., Schachner, M. K., Van de Vijver, A. J. R., & Juang, L. P. (2018). Equal but different: Effects of equality/inclusion and cultural pluralism on intergroup outcomes in multiethnic classrooms. *Cultural Diversity and Ethnic Minority Psychology*, *24*, 260–271. https://doi.org/10.1037/cdp0000173

Sharma, N., & Hussain, D. (2017). Current status and future directions for cultural intelligence. *Journal of Intercultural Communication Research*, *46*, 91–110. https://doi.org/10.1080/17475759.2016.126444

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Sieck, W. R., Smith, J. L., & Rasmussen, L. J. (2013). Metacognitive strategies for making sense of cross-cultural encounters. *Journal of Cross-Cultural Psychology*, *44*, 1007–1023. https://doi.org/10.1177/0022022113492890

Spencer-Oatey, H. (2008). *Culturally speaking. Culture, communication and politeness theory* (2nd ed.). London, UK: Continuum.

Straffon, D. A. (2003). Assessing the intercultural sensitivity of high school students attending an international school. *International Journal of Intercultural Relations*, *27*, 487–501. https://doi.org/10.1016/S0147-1767(03)00035-X.

Stanat, P. (2006). Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft [School performance of youth of immigrant background: The role of classroom composition]. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 189–219). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90082-7

Statistisches Bundesamt. (2018). Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund—Ergebnisse des Mikrozensus 2017-. [Population and employment. *Population of immigrant background - Results of the census 2017*]. Retrieved from https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/Migrationshintergrund2010220177004.pdf?__blob=publicationFile

Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood*, *3*, 375–387. https://doi.org/10.1177/2167696815587648

Tadmor, C. T., & Tetlock, P. E. (2006). Biculturalism: A model of the effects of second-culture exposure on acculturation and integrative complexity. *Journal of Cross-Cultural Psychology*, *37*, 173–190. https://doi.org/10.1177/0022022105284495

Taylor, M. C. (1998). How white attitudes vary with the racial composition of local populations: Numbers count. *American Sociological Review*, *63*, 512–535. https://doi.org/10.2307/2657265

Terenzini, P. T., Cabrera, A. F., Colbeck, C. L., Bjorklund, S. A., & Parente, J. M. (2001). Racial and ethnic diversity in the classroom: Does it promote student learning? *The Journal of Higher Education*, *72*, 509–531. https://doi.org/10.2307/2672879

Thijs, J., & Verkuyten, M. (2014). School ethnic diversity and students' interethnic relations. *British Journal of Educational Psychology*, *84*(Pt 1), 1–21. https://doi.org/10.1111/bjep.12032

Todd, A. R., & Galinsky, A. D. (2012). The reciprocal link between multiculturalism and perspective-taking: How ideological and self-regulatory approaches to managing diversity reinforce each other. *Journal of Experimental Social Psychology*, *48*, 1394–1398. https://doi.org/10.1016/j.jesp.2012.07.007

Tropp, L. R., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, *16*, 951–957. https://doi.org/10.1111/j.1467-9280.2005.01643.x

Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology*, *50*, 881–888. https://doi.org/10.1037/a0034325

van de Vijver, F. J. R. (2002). Inductive reasoning in Zambia, Turkey, and the Netherlands. *Intelligence*, *30*, 313–351. https://doi.org/10.1016/s0160-2896(02)00084-3

Van Dyne, L., Ang, S., Ng, K. Y., Rockstuhl, T., Tan, M. L., & Koh, C. (2012). Sub-dimensions of the four factor model of cultural intelligence: Expanding the conceptualization and measurement of cultural intelligence. *Social and Personality Psychology Compass*, *6*, 295–313. https://doi.org/10.1111/j.1751-9004.2012.00429.x

Verkuyten, M., & Thijs, J. (2013). Multicultural education and interethnic attitudes. *European Psychologist*, *18*, 179–190. https://doi.org/10.1027/1016-9040/a000152

Vorauer, J. D., Gagnon, A., & Sasaki, S. J. (2009). Salient intergroup ideology and intergroup interaction. *Psychological Science*, *20*, 838–845. https://doi.org/10.1111/j.1467-9280.2009.02369.x

Wagner, U., Van Dick, R., Peterson, D., & Christ, O. (2006). Ethnic prejudice in East and West Germany: The explanatory power of intergroup contact. *Group Processes & Intergroup Relations*, *6*, 22–36. https://doi.org/10.1177/1368430203006001010

Wang, M.-T., & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, *28*, 315–352. https://doi.org/10.1007/s10648-015-9319-1

APPENDIX 1

Fit of measurement models at the individual and classroom level

	χ^2/df	p	RMSEA	CFI	TLI	SRMR$_{within}$	SRMR$_{between}$
Study 1							
Self-reported CQ[a]							
Individual level	2.63	.10	.05	1.00	.99	.01	–
Same measurement model at individual and classroom level	5.96	<.001	.09	.97	.94	.02	.72
Loadings set equal across individual and classroom level	3.65	<.001	.07	.97	.97	.02	.39
SJT[b]							
Individual level	2.00	.05	.04	.98	.96	.03	–
Same measurement model at individual and classroom level	2.29	<.001	.05	.93	.91	.03	.24
Loadings set equal across individual and classroom level	3.29	<.001	.06	.85	.83	.05	.24
Study 2							
Self-reported CQ [a]							
Individual level	5.66	.02	.06	1.00	.98	.01	–
Same measurement model at individual and classroom level	4.71	<.001	.05	.98	.96	.01	.10
Loadings set equal across individual and classroom level	5.03	<.001	.06	.97	.96	.02	.10
SJT[b]							
Individual level	1.73	.04	.02	.99	.98	.02	–
Same measurement model at individual and classroom level	1.64	.01	.02	.98	.97	.02	.14
Loadings set equal across individual and classroom level	1.92	<.001	.03	.96	.95	.03	.15

Note: Study 1: *N* = 631 students in 29 classrooms, Study 2: *N* = 1,333 students in 66 classrooms.[a]Correlated error between CQ metacognition and CQ behavior at the individual level.[b]Since each SJT captures performance in one and the same situation, the ratings that the students receive for their responses to one and the same SJT are likely to be related. Moreover, it is likely that students are good at developing behavioral strategies in a diverse range of situations, which is why their scores on the intercultural behavior dimension are positively related across situations. Therefore, if modification indices indicated correlated errors between scores from the same SJT or between the same rating dimension, these were allowed.

APPENDIX 2 Equivalence of latent factor models across students of immigrant and non-immigrant background

Scale	Invariance	Fit statistics				
		χ^2/df	RMSEA	CFI	TLI	SRMR
Study 1						
Self-reported CQ[a]	Configural	1.65	.05	1.00	.99	.01
	Metric	1.49	.04	1.00	.99	.09
	Scalar	3.52***	.09	.95	.94	.15
SJT[b]	Configural	1.31	.03	.99	.97	.03
	Metric	1.14	.04	.98	.96	.04
	Scalar	1.42†	.04	.97	.96	.04
Study 2						
Self-reported CQ[a]	Configural	3.60*	.07	1.00	.97	.01
	Metric	1.81†	.04	1.00	.99	.03
	Scalar	2.55**	.05	.98	.98	.04
SJT[b]	Configural	2.43***	.05	.94	.90	.04
	Metric	2.42***	.05	.93	.90	.05
	Scalar	2.97***	.06	.88	.86	.06

Note: [a]Correlated error between CQ metacognition and CQ behavior at the individual level.[b]Since each SJT captures performance in one and the same situation, the ratings that the students receive for their responses to one and the same SJT are likely to be related. Moreover, it is likely that students are good at developing behavioral strategies in a diverse range of situations, which is why their scores on the intercultural behavior dimension are positively related across situations. Therefore, if modification indices indicated correlated errors between scores from the same SJT or between the same rating dimension, these were allowed.†*p* < .10, **p* < .05, ***p* < .01, ****p* < .001.

APPENDIX 3 ICCs

	ICC1	ICC2
Study 1		
1. Contact and cooperation climate	.05	.54
2. Multiculturalism climate	.10	.71
3. Self-reported CQ motivation	.02	.31
4. Self-reported CQ cognition	.03	.40
5. Self-reported CQ metacognition	.05	.54
6. Self-reported CQ behavior	.03	.40
7. SJTa suspending judgment	.11	.73
8. SJTa considering cultural influences	.12	.75
9. SJTa behavior	.04	.48
10. SJTb suspending judgment	.04	.48
11. SJTb considering cultural influences	.05	.54
12. SJTb behavior	.03	.40
Study 2		
1. Contact and cooperation climate	.16	.79
2. Color-evasion climate	.09	.66
3. Multiculturalism climate	.16	.79
4. Self-reported CQ motivation	.11	.71
5. Self-reported CQ cognition	.09	.66
6. Self-reported CQ metacognition	.06	.56
7. Self-reported CQ behavior	.05	.51
8. SJTb suspending judgment	.05	.51
9. SJTb considering cultural influences	.05	.51
10. SJTb alternative explanations	.07	.60
11. SJTb behavior	.07	.60
12. SJTc suspending judgment	.08	.63
13. SJTc considering cultural influences	.07	.60
14. SJTc alternative explanations	.08	.63
15. SJTc behavior	.09	.66

Note: Study 1: N = 631 students in 29 classrooms, Study 2: N = 1,333 students in 66 classrooms.