Longitudinal evaluation of friendship project: A multicultural – antiracist program for elementary school children

Elli Spyropoulou¹ · Angeliki Sourlantzi² · Aspasia Karakosta² · Maria Kotsampasoglou¹ · Theodoros Giovazolias¹

Published online: 9 June 2020

© The Author(s) 2020

Abstract

This study presents the results of the evaluation of Friendship Project-Greek version (FP-GR), a school based multicultural and anti-racist program aiming to develop intercultural skills and positive intergroup relationships, as well as familiarize and sensitize children with refugees' life circumstances and difficulties. The aim of the study was to adapt FP in the Greek context and evaluate its longitudinal effectiveness in Greek students. The sample of the present research consisted of 314 children aged 9–12 years from six elementary schools in Athens and Crete, Greece. Students were randomly assigned to condition (intervention or control group) and the effectiveness of the program was examined through three measurements (before, at the end of the intervention and 3 months later). Data analyses (Repeated measures ANCOVA's, Confirmatory Factor Analyses) were conducted using SPSS 25 and Mplusv.8.0. Results showed that FP-GR influenced participants' general attitudes towards refugees, affected their xenophobic and intolerant attitudes but did not seem to have any significant impact on participants' altruism. These results provide initial evidence that the FP-GR is an effective school-based prejudice-reduction program that promotes inclusive education for refugee-background elementary school children. Suggestions for the improvement of specific elements of FP-GR are also provided.

Keywords Refugee children · Inclusive education · Multicultural and anti-racist school program · Attitudes · Anti-racism policy

Introduction

Beyond doubt, modern societies are characterized by constant and broad changes on economical social, demographic and educational level (Cavounidis 2013). Up to the end of 2018, 70.8 million people had been forcibly displaced from their country of origin because of conflict, war and persecution. This number included nearly 25.9 million refugees with over half of them being children [United Nations High Commissioner for Refugees (UNHCR) 2019]. Thus, a global concern has been raised not only in terms of ensuring the safe settlement of refugee background people (UNHCR 2019) but also in terms of promoting their peaceful and harmonious social inclusion to the host countries (Sacramento 2015). Given that education is a key factor for social cohesion, the integration of refugees into the educational settings becomes

one societal challenge high in the international agenda (Beißert et al. 2019). Accumulating evidence suggest that education is among the most effective ways for refugee youth to become socially accepted and integrated in a new culture (Thomas et al. 2016). Unfortunately, youth with refugee background face often attitudinal barriers related to stereotypes, prejudice and xenophobia that make integration and quality education more difficult to attain (Pugh et al. 2012). Relevant literature suggests that refugee youth are usually stigmatized and become targets of bias-based bullying in the form of racist name calling, teasing, social isolation and segregation (Earnest et al. 2015). The social exclusion they usually face represents a problem with profound negative effects not only to their school adjustment but also to their overall psychosocial well-being (Szaflarski and Bauldry 2019). Social integration is considered a reciprocal process which can be successfully achieved when the host community adopts an open, positive and inclusive orientation towards cultural diversity (Berry 1997; Plexousakis et al. 2019). It is perceived to be the most effective strategy for immigrants' well-being that requires a readiness on the part of the receiving communities to welcome ethnic minorities without prejudice (Sengupta and Blessinger 2018). Moreover, inclusive school environments have been shown to promote students' school performance,

✉ Theodoros Giovazolias
giovazot@uoc.gr

¹ Department of Psychology, University of Crete, Gallos Campus, Rethymno, 74100 Crete, Greece

² Department of Early Childhood Education, National and Kapodistrian University of Athens, Athens, Greece

cultivate trust and acceptance of cultural differences, reduce prejudicial attitudes and enhance civic engagement in adulthood (Kourkoutas and Giovazolias 2015; Mickelson and Nkomo 2012). Particularly, it has been suggested that when children are actively involved in educational programs that establish acceptance and encourage the exchange of views, their general altruistic attitudes can be promoted (Büssing et al. 2013). Research has shown that altruism, which refers to voluntary actions that benefit another without personal interest or external awards (Batson et al. 2015), raises support for migration (Hansen and Legge 2016) and that children who care more about the well-being of others tend to increase helping behaviors towards migrant populations (Batson et al. 2009).

Given that schools can highly contribute to tackle the rise of racism, discrimination and xenophobia (De Paola and Brunello 2016), the implementation of educational programs that address ethnic diversity and foster a pluralistic school culture for host and refugees students has become an issue of critical importance for the research community (Barnett 2018).

School-Based Programs towards Ethnic Diversity

Culturally sensitive and inclusive interventions aim to alter students' racial attitudes and decrease prejudicial and xenophobic manifestations (Beelmann and Heinemann 2014). Multicultural and anti-racist school-based interventions support that prejudice stems from ignorance or misleading social information and can be reduced by providing knowledge that changes the way children think and feel about an out-group (White and Gleitzman 2006). Thus, by providing anti-bias information through socialization influences (e.g. story books, games) children acquire the skills needed to break down negative generalizations, overcome roadblocks to empathy (Aboud et al. 2012) and cultivate multicultural awareness (Killen and Smetana 2015). Further, social skills interventions supporting the active construction of attitudes focus on teaching children different ways of processing information underlying group dynamics (e.g. through role playing) (Rutland and Killen 2015).

The effectiveness of interventions aiming at changing children's racial attitudes is well documented in the literature (Brenick et al. 2019). Okoye-Johnson (2011) in a meta-analysis of 30 curricular and reinforcement multicultural intervention studies conducted in prekindergarten through twelfth graders, found that exposure to multicultural education led to a reduction in students' racial attitudes. Similarly, Aboud et al. (2012) in a systematic review of 32 studies on anti-prejudice interventions (14 contact and 18 media/instruction) delivered to young children found that somewhat fewer than half had positive outcomes in ethnic prejudice and discrimination reduction. Yet, most of these studies are conducted

within the context of more traditional majority/minority relations (e.g. White majority/Black minority, host majority/immigrants minority children). Surprisingly, there is a dearth of well-structured evaluated interventions targeting attitudes of majority group members towards other out-groups such as refugees (Ülger et al. 2018). To our knowledge, only three studies have evaluated interventions focusing on changing children's attitudes and behaviors towards refugees. Cameron et al. (2006) assessed the effectiveness of extended contact -through friendship storytelling - interventions in changing children's (5-to 11-years old) attitudes towards refugees and intended behaviors towards hypothetical refugee background children. Results supported the effectiveness of extended contact intervention only in improving attitudes towards refugees. Turner and Brown (2008) evaluated the longitudinal effectiveness of "The Friendship Project", a multicultural-antiracist school-based intervention in the UK designed to improve children's attitudes towards refugees. The program was found to be effective in improving short-term but not long-term attitudes. Finally, Glen et al. (2020) in a study with children aged 8 to 11 years, found that inducing empathy through brief narrative interventions promoted prosocial behavior towards refugees.

The Friendship Project-Greece (FP-GR)

The FP-GR is a multifaceted anti-prejudice program that focuses on promoting positive intergroup relations among Greek host primary students and refugee background children. The core elements of FP-GR are derived from "The Friendship Project" a multicultural-antiracist program, designed by a small voluntary organization [Kent Refugee Action Network (KRAN)] working to address the needs of refugees and asylum seekers in the Kent area, UK. FP-GR is considered to be an enriched version of the original "The Friendship Project" as it endorses a multi-strategic approach (Greco et al. 2010) that draws upon social learning and cognitive development theories regarding the development and change of children's prejudicial attitudes.

More precisely, FP-GR entails four weekly structured lessons concentrating on a threefold objective: a) knowledge: an understanding of concepts such as "refugee", "racism" and "discrimination", b) values of participants: to encourage children's respect towards refugees, and c) skills: for example, identification of prejudicial and egocentric attitudes in themselves and in society and empathy (Turner and Brown 2008). Based on the Empathy-Attitude-Action model which supports the role of empathy in promoting prosocial behaviors (Batson et al. 2002) FP-GR (for a detailed description of the original "The Friendship Project" see Turner and Brown 2008) was supplemented with extra empathy activities such as video (e.g. "Ivive and the Pillow"), collaboration tasks (e.g. group discussion, role-playing) and reflective writing (e.g. the FP-GR

diary, a personal booklet designed and distributed to children by the interventionists). More specifically, according to previous research, videos that induce empathy by portraying others in need decrease prejudice and enhance children's engagement in prosocial behaviors (Vezzali et al. 2019; Williams et al. 2014). Additionally, collaboration activities that teach students how to work effectively with others, despite of their differences in attitudes seem to promote empathetic skills (Hollingsworth et al. 2003). Finally, reflective writing has been suggested as an effective instructional tool for fostering empathy (Chen and Forbes 2014).

The content of each FP-GR's lesson is outlined below:

Lessons 1 & 2: Spot the refugee. The intervention begins with a child-friendly video animated by UNICEF entitled "Ivine and the Pillow" to motivate children and involve them in the program. This short video (2.35') depicts in an emotionally engaging way the true story of a 14-year-old refugee child after their perilous escape from Syria. Through brainstorming, children elaborate on what makes someone a refugee. Interventionists discuss the definition of the concept "refugee" as "someone who is unable or unwilling to return to their country of origin owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion" (United Nations General Assembly 1951). Next, LEGO® minifigures are distributed to the class and children are asked to identify the similarities that the figures share and suggest an identity for each figure based on its characteristics. The interventionist, using a UNHCR poster with LEGO® minifigures, encourages children to guess which one of the figures is a refugee, explain how they came up with that choice and think if their decision is in accordance with the given definition of "refugee". The above-mentioned structured learning tasks aim for a better understanding of the roots of racial biases and discriminative thinking as well as the harmful effects of distorted prejudicial beliefs on people's emotional and social well-being.

Lessons 3 & 4: How does it feel? Children working in small groups (4–5 persons) discuss how they would feel in several scenarios (e.g. "as a child of their age who has accidentally been separated from the family while holidaying in a foreign land") and portray their thoughts and feelings. A UNHCR's poster entitled "How does it feel" is used as a motivator, where one LEGO® figure is alone, away from the group. Interventionists use a series of questions as the basis for discussion (e.g. What similarities and differences are there between the situation of the lonely LEGO® figure and the hypothetical scenarios we considered?). Children are next encouraged to engage in role-playing, an effective cooperative learning technique (Stevens 2015) that enhances perspective taking (Mead 1934). Through all these collaboration activities participants are expected to explore different ways of thinking about diversity, develop feelings of empathy and cultivate the skills needed for handling possible social interactions with outgroups.

At the end of each lesson children are prompt to reflect upon and describe in their FP-GR diary their inner thoughts and feelings about what they had experienced and learned in the class.

Study Justification, Aim and Research Questions

As mentioned earlier in this literature review, social integration is considered to be a crucial factor for refugee's youth well-being and the establishment of long-term harmonious intergroup relations (Buchanan et al. 2018). Greece has recently received a substantial influx of refugees; according to the International Rescue Committee (IRC 2019), Greece currently hosts approximately 50 thousand refugees, with the most having stayed long enough to expect final resettlement. Official statistics show that the number of under-aged refugees studying in Greek public schools has already reached a record of 12.5 thousand with nearly 66% of them being asylum seekers or beneficiaries of international protection (UNHCR/Greece 2019).

However, the Greek society is still characterized by increased levels of racial discrimination. In particular, Greek youth seem to be prone to nationalist, xenophobic and authoritarian attitudes and intolerance towards ethnic minority groups (Koronaoui et al. 2015). Thus, given the increased hostility of Greek community towards refugees and the possible danger of socially segregating and marginalizing this group (Hangartner et al. 2019) it is more than vital to implement evaluated programs that focus on changing host students' anti-refugee sentiment. Accumulating evidence also suggest that prejudice reduction interventions are more effective when implemented in primary school students and adolescents (Aboud et al. 2012) as intergroup attitudes, favoring ingroup and excluding outgroup members seem to develop in middle-childhood (Nesdale 2008).

In this study, we evaluated the longitudinal effectiveness of FP-GR in a sample of Greek students, aged 9 to 11 years. This program was designed to: a) reduce xenophobia and prejudice towards refugees, b) cultivate tolerant attitudes towards ethnic diversity and, c) promote general altruism by fostering empathy. More specifically, it has been supported that xenophobia, defined as the fear or hatred against foreigners, decreases when the foreign and strange become more familiar (Suleman et al. 2018). Evidence also suggest that the reduction of xenophobic attitudes are associated with lower levels of prejudice against refugees (Obeid et al. 2019). Furthermore, tolerance towards refugees, described as the emergence of positive feelings and the acceptance of equality between the ingroup and outgroup members (Van Zalk et al. 2013) is an attitude that can be cultivated through activities that promote empathy (Hollingsworth et al. 2003). Similarly, altruism is a prosocial behavior that can be promoted through the development of empathetic understanding (Batson et al. 2015). The following hypotheses have been formulated based on the above objectives:

H1: The participation in FP-GR is expected to have a short- and a long-term positive impact on students' attitudes towards refugees.

H2: The participation in FP-GR is expected to reduce students' intolerant attitudes towards refugees both in a short and in a long-term basis.

H3: The participation in FP-GR is expected to decrease students' xenophobia towards refugees both in a short and in a long-term basis.

H4: The participation in FP-GR is expected to have no impact on Greek student's attitudes towards the in group.

H5: The participation in FP-GR is expected to have a short- and a long-term positive impact on Greek students' general altruism.

Methods

Overall Study Design and Data Collection

FP-GR involved four weekly sessions (approx. 90 min each), which were conducted in children's classrooms by four trained psychologists/counselors. The longitudinal effectiveness of the program was evaluated with three assessments or waves separated by 1 month and 3 months, respectively. More precisely, the baseline (Time 1) assessment was conducted in February of 2019 (before the intervention), the second (Time 2) in March of 2019 (after the completion of the intervention), while the third (Time 3) in June of 2019 (3 months after the completion of the intervention). An a priori G*power analysis ($ES = .25$, $\alpha = .05$, $df = 3$) indicated a sample size of 175 in order to achieve results with 80% power. In order to ensure a sufficient sample, our eligibility criteria included schools with over 300 students enrolled in grades 4, 5 and 6 that did not operate reception structures for refugee education. An initial call was sent to 20 schools and the intervention took place to six schools that responded positively. Students with parental consent from each grade willing to participate in the study were given a unique personal code and were randomly allocated to the intervention group (IG) and to the control group (CG) respectively, using Google's random number generator. For the data collection, for each assessment, a paper-based self-report anonymous questionnaire was administered to the total sample. The interventionist's academic profile in conjunction with a manual containing a detailed description of each lesson plan and a protocol recording intervention process at the end of each session ensured the fidelity of the program. Voluntary and anonymity of the participant's responses were clearly explained and ensured. Permission by teachers and school principals as well parental consent for children's participation in the study were also required. No compensation was given for the participation in the study. The research

obtained approval by the Ethics Committee of University of Crete and the Greek Ministry of Education.

Participants

The sample at Time 1 consisted of 367 Greek regular elementary students ($M_{age} = 10.79$; $SD = .87$, 51.2% girls) enrolled in grades 4, 5 and 6 in six primary public schools in Athens and Crete. The schools were selected through convenience sampling, as explained above, and participants were randomly assigned to IG ($n = 200$, 50.5% girls) and CG ($n = 167$, 51.2% girls). An attrition rate of 9.5% at Time 2 and Time 3 respectively, as well as deleted outliers ($n = 18$) reduced the analyzed sample to 314 participants ($M_{age} = 10.79$; $SD = .87$; 51.9% girls; 55.1% intervention group) (see Fig. 1). The two groups (intervention vs control) did not differ in terms of sex or school grade they attended ($\chi^2 = .406$, $p = .524$ and $\chi^2 = .401$, $p = .818$, respectively). Concerning the final sample's distribution in grade, 107 (34.1%) were 4th graders, 99 (31.5%) were 5th graders and 108 (34.4%) were 6th graders. Additionally, 273 (86.9%) were Greeks, and 36 (11.7%) were of "Other" racial heritage (e.g. Albanian, Italian).

Instruments

Attitudes towards Refugees Attitudes towards refugees (e.g. I like refugees) were measured by a six-item self-report instrument constructed by Turner and Brown (2008). The Greek version of the "Attitudes towards Refugees" scale was translated from the original, back-translated, and adjusted for cultural adaptation by the research team. Items were rated on a five-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*) with higher scores indicating more positive attitudes towards refugees. No missing values were reported in a sample of 177 Greek primary students ($M_{age} = 10.6$, $SD = .83$, 55.4% girls). The psychometric properties of this and the following scales were tested using Mplus v.8.0 software. One-factor confirmatory factor analysis (CFA) model using maximum likelihood estimator (MLR; Yuan and Bentler 2000) proved an excellent fit to the data, $\chi^2 = 5.14$, $df = 8$, $p > .05$. Metric ($\Delta\chi^2_{\text{metric vs configural}} = 6.62$, $df = 5$, $p > .05$) and scalar invariance ($\Delta\chi^2_{\text{2scalar vs metric}} = 5.25$, $df = 5$, $p > .05$) across gender were also confirmed. Internal consistency of the instrument was, $\omega = .85$. Test-retest reliability – 4 weeks interval – was also supported ($r = .77$, $p < .001$). In the present study the one-factor CFA model for the "Attitudes towards Refugees" scale provided an exact fit to the data ($\chi^2 = 8.50$, $df = 8$, $p > .05$). Metric ($\Delta\chi^2_{\text{2metric vs configural}} = 13.64$, $df = 10$, $p > .05$) and scalar invariance ($\Delta\chi^2_{\text{2scalar vs metric}} = 1.81$, $df = 9$, $p > .05$) across time were also confirmed. Omega values were as follows: (Time 1, $\omega = .83$; Time 2, $\omega = .85$; Time 3, $\omega = .86$).

Fig. 1 Flowchart of study participants

Attitudes towards Greek People Attitudes towards Greek people (e.g. I like Greek people) were measured in order to provide evidence for the discriminant validity of the intervention. Following the paradigm of Turner and Brown (2008) we used the same items as before, but we replaced the word “Refugees” with “Greek people”. The Greek version of the “Attitudes towards Greek people” scale was translated from the original, back-translated, and adjusted for cultural adaptation by the research team. No missing values were reported in a sample of 177 Greek primary students ($M_{age} = 10.6$, $SD = .83$, 55.4% girls). One-factor CFA model using MLR estimator (Yuan and Bentler 2000) proved an excellent fit to the data, $\chi^2 = 8.50$, $df = 8$, $p > .05$. Metric ($\Delta\chi^2_{metric \text{ vs configural}} = 3.88$, $df = 5$, $p > .05$) and scalar invariance ($\Delta\chi^2_{scalar \text{ vs metric}} = 1.79$, $df = 5$, $p > .05$) across gender were also confirmed. Internal consistency of the instrument was $\omega = .76$. Test-retest reliability - 4 weeks interval - was supported ($r = .78$, $p < .001$). In the present study the one-factor CFA model for the “Attitudes towards Greek people” scale provided a reasonable fit to the data, $\chi^2 = 23.56$, $df = 8$, $p < .05$, RMSEA = .08, CFI = .93, SRMR = .04. Metric ($\Delta\chi^2_{metric \text{ vs configural}} = 13.76$, $df = 10$, $p > .05$) and partial scalar invariance ($\Delta\chi^2_{partial \text{ scalar vs metric}} = 11.06$, $df = 8$, $p > .05$) across time were also confirmed. Omega values were as follows: (Time 1, $\omega = .70$; Time 2, $\omega = .80$; Time 3, $\omega = .81$).

Tolerance and Xenophobia Tolerant (e.g. Immigrants are good for the Greek economy) and xenophobic attitudes (e.g. Immigrants increase criminality) towards refugees were measured with eight items from the Tolerance and Xenophobic Questionnaire (TRQ, see Van Zalk et al. 2013). The Greek version of the TRQ scale was translated from the original, back-translated, and adjusted for cultural adaptation

by the research team. Each item was rated on a four-point Likert scale (1 = *don't agree at all* to 4 = *agree completely*) with higher scores indicating more frequent use of the corresponding attitude. For the purpose of our study we changed the words “Swedish”, “Sweden” and “Immigrants” with “Greek”, “Greece” and “Refugees” respectively. We also recoded Tolerance items so that lower scores reflected increased level of intolerance. The Little’s Missing Completely at Random (MCAR) test in a sample of 177 Greek primary students ($M_{age} = 10.6$, $SD = .83$, 55.4% girls) was not significant, $\chi^2 = 5.51$, $df = 6$, $p = .48$, indicating that missing values were distributed completely at random. A factor loading cutoff of .32 (Tabachnick and Fidell 2001) suggested the deletion of item 8 ($\beta = .11$, $p > .05$) of the Xenophobia subscale. The two-factor CFA model proposed by previous studies (Van Zalk et al. 2013) provided using MLR estimator (Yuan and Bentler 2000) a reasonable fit to the data, $\chi^2 = 28.01$, $df = 12$, $p < .05$, RMSEA = .09, CFI = .91, SRMR = .05. Discrimination of the two constructs was accomplished as the factor correlation coefficient ($r = -.74$, $p < .001$) did not exceed the cutoff value of .85 (Brown 2006). Metric ($\Delta\chi^2_{metric \text{ vs configural}} = 5.15$, $df = 5$, $p > .05$) and scalar invariance ($\Delta\chi^2_{scalar \text{ vs metric}} = 2.28$, $df = 5$, $p > .05$) of the questionnaire across gender were also confirmed. Internal consistency for the four-item Tolerance subscale was $\omega = .67$ and for the three-item Xenophobia subscale, $\omega = .59$, respectively. Test-retest reliability was also supported (Tolerance, $r = .72$, $p < .001$; Xenophobia, $r = .75$, $p < .001$). In the present study the two-factor CFA provided a satisfactory fit to the data, $\chi^2 = 33.94$, $df = 12$, $p < .05$, RMSEA = .08, CFI = .91, SRMR = .06. Metric ($\Delta\chi^2_{metric \text{ vs configural}} = 11.51$, $df = 10$, $p > .05$) and partial scalar invariance ($\Delta\chi^2_{partial \text{ scalar vs metric}} = 12.97$, $df = 8$, $p > .05$) across time

were also confirmed. Omega values for the Xenophobia subscale were (Time 1, $\omega = .57$; Time 2, $\omega = .60$; Time 3, $\omega = .72$) and for the Tolerance subscale were (Time 1, $\omega = .65$; Time 2, $\omega = .67$; Time 3, $\omega = .68$), respectively.

Altruism The Generative Altruism Scale (GALS, Büssing et al. 2013) is a self-report instrument designed to measure adolescents' and young adults' altruism, (e.g. When I see needy persons, I ask them how I can help). The Greek version of the GALS scale was translated from the original, back-translated, and adjusted for cultural adaptation by the research team. All items were scored on a four-point Likert scale measuring the intensity of the respective attitude or behavior (0 = *Never* to 3 = *Very Often*). There are two versions of the GALS scale, the 11 and the 9-item version respectively. The MCAR test in a sample of 177 Greek primary students ($M_{age} = 10.6$, $SD = .83$, 55.4% girls) was not significant, $\chi^2 = 17.23$, $df = 10$, $p = .07$, indicating that missing values were distributed completely at random. One-factor CFA model using MLR estimator (Yuan and Bentler 2000) for the 11-item scale proved an acceptable model fit with characteristics of $\chi^2 = 62.385$, $df = 43$, $p < .001$, RMSEA = .05, CFI = .97, SRMR = .04. Metric ($\Delta\chi^2_{metric \text{ vs } configural} = 3.45$, $df = 10$, $p > .05$) and scalar invariance ($\Delta\chi^2_{scalar \text{ vs } metric} = 6.82$, $df = 10$, $p > .05$) across gender were also confirmed. Internal consistency for GALS was, $\omega = .87$. Test-retest reliability was also supported ($r = .74$, $p < .01$). In the present study the one-factor CFA model for the GALS provided a satisfactory fit to the data, $\chi^2 = 68.55$, $df = 43$, $p < .001$, RMSEA = .04, CFI = .95, SRMR = .04. Partial metric ($\Delta\chi^2_{partial \text{ metric vs } configural} = 17.71$, $df = 19$, $p > .05$) and partial scalar invariance ($\Delta\chi^2_{partial \text{ scalar vs } partial \text{ metric}} = 24.64$, $df = 16$, $p > .05$) across time were supported. Omega values were as follows: (Time 1, $\omega = .80$; Time 2, $\omega = .86$; Time 3, $\omega = .85$).

Data Analyses

Attrition analyses with Independent-Samples T-tests were carried out to test the differences between participants and no participants in the baseline variables. A Kolmogorov-Smirnov test was also conducted to assess the normality of the study variables. Furthermore, descriptive statistics were examined and zero-order correlations were carried out among all the research variables in the three waves. Multivariate analyses were conducted to examine the differences by demographics (i.e., gender, grade and parental level of educational attainment) in the study variables in the three assessment points. Repeated measures ANCOVAs, controlling for the influence of the four interventionists, were carried out in order to evaluate the longitudinal effectiveness of the FP. Interventionists were treated as continuous variables by dummy coding (the first interventionist was specified to be the reference group). Moreover, univariate and multivariate

outliers were assessed by boxplot and mahalanobis distance and were treated accordingly. Homogeneity of variances (Levene's test) and homogeneity of covariances (Box's M test) were tested in the research variables in all assessment points. When the Mauncly test of sphericity was violated ($p < .05$), the Greenhouse-Geisser correction was used to correct for this prevalent violation. Effect sizes (ES) were assessed using the Partial Eta-Squared (η_p^2) and the Cohen's d measures. The η_p^2 indicates the proportion of variance in the dependent variable accounted by some effect (small ES = .01; medium ES = .09; large ES = .25) while the Cohen's d (Cohen 1988) assesses the size of difference between two means in SD units (small ES = .20; medium = .50; large = .80 and above). The η_p^2 provides valuable information when comparing effects between studies with similar experimental designs while the Cohen's d seems more robust to these differences (see Laken 2013). All statistical analyses were performed with SPSS 25. The Cohen's d was computed using an online calculator (www.socscistatistics.com) and the η_p^2 was converted into Cohen's d with De Coster (2012) spreadsheet.

Results

Normality Test Attrition Analyses

The Kolmogorov-Smirnov test indicated that all study variables were non-normally distributed ($p < .05$). Nevertheless, F-test is supported to be robust to violations of normality (Blanca et al. 2017) and thus, ANCOVAs were conducted in subsequent analyses. Attrition analyses indicated no group differences in baseline assessment (Time 1), for Attitudes towards refugees $t(312) = .783$, $p = .624$; Intolerance, $t(45.91) = .154$, $p = .88$; Xenophobia, $t(342) = -.94$, $p = .35$ and Generative Altruism, $t(341) = -.901$, $p = .40$. Nevertheless, group differences were found in Attitudes towards the Greek people, $t(31.33) = -2.47$, $p < .05$, with participants demonstrating higher levels of positive attitudes ($M_{participants} = 26.20$, $M_{noparticipants} = 24.00$).

Descriptive Statistics and Bivariate Correlations among the Study Variables

Means, standard deviations and bivariate correlations among the study variables are presented in Table 1. Regarding the effect of demographics on the research variables in all waves, a multivariate analysis of variance indicated no significant effects of gender, $F(15, 103) = 1.09$, $p = .374$, Wilk's $\Lambda = .77$, grade, $F(30, 206) = .97$, $p = .513$, Wilk's $\Lambda = .77$, paternal educational level, $F(75, 535) = .89$, $p = .727$, Wilks' $\Lambda = .55$, or maternal educational level, $F(75, 497) = .76$, $p = .93$, Wilks' $\Lambda = .60$.

Table 1 Means, standard deviations and Pearson zero-order bivariate correlations among the study variables in each assessment time

Variables	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.ATR1	17.82	5.15	–														
2.ATR2	19.16	5.21	.65**	–													
3.ATR3	18.35	5.30	.58**	.64**	–												
4.XEN1	5.66	1.98	-.42**	-.30**	-.28**	–											
5.XEN2	5.36	1.91	-.28**	-.36**	-.27**	.56**	–										
6.XEN3	5.50	2.04	-.35**	-.32**	-.36**	.49**	.55**	–									
7.INT1	9.83	2.54	-.42**	-.34**	-.32**	.26**	.19**	.26**	–								
8.INT2	9.71	2.58	-.41**	-.55**	-.49**	.25**	.31**	.32**	.57**	–							
9.INT3	9.82	2.54	-.37**	-.42**	-.56**	.21**	.22**	.40**	.50**	.59**	–						
10.ALTR1	19.30	5.63	.40**	.31**	.38**	-.19**	-.17**	-.30**	-.40**	-.33**	-.36**	–					
11.ALTR2	19.20	6.47	.34**	.50**	.45**	-.12**	-.22**	-.25**	-.22	-.44**	-.34**	.59**	–				
12.ALTR3	18.43	6.19	.26**	.35**	.50**	-.08	-.09	-.23**	-.13*	-.30**	-.41**	.58**	.62**	–			
13.ATG1	26.21	3.08	.01	-.01	-.01	.06	.20**	.03	.01	.01	-.04	.19**	.07	.17**	–		
14.ATG2	26.13	3.55	-.04	.03	.10	.04	.05	.00	-.01	-.04	-.11	.14*	.15**	.12*	.56**	–	
15.ATG3	25.88	3.66	-.07	.01	.11	.03	.05	-.02	-.08	-.08	-.16**	.12*	.05	.22**	.51**	.60**	–

ATR1 (Attitudes towards Refugees Time 1); ATR2 (Attitudes Towards Refugees Time 2); ATR3 (Attitudes Towards Refugees Time 3); XEN1 (Xenophobia Time 1); XEN2 (Xenophobia Time 2); XEN (Xenophobia Time 3); INT1 (Intolerance Time 1); INT2 (Intolerance Time 2); INT3 (Intolerance Time 3); ALTR1 (Altruism Time 1); ALTR2 (Altruism Time 2); ALTR3 (Altruism Time 3); ATG1 (Attitudes towards Greeks Time 1); ATG2 (Attitudes Towards Greeks Time 2); ATG3 (Attitudes Towards Greeks Time 3)

* $p < .05$. ** $p < .01$

Covariance Analyses

Effect of FP-GR on Attitudes Towards Refugees in the Two Study Groups

Homogeneity of variances [$F^{1st\ wave}(1, 312) = .004, p > .05$; $F^{2nd\ wave}(1, 312) = .20, p > .05$; $F^{3rd\ wave}(1, 312) = .08, p > .05$], and covariances (Box's $M = .517, p > .05$) was confirmed by the data. The Greenhouse-Geisser correction [Mauchly's test: $\chi^2(2) = 6.78, p < .05$] showed a significant interaction between time and group, $F(1.96, 604.83) = 5.29, p < .01, \eta_p^2 = .03$, Cohen's $d = .34$, whilst controlling for the effect of the interventionists. More specifically, the IG demonstrated significantly higher levels of positive attitudes towards refugees in the 2nd wave of assessment and the 3rd wave respectively compared to the 1st wave ($M_{intervention\ time2-time1} = 2.08, p < .001$, Cohen's $d = .42$; $M_{intervention\ time3-time1} = .93, p < .05$, Cohen's $d = .18$). Nevertheless a significant but small reduction in positive attitudes towards refugees was found from the 2nd to the 3rd wave of assessment for this treatment group ($M_{intervention\ time2-time3} = 1.14, p < .01$, Cohen's $d = .22$). On the contrary, no significant attitudinal change across time was supported for the CG ($p > .05$). Furthermore, the IG showed significantly higher levels of positive attitudes towards refugees compared to the CG on the 2nd wave of assessment ($M_{intervention-control} = 1.62, p < .01$, Cohen's $d = .33$).

Effect of FP-GR on Attitudes towards Greek People in the Two Study Groups

Although homogeneity of covariances was supported (Box's $M = 8.51, p > .05$), homogeneity of variances was not confirmed across all assessment points [$F^{1st\ wave}(1, 312) = .021, p = .86$; $F^{2nd\ wave}(1, 312) = .009, p = .93$; $F^{3rd\ wave}(1, 312) = 5.87, p < .02$]. Results [Mauchly's test: $\chi^2(2) = 1.80, p > .05$] showed a no significant interaction between time and group $F(2, 618) = 1.30, p = .27, \eta_p^2 = .00$, while controlling for the effect of the interventionists. Additionally, no significant main effects were supported for time [$F(2, 618) = .64, p > .05$] and group [$F(1, 309) = .051, p > .05$] respectively.

Effect of FP-GR on Xenophobia in the Two Study Groups

Homogeneity of variances was not equally supported across all assessment points [$F^{1st\ wave}(1, 312) = 6.73, p < .01$; $F^{2nd\ wave}(1, 312) = .60, p > .05$; $F^{3rd\ wave}(1, 312) = 2.56, p > .05$]. Instead homogeneity of covariances was confirmed (Box's $M = 10.35, p > .05$). The Greenhouse-Geisser correction [Mauchly's test: $\chi^2(2) = 8.27, p < .05$] indicated a significant interaction between time and group, $F(1.95, 602.05) = 6.56, p < .01, \eta_p^2 = .04$, Cohen's $d = .41$, whilst controlling for the effect of the interventionists. More specifically the IG demonstrated significantly lower levels of xenophobia in the 2nd wave of assessment and the 3rd wave respectively compared

to 1st wave ($M_{\text{intervention}}^{\text{time1} - \text{time2}} = .61, p < .001$, Cohen's $d = .29$; $M_{\text{intervention}}^{\text{time1} - \text{time3}} = .45, p < .01$, Cohen's $d = .23$). Additionally no significant change in xenophobic attitudes was found between the 2nd and the 3rd wave for this treatment group ($M_{\text{intervention}}^{\text{time2} - \text{time3}} = -.15, p > .05$). The CG showed no significant attitudinal change across time ($p > .05$). Furthermore, the IG demonstrated significantly lower levels of xenophobic attitudes compared to the control group on the 2nd wave of assessment ($M_{\text{intervention-control}} = -.49, p < .05$, Cohen's $d = .25$).

Effect of FP-GR on Intolerance in the Two Study Groups

Homogeneity of variances [$F^{1\text{st wave}}(1, 312) = .028, p > .05$; $F^{2\text{nd wave}}(1, 312) = .215, p > .05$; $F^{3\text{rd wave}}(1, 312) = .090, p > .05$] and covariances (Box's $M = 12.81, p < .05$) was confirmed by the data. Results [Mauchly's test: $\chi^2(2) = 4.00, p > .05$] indicated a significant interaction between time and group, $F(2, 618) = 4.23, p < .05, \eta_p^2 = .03$, Cohen's $d = .34$, while controlling for the effect of the interventionists. The IG showed significantly lower levels of intolerance towards refugees on the 2nd wave of assessment compared to the 1st wave respectively ($M_{\text{intervention}}^{\text{time1-time2}} = .46, p < .05$, Cohen's $d = .18$). The CG demonstrated no significant change in intolerant attitudes across time ($p > .05$). Significant differences between the two treatment groups were found in neither time point of assessment ($p > .05$).

Effect of FP-GR on the Generative Altruism in the Two Study Groups

Homogeneity of variances [$F^{1\text{st wave}}(1, 312) = .01, p > .05$; $F^{2\text{nd wave}}(1, 312) = .17, p > .05$; $F^{3\text{rd wave}}(1, 312) = 1.83, p = .18$] and covariances (Box's $M = 5.19, p > .05$) was supported. Results showed [Mauchly's test: $\chi^2(2) = 5.80, p > .05$] a significant interaction between time and group, $F(2, 618) = 7.94, p < .001, \eta_p^2 = .03$, Cohen's $d = .34$. The IG demonstrated no significant change in general altruism between the 1st and the 2nd wave of assessment ($M_{\text{intervention}}^{\text{time1} - \text{time2}} = -.95, p > .05$). Surprisingly, a significant reduction in general altruism was found for this treatment group from the 2nd wave of assessment to the 3rd wave respectively ($M_{\text{intervention}}^{\text{time2} - \text{time3}} = -1.03, p < .05$, Cohen's $d = .15$). Additionally, analyses of the data indicated a significant reduction of the altruistic behavior for the CG across time ($M_{\text{control}}^{\text{time2} - \text{time1}} = -1.4, p < .01$, Cohen's $d = .22$) and Time 3 ($M_{\text{control}}^{\text{time3} - \text{time1}} = -.44, p < .01$, Cohen's $d = .31$).

Discussion

The current study evaluated the longitudinal impact of the FP-GR, a structured theory-driven intervention program designed

to prevent prejudice and generate children's positive attitudes towards refugees. Our findings provide initial support regarding the positive short-term impact of FP-GR on children's general attitudes towards refugees. The long-term effectiveness of the program, though significantly smaller in magnitude compared to the short-term, can also be supported by the results of this study. Additionally, discriminant validity of the FP-GR was confirmed as IG participant's attitudes towards Greek people remained stable. Furthermore, no statistical longitudinal attitudinal change was detected for the CG. These effects appear to be in agreement with previous results of studies aiming to reduce prejudicial attitudes and cultivate social and emotional skills (e.g. Brenick et al. 2019; Turner and Brown 2008). Indeed, most of the reinforcement programs designed to promote positive intergroup attitudes report only low to moderate effects (e.g. mean $ES = .08$; Okoye-Johnson 2011) which often tend to weaken or disappear after a time delay (Beelmann and Heinemann 2014). Our results, although small in effect size (mean $ES = .30$), should be considered educationally significant given the cut-off .25 ES value proposed by Tallmadge (1977). Additionally, it should be mentioned that they are in line with suggestions pertaining that in a low intensity and cost program, even a slight attitudinal change could have impressive implications and practical importance (Ellis 2010). Furthermore, the FP-GR showed a positive moderate longitudinal impact on children's xenophobic attitudes towards refugees. More specifically, the IG demonstrated significantly lower levels of xenophobia in the short-term and in the long-term, respectively (mean $ES = .26$). Furthermore, no statistical longitudinal attitudinal change was found for the CG. These findings could have implications for the effectiveness of FP-GR in combating xenophobia, the "psychological state of hostility or fear towards outsiders" (Reynolds and Vine 1987, p.28). It has been documented that misconceptions and threat narratives about refugees widespread in society can evoke feelings of anxiety and fear in children and fuel anti-immigrant attitudes (Glen et al. 2020). The effectiveness of direct and manipulated exposure to anti-bias information in reducing negative generalizations and cultivating empathy is well supported (Bigler and Liben 2006). Indeed, media/instruction anti-prejudice interventions in young children have yielded promising results (see Aboud et al. 2012). The FP-GR, acknowledging that ignorance is a key factor of xenophobia (Van der Westhuizen and Kleintjes 2015) provided children with the accurate and developmentally appropriate knowledge to dispel false myths and to reconstruct possible distorted beliefs about refugees. To our knowledge, this is the first study that assesses the effectiveness of an anti-prejudice program in combating the xenophobic attitudes of the majority children using a psychometrically sound instrument. Thus, our results not only contradict previous research supporting the weak impact of multicultural interventions on changing negative ethnic stereotypes and attitudes

(e.g. Killen et al. 2011) but also confirm the potential efficacy of the FP-GR in decreasing xenophobic attitudes.

The FP-GR demonstrated a negative but small in effect size ($ES = .18$) short-term impact on children's intolerant attitudes towards refugees. Contrary to our hypothesis, the mean difference between the initial assessment and that conducted 3 months after the end of the intervention was found to be insignificant. Tolerance towards immigrants is understood as a basic democratic principle that is based on the abstract understanding and endorsement of equality between immigrants and non-immigrants (Rapp 2017). The weak impact of the FP-GR on changing participants' intolerant attitudes towards refugees could be attributed to age differences regarding abstract and critical level of thought (Walker et al. 2018) as well as the limited time of the program to cultivate children's abstract thinking about equality and social justice, two concepts closely related to tolerance. Self-report measures are not also sensitive enough to determine whether the profile of attitudinal change manifested by participants reflects changes in actual behavior (Brenick et al. 2019). That means that the FP-GR may be effective in combating children's intolerant attitudes towards refugees but still in an implicit and unconscious way.

Contrary to our expectations, the FP-GR did not seem to have any significant positive short-term impact on IG's altruism. Unexpectedly, IG participants showed a significant reduction in their motivation to behave altruistically, 3 months after the completion of the intervention. Additionally, the CG demonstrated significantly lower levels in altruism in the long run. As it has been already mentioned, the FP-GR aimed at fostering children's altruistic behavior by implementing several well-structured empathy inducing activities. More specifically, it was anticipated that by stimulating participants' empathetic concern towards refugees, there would be an improvement in their prosocial helping or caring for others (Fourie et al. 2017). The unstable levels of altruism found in both groups contradict theories supporting the biological roots of altruism (Rajhans et al. 2016). Furthermore, the insignificant impact of the program on enhancing children's generative altruism, conceptualized as the non-conflictual pleasure of fostering the success and/or welfare of another (Büssing et al. 2013) provides support for the contextual nature of altruistic behavior. More specifically, our results confirm previous findings suggesting that individual's altruistic actions are not intrinsic and uniform but depend partially on the societally shared conceptions and the established positive relationships with particular social identity groups (Jenkins et al. 2018). We acknowledge that participants had few opportunities for intergroup contact, which has been proposed as the most effective way to reduce prejudice (Brown and Hewstone 2005; Pettigrew and Tropp 2006), as they were enrolled in schools that did not operate reception structures for refugee education. Results could also be explained from a developmental standpoint, given that acting altruistically requires acquisition of

Theory of Mind, an ability necessary to attribute mental states and put oneself into another's shoes (Wellman et al. 2001; Wellman and Liu 2004). Elementary students are not yet able to perform hypothetical reasoning and it is not before 12 years of age that they acquire the ability to reason abstractly (Piaget 1972, 1981; Rafetseder et al. 2013). Additionally, given that our study assessed neither empathy nor altruism specifically towards refugees (as to our knowledge there are no available research measures), we do not have enough evidence to provide support for the Empathy-Attitudes-Action model (Batson et al. 2002), which states that increasing empathy towards one social out-group (e.g. refugees), improves positive attitudes and in turn promotes prosocial behaviour (e.g. altruism) towards that group.

Taken together, these findings confirm that the FP-GR not only influenced participants' general attitudes towards refugees, but also affected their xenophobic and intolerant attitudes. The small effect sizes found in our study could be partially attributed to several reasons. First, as it has been already mentioned, the FP-GR is a brief and low intensity preventative anti-prejudice program. In a meta-analysis conducted by Okoye – Johnson in 2011, it was found that multicultural programs were longitudinally effective in changing deep-rooted stereotypes, when embedded to the school everyday life, from curriculum and instructional materials through every component of the school experience. Second, following the ethical principle of respect for person, parental consent was a requirement for children's participation in this research. It is, then, possible that parents who permitted their child participation in the present multicultural – antiracist program, already demonstrated low levels of xenophobic and intolerant attitudes towards refugees. According to the literature, parents' subtle prejudicial cognition and automatic behavior are positively associated with children's prejudice towards an out-group (Pirchio et al. 2018). Thus, the effectiveness of the FP-GR in promoting intergroup understanding and acceptance was possibly assessed in a sample of children with pre-existing positive perceptions of cultural diversity, giving little room for improvement. Finally, the FP-GR is a program implemented exclusively at the individual level. Hence, it could be supported that the small effect sizes indicate those pertaining the effectiveness of multilevel – individually and systematically- approaches in addressing race discrimination (Greco et al. 2010). More specifically, according to ecological developmental models, development should be addressed as a result of complex links between individuals and multiple, causally interactive proximal socialization contexts (Bronfenbrenner 1977). Thus, it is possible that the effectiveness in changing participants' intergroup attitudes of anti-prejudice programs is moderated by various social factors associated with children's development of racism and discrimination such as media negative stereotypes about the out-group and teachers lack of intercultural competence (Hajisoteriou et al. 2019).

This intervention study contributes to existing literature in several ways. First, taking into account the critical role that a reliable and valid measurement plays in intervention studies (Gabrera-Nguyen 2010), the outcomes of the present program were assessed using developmentally appropriate and psychometrically sound self-report instruments. Second, given the multifaceted character of attitudes (Rutland and Killen 2015), several measures and indicators were employed in order to investigate and capture any potential attitudinal changes. Third, in contrast to the majority of studies assessing the immediate effects of interventions on intergroup attitudes (Beelmann and Heinemann 2014), our research considered the short-term as well as the long-term stability of the outcomes. Finally, to our knowledge, this is one of the few evaluations of a preventing program that prepares children to engage in a meaningful way with their outgroup refugee peers (e.g. Turner and Brown 2008).

Limitations and Recommendations

This study has also some limitations that are worth mentioning. First, we did not control for participants' pre-existing levels of contact with refugees. This is an important limitation given that being socially engaged with many cross-cultural companions has been linked to unbiased racial attitudes (Aboud et al. 2003). Furthermore, the implementation of the FP-GR took place by professionals, culturally competent and expertized in educational interventions towards diversity. Hence, the effectiveness of this program might not be confirmed in the future, especially when implemented by low levels of intercultural literacy and sensitivity teachers. Indeed, teachers' negative implicit racial attitudes not only serve as a potential predictor of implicit prejudicial attitudes displayed by children (Vezzali et al. 2012) but could also act as a "hidden and well camouflaged" obstacle for the successful implementation of anti-prejudiced school based programs. Additionally, children's attitudes towards refugees were assessed via self-report instruments. Although this type of measures represent one of the most frequently used research tools in relevant psychological research, they tend to be susceptible to response bias. In our case participants might have felt accountable for the anti-prejudice information they have received and thus provided socially desirable responses. Self-report measures also evaluate only explicit attitudes (Axt 2018). Nevertheless, research has shown that implicit attitudes, activated by the mere presence of the attitude object without the person's full awareness or control, provide more information than explicit attitudes (Dovidio et al. 2002). The Implicit Association Test (IAT) and several "priming" methods (for a description see McKeague et al. 2013) have the potential to provide a useful way of measuring children's unconscious racial attitudes. Further, given that attitudes have a cognitive, emotional and behavioral component, well-conceived and well-executed behavioral observations can

offer a valuable insight into them (Furr and Funder 2007). Taking all these into consideration, a multimethod approach could be examined, if researchers are really interested in determining the true change in attitudes. Finally, children's empathetic and altruistic behavior towards refugees was not evaluated in this study, as to our knowledge there are no available instruments reported in the literature.

Despite the limitations of this study, the FP-GR did have an impact on children's general, xenophobic and intolerant attitudes towards refugees. The next step is to improve the longitudinal effectiveness of this particular intervention and similar anti-prejudice programs, by taking into consideration the following suggestions. First, it seems important that multicultural and moral development programs be fundamentally embedded in the everyday school curriculum. Second, where possible, these programs could be combined with structured intergroup contact learning experiences, which have been found to promote intergroup relations (Liebkind et al. 2019). Finally, the professional development of teachers on cultural diversity issues should be an educational priority (Giovazolias et al. 2019; Szelei et al. 2019). Overall, investing in preventive school-based programs that prepare host society students to interact successfully with refugee background children results in the peace-building and the social cohesion of multicultural societies. Our study supports that the FP-GR is indeed a promising anti-prejudice program that could be implemented in elementary school settings.

Funding This research has been co-financed by the Operational Program "Human Resources Development, Education and Lifelong Learning" and is co-financed by the European Union (European Social Fund) and Greek national funds (MIS 5006484, code 4507).

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

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

References

- Aboud, F., Mendelson, M., & Purdy, K. (2003). Cross-race peer relations and friendship quality. *International Journal of Behavioral Development*, 27(2), 165–173.

- Aboud, F. E., Tredoux, C., Tropp, L. R., Brown, C. S., Niens, U., & Noor, N. M. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review*, 32(4), 307–336.
- Axt, J. R. (2018). The best way to measure explicit racial attitudes is to ask about them. *Social Psychological and Personality Science*, 9(8), 896–906.
- Barnett, M. (2018). Human rights, humanitarianism, and the practices of humanity. *International Theory*, 10(3), 314–349.
- Batson, C., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes, and action: Can feeling for a member of a stigmatized group motivate one to help the group? *Personality and Social Psychology Bulletin*, 28, 1656–1666.
- Batson, C. D., Ahmad, N., & Lishner, D. A. (2009). Empathy and altruism. In S. J. Lopez & C. R. Snyder (Eds.), *Oxford library of psychology. Oxford handbook of positive psychology* (pp. 417–426). Oxford: Oxford University Press.
- Batson, C. D., Lishner, D. A., & Stocks, E. L. (2015). The empathy–altruism hypothesis. In D. A. Schroeder & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior* (pp. 259–268). Oxford: Oxford University Press.
- Beelmann, A., & Heinemann, K. S. (2014). Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs. *Journal of Applied Developmental Psychology*, 35(1), 10–24.
- Beißert, H., Gönültaş, S., & Mulvey, K. L. (2019). Social inclusion of refugee and native peers among adolescents: It is the language that matters! *Journal of Research on Adolescence*, 30(1), 219–233.
- Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, 46, 5–68.
- Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 34, pp. 39–89). Elsevier: San Diego.
- Blanca, M., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 552–557.
- Brenick, A., Lawrence, S. E., Carvalho, D., & Berger, R. (2019). Teaching tolerance or acting tolerant? Evaluating skills-and contact-based prejudice reduction interventions among Palestinian-Israeli and Jewish-Israeli youth. *Journal of School Psychology*, 75, 8–26.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32(7), 513–531.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 255–343). San Diego: Elsevier Academic Press.
- Buchanan, Z. E., Abu-Rayya, H. M., Kashima, E., Paxton, S. J., & Sam, D. L. (2018). Perceived discrimination, language proficiencies, and adaptation: Comparisons between refugee and non-refugee immigrant youth in Australia. *International Journal of Intercultural Relations*, 63, 105–112.
- Büssing, A., Kerksieck, P., Günther, A., & Baumann, K. (2013). Altruism in adolescents and young adults: Validation of an instrument to measure generative altruism with structural equation modeling. *International Journal of Children's Spirituality*, 18(4), 335–350.
- Cameron, L., Rutland, A., Brown, R., & Douch, R. (2006). Changing children's intergroup attitudes toward refugees: Testing different models of extended contact. *Child Development*, 77(5), 1208–1219.
- Cavounidis, J. (2013). Migration and the economic and social landscape of Greece. *South-Eastern Europe Journal of Economics*, 11(1), 59–78.
- Chen, I., & Forbes, C. (2014). Reflective writing and its impact on empathy in medical education: Systematic review. *Journal of Educational Evaluation for Health Professions*, 11(20), 1–6.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge Academic.
- De Coster, J. (2012). Spreadsheet for converting effect size measures. Available from <http://www.stathelp.com/spreadsheets>
- De Paola, M., & Brunello, G. (2016). *Education as a Tool for the Economic Integration of Migrants*, IZA discussion papers, no. 9836, Institute for the Study of labor (IZA), Bonn.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62–68.
- Earnest, J., Mansi, R., Bayati, S., Earnest, J. A., & Thompson, S. C. (2015). Resettlement experiences and resilience in refugee youth in Perth, Western Australia. *BMC Research Notes*, 8(1), 236.
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.
- Fourie, M. M., Subramoney, S., & Gobodo-Madikizela, P. (2017). A less attractive feature of empathy: Intergroup empathy bias. *Empathy: An evidence based interdisciplinary perspective. INTECH* (London), 45–61.
- Furr, R. M., & Funder, D. C. (2007). Behavioral observation. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (p. 273–291). The Guilford Press.
- Gabrera-Nguyen, P. (2010). Author guidelines for reporting scale development and validation results in the journal of the Society for Social Work and Research. *Journal of the Society for Social Work and Research*, 1(2), 99–103.
- Giovazolias, T., Syggelaki, E., & Papastilianou, A. (2019). Effectiveness of a teachers' training program in bullying on their Core self evaluations, relationship with students and work engagement. *Journal of Classroom Interaction*, 54(2), 4–23.
- Glen, C., Taylor, L. K., & Dautel, J. B. (2020). Promoting Prosocial behavior toward refugees: Exploring the empathy-attitude-action model in middle childhood. In *Children and Peace* (pp. 71–87). Springer, Cham.
- Greco, T., Priest, N., & Paradies, Y. (2010). *Review of strategies and resources to address race-based discrimination and support diversity in schools*. Victorian Health Promotion Foundation.
- Hajisoteriou, C., Maniatis, P., & Angelides, P. (2019). Teacher professional development for improving the intercultural school: An example of a participatory course on stereotypes. *Education Inquiry*, 10(2), 166–188.
- Hangartner, D., Dinas, E., Marbach, M., Matakos, K., & Xefteris, D. (2019). Does exposure to the refugee crisis make natives more hostile? *American Political Science Review*, 113(2), 442–455.
- Hansen, O.-P. M., & Legge, S. (2016). Drawbridges down: Altruism and immigration preferences. *The international platform of Ludwig-Maximilians University's center for economic studies and the IFO institute*, 6204, 1–43.
- Hollingsworth, L. A., Didelot, M. J., & Smith, J. O. (2003). REACH beyond tolerance: A framework for teaching children empathy and responsibility. *The Journal of Humanistic Counseling, Education and Development*, 42(2), 139–151.
- International Rescue Committee (IRC). 2019 Year in Review. Retrieved December 6, 2019 from <https://www.rescue.org/announcement/2019-year-review>.
- Jenkins, A. C., Karashchuk, P., Zhu, L., & Hsu, M. (2018). Predicting human behavior toward members of different social groups. *Proceedings of the National Academy of Sciences*, 115(39), 9696–9701.
- Killen, M., & Smetana, J. G. (2015). Origins and development of morality. In R. M. Lerner & M. E. Lamb (Eds.), *Handbook of Child*

- Psychology and Developmental Science* (Vol. 3, 7th ed., pp. 701–749). New York: Wiley-Blackwell.
- Killen, M., Rutland, A., & Ruck, M. (2011). Promoting equity and justice in childhood: Policy implications. *Social Policy Report*, 25(4), 1–25.
- Koronaoui, A., Lagos, E., Sakellariou, A., Kymionis, S., & Chiotaki-Poulou, I. (2015). Golden Dawn, austerity and young people: The rise of fascist extremism among young people in contemporary Greek society. *The Sociological Review*, 63, 231–249.
- Kourkoutas, E., & Giovazolias, T. (2015). School-based counselling work with teachers: An integrative model. *European Journal of Counselling Psychology*, 3(2), 137–158.
- Laken, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Liebkind, K., Mäkinen, V., Jasinskaja-Lahti, I., Renvik, T. A., & Solheim, E. F. (2019). Improving outgroup attitudes in schools: First steps toward a teacher-led vicarious contact intervention. *Scandinavian Journal of Psychology*, 60(1), 77–86.
- McKeague, L., O'Driscoll, C., Hennessy, E., & Heary, C. (2013). Using implicit measures to explore children's intergroup attitudes: Methodological and practical considerations for researchers. *International Journal of Social Research Methodology*, 18(1), 1–13.
- Mead, G. H. (1934). *Mind, self, and the society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.
- Mickelson, R. A., & Nkomo, M. (2012). Integrated schooling, life course outcomes, and social cohesion in multiethnic democratic societies. *Review of Research in Education*, 36(1), 197–238.
- Nesdale, D. (2008). Peer group rejection and children's intergroup prejudice. In S. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 32–46). Oxford: Oxford University Press.
- Obeid, S., Haddad, C., Salame, W., Kheir, N., & Hallit, S. (2019). Xenophobic attitudes, behaviors and coping strategies among Lebanese people toward immigrants and refugees. *Perspectives in Psychiatric Care*, 55, 710–717.
- Okoye-Johnson, O. (2011). Does multicultural education improve students' racial attitudes? *Implications for closing the achievement gap. Journal of Black Studies*, 42(8), 1252–1274.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783.
- Piaget, J. (1972). Intellectual evolution from adolescence to adulthood (J. Bliss & H. Furth, trans.). *Human Development*, 15(1), 1–12.
- Piaget, J. (1981). *The psychology of intelligence*. Totowa: Littlefield, Adams.
- Pirchio, S., Passiatore, Y., Panno, A., Maricchiolo, F., & Carrus, G. (2018). A Chip off the old block: Parents' subtle ethnic prejudice predicts Children's implicit prejudice. *Frontiers in Psychology*, 9, 110.
- Plexousakis, S., Kourkoutas, E., Giovazolias, T., Chatira, K., & Nikolopoulos, D. (2019). School bullying and post-traumatic stress disorder symptoms: The role of parental bonding. *Frontiers in Public Health*, 7, 75.
- Pugh, K., Every, D., & Hattam, R. (2012). Inclusive education for students with refugee experience: Whole school reform in a south Australian primary school. *The Australian Educational Researcher*, 39(2), 125–141.
- Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, 114(3), 389–404.
- Rajhans, P., Altvater-Mackensen, N., Vaish, A., & Grossmann, T. (2016). Children's altruistic behavior in context: The role of emotional responsiveness and culture. *Scientific Reports*, 6, 24089.
- Rapp, C. (2017). Shaping tolerant attitudes towards immigrants: The role of welfare state expenditures. *Journal of European Social Policy*, 27(1), 40–56.
- Reynolds, V., & Vine, I. (1987). *The socio-biology of ethnocentrism: Evolutionary dimensions of xenophobia, discrimination, racism, and nationalism*. London: Croom Helm.
- 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(1), 121–154.
- Sacramento, O. (2015). Schengen and the security obsession: Selective citizenship, exclusion and the ironies of control. The borders of Schengen (pp. 115–127). Retrieved from <https://www.scopus.com>.
- Sengupta, E., & Blessinger, P. (2018). Introduction to refugee education: Integration and acceptance of refugees in mainstream society. In *Refugee Education: Integration and Acceptance of Refugees in Mainstream Society* (pp. 3–15). Emerald publishing limited.
- Stevens, R. (2015). Role-play and student engagement: Reflections from the classroom. *Teaching in Higher Education*, 20(5), 481–492.
- Suleman, S., Garber, K. D., & Rutkow, L. (2018). Xenophobia as a determinant of health: An integrative review. *Journal of Public Health Policy*, 39(4), 407–423.
- Szaflarski, M., & Bauldry, S. (2019). The effects of perceived discrimination on immigrant and refugee physical and mental health. In R. Frank (Ed.), *Immigration and Health* (pp. 173–204). Emerald Publishing Limited.
- Szelei, N., Tinoca, L., & Pinho, A.-S. (2019). Professional development for cultural diversity: The challenges of teacher learning in context. *Professional Development in Education*, 1, 17.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon.
- Tallmadge, G. (1977). *The joint dissemination Review panel ideabook*. Washington, DC: National Institute of Education, U.S. Office of Education.
- Thomas, R. L., Chiarelli-Helminiak, C. M., Ferraj, B., & Barrette, K. (2016). Building relationships and facilitating immigrant community integration: An evaluation of a cultural navigator program. *Evaluation and Program Planning*, 55, 77–84.
- Turner, R. N., & Brown, R. (2008). Improving Children's attitudes toward refugees: An evaluation of a school-based multicultural curriculum and an anti-racist intervention. *Journal of Applied Social Psychology*, 38(5), 1295–1328.
- Ülger, Z., Dette-Hagenmeyer, D. E., Reichle, B., & Gaertner, S. L. (2018). Improving outgroup attitudes in schools: A meta-analytic review. *Journal of School Psychology*, 67, 88–103.
- UN General Assembly. (1951). *Convention Relating to the Status of Refugees*, 28 July, United Nations, Treaty Series, vol. 189, p. 137, available at: <https://www.refworld.org/docid/3be01b964.html>. Accessed 8 May 2020.
- United Nations High Commissioner for Refugees (UNHCR). (2019). *Figures at a Glance*. Retrieved October 12, from <https://www.unhcr.org/figures-at-a-glance.html>
- Van der Westhuizen, M., & Kleintjes, L. (2015). Social work services to victims of xenophobia. *Social Work*, 51(1), 115–134.
- Van Zalk, M. H. W., Kerr, M., Van Zalk, N., & Stattin, H. (2013). Xenophobia and tolerance toward immigrants in adolescence: Cross-influence processes within friendships. *Journal of Abnormal Child Psychology*, 41(4), 627–639.
- Vezzali, L., Giovannini, D., & Capozza, D. (2012). Social antecedents of children's implicit prejudice: Direct contact, extended contact, explicit and implicit teachers' prejudice. *European Journal of Developmental Psychology*, 9(5), 569–581.
- Vezzali, L., Di Bernardo, G. A., Stathi, S., Visintin, E. P., & Hewstone, M. (2019). Using intercultural videos of direct contact to implement vicarious contact: A school-based intervention that improves intergroup attitudes. *Group Processes & Intergroup Relations*, 22(7), 1059–1076.

- Walker, C. M., Hubachek, S. Q., & Vendetti, M. S. (2018). Achieving abstraction: Generating far analogies promotes relational reasoning children. *Developmental Psychology*, 54(10), 1833–1841.
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75(2), 523–541.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta- analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655–684.
- White, F. A., & Gleitzman, M. (2006). An examination of family socialization processes as moderators of racial prejudice transmission between adolescents and their parents. *Journal of Family Studies*, 12(2), 247–260.
- Williams, A., O'Driscoll, K., & Moore, C. (2014). The influence of empathic concern on prosocial behavior in children. *Frontiers in Psychology*, 5, 425.
- Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with non-normal missing data. *Sociological Methodology*, 30(1), 165–200.

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