

Sensitizing Children to the Social and Emotional Mechanisms involved in Racism: a program evaluation

Sofia Triliva^{a1} Tanya Anagnostopoulou^b & Georgios Vleioras^c

^a*Department of Psychology, University of Crete, Greece*

^b*Hellenic Institute of Psychology and Health, Thessaloniki, Greece.*

^c*Department of Psychology, University of Groningen, The Netherlands*

This paper describes and discusses the results of an intervention aiming to sensitize children to the social and emotional processes involved in racism. The intervention was applied and evaluated in 10 Greek elementary schools. The goals and the intervention methods of the program modules are briefly outlined and the results of the program evaluation are elaborated and discussed. Two-hundred students participated in the program and 180 took part in the pre-and-post-testing which assessed their ability to identify emotions associated with prejudice, discrimination and stereotypical thinking; to understand similarities and differences between people; and to develop perspective taking and empathic skills in relation to diverse others. Results indicate gains in all three areas of assessment although the increased ability to identify similarities between people can also be attributed to age/grade effects. The implications of the findings are discussed with regard to antiracism intervention methods and evaluation strategies.

Keywords: racism, program evaluation, anti-racism program, social and emotional processes

First submission December 30th 2013; Accepted for publication April 3rd 2014.

Introduction

Interventions for sensitizing children to issues related to diversity, social justice, and antiracism have been developed and applied in schools in many countries around the world (Banks, 2006; Benson & Poliner, 1997; Connolly, Fitzpatrick, Gallagher, & Harris, 2006; Derman-Sparks & Brunson-Phillips, 1997; Grant &

¹ Corresponding author. Email address: triliva@uoc.gr

Sleeter, 1998; Harris, Nikels, Mims, & Mims, 2007; McFarland, Siebold, Aguilar, & Sarmiento, 2007). An array of intervention paradigms have been formulated in order to understand and to confront racism and ethnocentrism, major challenges in contemporary societies (Boyd & Arnold, 2000; Sefa Dei, 2008). According to Levy Paluck (2006) curriculum intervention studies have been conducted, yet, evidence of their impact is rarely reported and their theoretical underpinnings are, often not clearly delineated. Boyd and Arnold (2000) point out that a great deal of discrepancy exists between theory, practice and teachers' understandings, and this adds to the difficulties in conducting empirical validation of the interventions. Moreover, in their review of violence prevention strategies, Eisner and Malti (2012) report that little progress has been made in "expanding the evidence base" (p. 170), "developing and testing prevention strategies" and in "improving knowledge of mechanisms and active components" (p.172). Yet, administrators, educators, psychologists and policy makers in many parts of the world are discussing and developing programmes to deal with social justice, inclusion, and the social and emotional issues involved in racism (Aveling, 2007; Bartolo, 2010; Bartolo, Borg, Cefai, & Martinelli, 2010; Daniels, 2010; Forlin, 2010). This sort of problematization is also happening in Greece.

Educators, parents, pupils and administrators in Greece have had to confront and deal with the complex problems that contemporary societies face when social and economic upheaval results in breakdown in social relations and increase in the incidence of racism, discrimination, and conflict. Our research and intervention originated from a call for support from a group of educators in Thessaloniki Greece. These teachers noted that their 'diverse' classrooms were difficult to manage and that they needed help in developing an intervention that would help the school-wide community to develop positive attitudes and behaviors towards 'different' others. The objective of this paper is to describe an antiracism intervention and an evaluation of its effectiveness. To the authors' knowledge, there are virtually no other empirically-validated anti-racism interventions for children described in the literature. Given the dearth of research on anti-racism interventions for children, this study thus aims to contribute to filling the gap in this area.

Program description

The intervention was first developed, applied and evaluated in order to help students and educators learn to respect and include children with differences in abilities within their classrooms (Triliva, Anagnostopoulou, Hatzinikolaou, Chimienti, & Mastorakou, 2009). It was later expanded to include differences and commonalities in gender, ethnicity/ culture, race, appearance, and religious beliefs. It was during this problematization phase regarding both theory and intervention strategies, that a paradigm to sensitize children to the social and emotional processes involved in racism, was designed and implemented. This development was based on a needs assessment conducted in the schools, as well as research which underlined the need for such programming (Koroni, Garagouni-Areou, Roussi-Vergou, Zafiropoulou, & Piperakis, 2008; Mattheou, Roussakis, & Theocharis, 2006; Motti-Stefanidi et al., 2008; Moutsios, 2003; Zoniou-Sideri & Vlachou, 2006). The newly expanded program called '*Neither better nor worse, just different*' (Triliva, Anagnostopoulou, & Hatzinikolaou 2008) consisted of seven modules aiming:

- To help students recognize and express emotions and to apply these skills in understanding instances of discrimination, bias, prejudice, and stigmatization;
- To celebrate similarities and differences;
- To promote social awareness—empathy, perspective taking, and valuing one’s own and others’ experiences as meaningful sources of knowledge;
- To understand the mechanisms of how stereotypes develop;
- To recognize ways through which stereotypes can be overturned and how marginalized people’s voices, histories and contributions can be reclaimed;
- To develop effective communication techniques in dealing with controversy, conflict and bias and in challenging inequality and oppression;
- To take positive social action against biases, stereotypes, and prejudice, and critically appraise how inequalities, marginalization, and oppression become part of the social milieu (critical consciousness raising).

Intervention methods

Experiential group exercises, inter-group contact, and a reflective practice approach, were the pedagogical methods applied to help the students analyze and think critically about their experiences regarding diversity, social justice, and racism. Experiential approaches have been used widely in diversity training (Levy Paluck, 2006). They allow for the co-construction and dialectical generation of meanings and enable learners to value that mutual understanding flourishes through dialogue and critical reflection; processes which are relational in nature (Heron, 1996; Postle, 1993). These forms of experience provide the participants with opportunities to learn from each other and to co-construct knowledge regarding prejudice from their particular context. The cooperative and collective nature of the activities privileged knowledge that develops from the students’ own initiatives, insights, and praxis (Freire, 2005). Engagement in experiential activities, along with the concomitant critical reflection, is reported to cultivate a sense of community and collective responsibility (Levy Paluck, 2006). The teachers who implemented the program activities were sensitized to their own biases, provided with supervision as to how to bring bias and discrimination issues into discussion through experiential activities.

In this paper we present the outcomes of the first four modules outlined above. The impact of the last three modules was not assessed because the evaluation methodology required was difficult to implement. The modules used in the inquiry reported here focus on knowledge acquisition and the development of social skills in dealing with discrimination and racism. The modules which were not included focus on inter-group dialogue and collaboration, and the application of social justice initiatives in school communities. Teachers implemented exercises consisting of role-plays, simulations, artistic expression, pantomimes, inter-group-contact and educational drama. For example, in ‘the sorceress’ exercise, the children were asked to use their imagination to draw a sorceress and to list all the attributes that characterize her. Using reflective questions, the teacher helped the children gain insight into the processes of stereotypical thinking and the socio-

emotional mechanisms involved. Finally, the children were asked to draw ‘the sorceress’ without the stereotypical characteristics that they originally attributed to her.

In accordance with the learning objectives of the first four modules, we expected that after the intervention, the students will have:

1. Increased their ability to recognize and name affective states involved in dealing with diversity, bias, prejudice and discrimination;
2. Increased their understanding that people are more similar than different;
3. Increased social perspective and empathy for diverse others; and
4. Increased knowledge that generalization is a basic component of biased behavior and the subsequent development of fear and avoidance responses.

Methodology

Study design

A pre–post program evaluation design was utilised in assessing possible changes in the variables defined above. The pre-intervention assessment took place a few weeks prior to the commencement of the antiracism program, and was carried out by the classroom teachers. The second assessment took place three weeks after the end of the intervention and was administered by volunteer research assistants. The intervention was implemented during one full academic year, and the experiential activities were carried out on a weekly basis for approximately two hours.

Participants

The sample consisted of 202 elementary school students from 10 public schools in Thessaloniki, Greece. Only 180 of the students participated in both pre and post assessments, as the remaining 22 were absent at either the pre or post assessment. Of the 180 students, 69 (38.3%) were 2nd graders (from classrooms of 15, 13, 23, and 18 students respectively), 51 (28.3%) were 3rd graders (from classrooms of 10, 16, and 11 students respectively), 17 (9.4%) were 4th graders (from one classroom), and 43 (23.9%) were 5th graders (from classrooms of 16, 12, and 15 students respectively). Eighty-eight (48.8%) were male and ninety-two (51.0%) were female. The vast majority of the students were Greek (85.6%) and the rest came from Albania, Georgia, Bulgaria or Turkey. The sample consisted of varying abilities and disabilities and religious backgrounds. The children’s parents provided informed consent for their children to participate in the study. The children’s responses were number coded and remained anonymous, adhering to the relevant ethical code of conduct. The classroom teachers received bi-weekly group supervision conducted by two psychologists and a specialist in the application of health promotion programming at school.

Measures

The pre and post intervention variables included students’ understanding that people are more similar than different, their ability to recognize and name affective states accurately, the development of social

perspective and empathy, and lastly, their comprehension that generalization is a basic component of biased behavior and the subsequent development of fear and avoidance responses.

Two short stories were developed to measure these variables. Both stories contained fictional animal characters, which were treated unfairly based on popular stereotypes. After reading each story, the students were asked a series of open-ended questions. This format allowed children to speak freely and to evaluate more precisely the particular nuances of their understandings. The type of moral reasoning vignettes used in this study have been found to be valid and reliable methods in assessing judgment and emotional attributions in social exclusion (Malti, Killen, & Gasser, 2012), children's moral judgments and emotion attributes in retaliation as opposed to unprovoked aggression (Gasser, Malti, & Gutzwiller-Helfenfinger, 2012), social understanding and perspective taking (Farr Darling, 2002; Selman, 2003; Woolgar, Steele, Steele, Yabsley, Fonagy, 2001), moral orientation in judgments dealing with culture and ethnic conflict (Garrod et al., 2003) and in solving complex social problems (Kim, 2012). The two stories used in the pre and post evaluation were the following:

Story 1.

Riri is one of four kittens of a stray cat living in a neighborhood of our city. The children of the neighborhood named all the kittens, and tried to find people who were interested in adopting them. All of the kittens were adopted with the exception of Riri. In contrast to the other kittens, Riri was black, and the people who adopted the other kittens refused to take her into their homes, saying: 'No, we don't want a black kitten; it will bring us bad luck!'

The respective questions were:

1. How do the kittens feel now that they have a place to stay?
2. How does Riri feel?
3. Does Riri look like the other kittens?

Questions 1 and 2 aimed at investigating the children's ability to recognize and name affective states accurately. Question 3 aimed at investigating the children's understanding that people are more similar than different.

Story 2.

Petros is a boy living in an apartment with his family. No animals are allowed in his apartment building, and he has never come into close contact with animals. One day, on his way to school, he encountered a dog named Lina. The dog barked and ran after him. Since then, whenever Lina approached him, Petros ran away as fast as he could. Moreover, since that incident, Petros did not only become afraid of Lina, but whenever he heard a dog barking he became frightened and started crying. His parents told him not to be afraid of dogs, but his fear did not subside. Petros started avoiding all places where it was likely to encounter dogs, such as the local park. He also started thinking of ways to defend himself in case he would be 'attacked' by a dog, and began carrying a stick that could be used in defense against a possible attack.

The respective questions were:

1. Do you think that Petros should avoid Lina? Why?
2. Do you think that Petros should avoid all dogs? Why?

Question 1 aimed at investigating social perspective taking and empathy. Question 2 aimed at investigating knowledge that generalization is a basic component of biased behavior and the subsequent development of fear and avoidance responses.

Procedures

The instructions to the teachers and research assistants were the following:

Inform the children that they will participate in a fantasy game, during which they will read stories and will answer some questions on their comprehension of the stories. Explain that there are no 'right' or 'wrong' answers. Each student is free to answer as he/she feels. If a student asks you, 'What should I write?' please respond, 'This is a fantasy game, you can think or imagine or guess. There are no right or wrong answers.' Make sure that each child answers for her/himself and that they do not copy other students' answers. If you are teaching first or second grade, you will have to read the stories aloud. If the students ask you to explain the meaning of a word, you should assist them, but please do not help students by giving them verbal or non-verbal hints (e.g., If they do not understand the question, 'How does Riri feel?' you shouldn't assist them by saying, 'How do you think Riri feels now that she is all alone? Is she happy or unhappy?'). Make sure that all students have answered all the questions. Keep in mind that questions are twofold, requiring both a simple answer and a justification of the answer. The students should answer both. They can also answer, 'I do not know.'

Data transformation

Data consisted of the children's responses to the questions that followed the stories. A coding scheme was developed *a posteriori* to code the answers to the open-ended questions. Two undergraduate and two graduate psychology students and one of the principal researchers coded the data and developed the categories. All answers were carefully scrutinized and all cases examined. The criteria developed were then applied by the research team to all questions to ensure that no responses were left out. All responses that were difficult to classify were thoroughly examined and the criteria were further refined until all agreed regarding the appropriateness of the category. Finally, one of the principal investigators ran a reliability check on a subsample of 30% of the data to ensure that they were correctly categorized. A concordance rate of 100% was tabulated on the classified responses. These procedures were undertaken to ensure the reliability of the coding and categories.

In the final coding scheme four categories were created for the answers to the questions:

'How do the kittens feel now that they have a place to stay?' and 'How does Riri feel?'

1. Vague emotion: This category included answers which describe emotional experience in a general and vague way (e.g., 'nice' or 'well')
2. Single emotion: This category included answers which accurately name one emotion (e.g., 'they are happy' or 'she is sad')

3. Complex emotion: This category included answers which accurately describe two or more affective states (e.g., ‘the kittens are happy to have a home but lonely because they do not have each other’)
4. Other: This category included answers that could not be coded in the previous categories (e.g., ‘I do not know,’ no answer, incomprehensible answer.)

Three categories were created for the answers to the question: ‘Does Riri look like the other kittens?’

1. Yes: This category included answers that could be summarized into ‘Yes, she is similar in everything but the color.’
2. No: This category included answers that could be summarized into ‘No, she is different.’ (e.g., ‘she is a black thing’ or ‘she is a monster’)
3. Other: This category included answers that could not be coded in the previous categories (e.g., ‘I do not know’ or ‘nothing’ or no answer or an incomprehensible answer)

Three categories were created for the answers to the questions: ‘Do you think that Petros is doing the right thing by avoiding Lina?’ and ‘Do you think that Petros is doing the right thing by avoiding all dogs?’

1. Yes: This category included affirmative answers (e.g., ‘yes’ or ‘sure’)
2. No: This category included negative answers (e.g., ‘no’ or ‘no way’)
3. Other: This category included answers that could not be coded in the previous categories (e.g., ‘I do not know’ or no answer or a contradictory answer or an incomprehensible answer)

Finally, three categories were created for the justification of the answers to the previous questions:

1. Fear: This category included answers revolving around fear of the dog (e.g., ‘Petros is shaking with fear’), or showing an attempt to avoid physical harm from a barking dog or getting contaminated by a sick dog (e.g., ‘She will attack him’ or ‘She has fleas’).
2. Understanding: This category contained answers which went beyond Petros and his behavior to commenting on issues related to aggression. Some answers included:
 - a. justified the dog’s aggressive behavior (e.g., ‘She didn’t want to harm him’ or ‘She only wanted to play’)
 - b. showed increased understanding of the harmful consequences of aggression (e.g., ‘If he attacks, she will strike back’)
 - c. showed understanding that Lina is one dog and her behavior characterizes only her (e.g., ‘Not all dogs are the same’ or ‘Some dogs are not aggressive’)
 - d. contained the possibility of a positive outcome (e.g., ‘he may become friends with Lina’).
3. Other: This category contained answers that could not be coded in the previous categories (e.g., ‘I do not know’, ‘nothing’, no answer, and an incomprehensible answer).

Recapitulation of hypotheses

Having explicated and delineated the children’s answers and how they were evaluated, it is important now to take a look at the research hypotheses and to operationalize and recapitulate them accordingly.

- ‘How do the kittens feel?’ and ‘How does Riri feel?’: it was hypothesized that after the intervention, fewer children would employ a *vague* emotion and more children would employ either a *single* or a *complex* emotion
- ‘Does Riri resemble the other kittens?’: It was hypothesized that more children would answer ‘Yes’ and fewer children would answer ‘No’
- ‘Do you think that Petros is doing the right thing by avoiding Lina?’ and ‘Do you think that Petros is doing the right thing by avoiding all dogs?’: It was hypothesized that more children would reply ‘No’ that Petros is not doing the right thing in avoiding Lina and in avoiding all dogs.
- Regarding the justification for the answers to the last two questions, it was hypothesized that fewer children would provide answers in the category of ‘*fear*’ and that more children would provide answers in the category of ‘*understanding*’.

Analyses

All answers were transformed into dichotomous format (0 for non-occurrence and 1 for occurrence of an answer). This means that new items were created, one for each possible answer to the original questions. For example, regarding the question: ‘Do you think that Petros is doing the right thing in avoiding Lina?’ three items were created, representing the occurrence vs. non-occurrence of the answers ‘yes,’ ‘no’ and ‘other’. Because of the dichotomous nature of the answers, the dependent variables modeled in the analyses hereunder, are the ‘success’ probabilities ϕ , i.e., the probabilities that an answer occurred, not the numbers of participants giving these answers. The hypotheses of the study were transformed accordingly to represent changes in probability of answers rather than in numbers of participants.

Typical inferential statistical analyses make a strong assumption that measurements are independent from each other. As Arnold (1992) explains, however, this assumption is particularly problematic in research in educational settings, where the measures are influenced not only by the participants’ characteristics, but also the classroom, school and/or the district qualities. Hence, since students’ answers are affected by these contexts, they are interrelated (McCoach, 2010). The statistical analysis that deals with this problem and allows the modeling of this non-independence is Hierarchical Linear Modeling (see Nezlek & Zyzniewski, 1998; Raudenbush & Bryk, 2002). A special case of these models is used in this study because the outcome is a binary variable taking on a value of either 0 or 1 (see Raudenbush & Bryk, 2002, 294-309). HLM software was used to run these analyses.

A presentation of the model that tests each of our hypotheses is described below. The dependent variable is ϕ , is a probability and takes values from the interval [0,1]. On account of this constraint in the values of ϕ , the following transformation is necessary: $\eta = \log [\phi/(1-\phi)]$ (1). The added value of this transformation is that the dependent variable η can now take on any real value. The level-1 equation of this model is: $\eta = \beta_0 + \beta_1(\text{INTERVENTION})$ (2). In this equation, INTERVENTION is a binary variable that takes the values 0 (denoting the condition before the intervention) or 1 (denoting the condition after the intervention). Note that, because no control group was included in this study, $\beta_1(\text{INTERVENTION})$ is

equivalent to $\beta_1(\text{TIME})$. Consequently, β_0 represents the individual-level η before the intervention (Time 1), while β_1 represents the effect of the intervention on the individual-level η (Time 2).

If the data were not nested, the level-1 equation would represent the entire model, resembling a logistic regression analysis (see Pampel, 2000). However, because the data are nested, two equations are added to take this into account: $\beta_0 = \gamma_{00} + u_0$ (3) and $\beta_1 = \gamma_{10}$ (4). These equations model the level-1 (individual-level) coefficients as a linear function of level-2 (class-level) coefficients. In particular, in equation (3), β_0 is modeled as a linear function of γ_{00} , which represents the grand η , i.e., the η for all the participating classes, and u_0 , which represents the related error term. Moreover, in equation (4), β_1 is modeled as a linear function of γ_{10} , which represents the change in the η for all the participating classes. Equation (3) was extended to include the effect γ_{01} of the participants' grade before the intervention (GRADE): $\beta_0 = \gamma_{00} + \gamma_{01}(\text{GRADE}) + u_0$ (5). Thus, the full model tested here comprised equations (2), (4), and (5).

A typical HLM analysis uses a t-ratio to test the null hypotheses that each of the related coefficients is 0. Here, we are only interested in the coefficients γ_{01} and γ_{10} . In particular, a γ_{10} that is different from 0 denotes that there is a statistically significant intervention effect. In addition, a γ_{01} that is different from 0 denotes that there is a statistically significant grade effect. If grade has a significant effect on β_0 , it is necessary to test whether the intervention effect is larger. The null hypothesis that $\gamma_{01} = \gamma_{10}$ can be tested by means of a chi-square test.

Results

Descriptive statistics

The following tables 1-3 illustrate the relative and the absolute frequencies at the different grades and at the pre- and post-intervention measurements for students' answers that are relevant to the study's hypotheses. There are three important pieces of information that are of interest in these tables. First, the relative frequencies in the pre-intervention measurements at different grades provide information on possible developmental trajectories. Second, the pre- and the post-intervention frequencies per grade provide information on the intervention effects. Finally, the comparison between the post-intervention and pre-intervention frequencies at the subsequent grade provide information on the relative contributions of development in conjunction with the intervention.

Table 1 displays the frequencies of the answers on the feelings of the kittens and Riri and the similarity between them. There was a considerable change in the feelings of the kittens between 2nd and 3rd grades, with a decrease in the incidence of *vague* emotions and an increase in the incidence of *simple* emotions. On the other hand, in the pre-intervention relative frequencies there was no unidirectional change in the feelings of Riri. The incidence of *vague* emotions decreases at 3rd grade, increases at 4th grade and decreases again at 5th grade. In the question on the similarity between Riri and the other kittens, negative answers peaked at 2nd grade, while positive answers peaked at 4th grade.

Table I. Relative (and absolute) frequencies in pre- and post-intervention measurements of feelings and similarity

Grade		<i>How do the kittens feel?</i>			<i>How does Riri feel?</i>			<i>Does Riri resemble the other kittens?</i>	
		Vague	Single	Complex	Vague	Single	Complex	Yes	No
2 nd	Pre	81.2 (56)	10.1 (7)	5.8 (4)	84.1 (58)	10.1 (7)	2.9 (2)	11.6 (8)	84.1 (58)
	Post	72.5 (50)	24.6 (17)	2.9 (2)	63.8 (44)	29.0 (20)	7.2 (5)	14.5 (10)	79.7 (55)
3 rd	Pre	62.7 (32)	31.4 (16)	3.9 (2)	56.9 (29)	41.2 (21)	2.0 (1)	25.5 (13)	45.1 (23)
	Post	37.3 (19)	45.1 (23)	15.7 (8)	39.2 (20)	31.4 (16)	29.4 (15)	66.7 (34)	13.7 (7)
4 th	Pre	64.7 (11)	29.4 (5)	5.9 (1)	82.4 (14)	17.6 (3)	0.0 (0)	41.2 (7)	41.2 (7)
	Post	58.8 (10)	23.5 (4)	17.6 (3)	64.7 (11)	35.3 (6)	0.0 (0)	52.9 (9)	41.2 (7)
5 th	Pre	44.2 (19)	39.5 (17)	9.3 (4)	62.8 (27)	20.9 (9)	9.3 (4)	30.2 (13)	58.1 (25)
	Post	53.5 (23)	32.6 (14)	7.0 (3)	53.5 (23)	30.2 (13)	11.6 (5)	65.1 (28)	20.9 (9)

The effect of the intervention varies according to age and according to the question, but not at all grade levels. For instance, at 2nd grade, there was an increase in the incidence of simple emotions in the questions concerning both Riri and the other kittens. At 3rd grade, on the other hand, there was an increase in the incidence of simple and of complex emotions regarding the kittens, while an increase in the incidence of complex emotions was only evident with regards to Riri. On the similarity question, the intervention effects were in the expected direction at all grades and they were more pronounced in grades 3 and 4.

Finally, the comparison between the post-intervention and pre-intervention measurements at the subsequent grade yields mixed results. In the kittens question, at 2nd grade, there was a decrease in the incidence of the vague emotion from pre- to post-intervention, and the post-intervention incidence was larger than the incidence at the pre-intervention at 3rd grade. This implies that the intervention effect was larger than the developmental effect. On the other hand, within the same age groups, there was an increase in the incidence of the simple emotion from pre- to post-intervention, but the post-intervention incidence was lower than the incidence at pre-intervention at 3rd grade. This suggests that the intervention effect is stronger than the developmental effect. Similar conclusions are reached concerning the similarity question. Thus, the developmental effect between 2nd and 3rd grade is larger than the intervention effect at 2nd grade, while the intervention effect at 3rd grade is larger than the developmental effect between 3rd and 4th grades.

Table 2 displays the frequencies of the answers on Petros, Lina and all dogs. The following developmental trajectories were observed: avoidance peaked at grade 5 and, somewhat less, at grade 3; generalization peaked at grade 4 and was less apparent at grade 5; fear as justification for avoiding Lina

peaked at grade 3 and, somewhat less, at grade 5; fear as a justification for avoiding all dogs peaked at grade 4; and finally, understanding as a justification for not avoiding Lina and all dogs peaked at grade 5. The intervention effects regarding Petros, Lina and all dogs were in the expected direction. The most pronounced effects in relation to avoiding Lina and all dogs and the related justifications appear at grade 5. Finally, the comparison between the pre and post-intervention at the next grade yields mixed conclusions. In some grades, the intervention effects were in a direction different from the development effect. For instance, in grade 2, the intervention brought a decrease in the incidence of the avoidance of Lina, while development led to an increase. In some other grades, the intervention and the development effects were in the same direction. For instance, at grade 4, the intervention and development brought a decrease in the incidence of the avoidance of all dogs, but the intervention effect was larger.

Table II. Relative (and absolute) frequencies in the pre- and post-intervention measurements of avoidance and generalization

Grade		<i>Do you think Petros is doing the right thing by avoiding Lina?</i>		<i>Why?</i>		<i>Do you think Petros is doing the right thing by avoiding all dogs?</i>		<i>Why?</i>	
		Yes	No	Fear	Understanding	Yes	No	Fear	Understanding
2 nd	Pre	24.6 (17)	69.6 (48)	39.1 (27)	17.4 (12)	20.3 (14)	76.8 (53)	23.2 (16)	34.8 (24)
	Post	20.3 (14)	75.4 (52)	20.3 (14)	43.5 (30)	10.1 (7)	85.5 (59)	11.6 (8)	46.4 (32)
3 rd	Pre	47.1 (24)	39.2 (20)	56.9 (29)	7.8 (4)	33.3 (17)	60.8 (31)	39.2 (20)	27.5 (14)
	Post	23.5 (12)	66.7 (34)	25.5 (13)	43.1 (22)	11.8 (6)	84.3 (43)	11.8 (6)	54.9 (28)
4 th	Pre	35.3 (6)	52.9 (9)	41.2 (7)	23.5 (4)	41.2 (7)	47.1 (8)	47.1 (8)	17.6 (3)
	Post	17.6 (3)	70.6 (12)	17.6 (3)	52.9 (9)	17.6 (3)	64.7 (11)	17.6 (3)	41.2 (7)
5 th	Pre	58.1 (25)	37.2 (16)	51.2 (22)	27.9 (12)	39.5 (17)	53.5 (23)	32.6 (14)	37.2 (16)
	Post	14.0 (6)	81.4 (35)	16.3 (7)	53.5 (23)	7.0 (3)	90.7 (39)	4.7 (2)	55.9 (24)

Inferential statistics

Table 3 displays the values of the coefficients γ_{01} , γ_{10} , and the related significance tests for each of the answers hypothesized to change. The results show that grade is not a significant predictor of the answers in all but two items. This means that the value of η (and, consequently, the probability ϕ that a participant gives a particular answer) is on average the same in all grades. The one exception concerns the answers to the question ‘Does Riri resemble the other kittens?’ In this case, there are significant grade effects. The directions of these effects are not important; what matters are their values compared to the intervention effects. Chi-

square values were 1.19 and 3.14, which were not statistically significant ($df=1$, $p>0.05$), suggesting that, on average, the intervention and grade effects are equal.

Table III. Effects of the grade (γ_{01}) and the intervention (γ_{10}) on the dependent variable η for each of the items and related significance

Question / related answers	γ_{01}	t	p level	γ_{10}	t	p level
<i>How do the kittens feel?</i>						
Vague emotion	-0.41	-2.0	n.s.	-0.41	-1.8	n.s.
Complex emotion	0.19	1.0	n.s.	0.41	1.0	n.s.
<i>How does Riri feel?</i>						
Vague emotion	-0.20	-0.9	n.s.	-0.80	-3.4	0.001
Complex emotion	0.15	0.5	n.s.	1.44	3.2	0.01
<i>Does Riri resemble the other kittens?</i>						
Yes	0.59	3.2	0.05	1.15	4.6	0.001
No	-0.62	-2.5	0.05	-0.99	-4.0	0.001
<i>Do you think Petros is doing the right thing by avoiding Lina? Why?</i>						
Yes	0.20	1.1	n.s.	-1.08	-4.3	0.001
No	-0.16	-0.9	n.s.	1.03	4.4	0.001
<i>Why?</i>						
Fear	0.02	0.1	n.s.	-1.29	-5.3	0.001
Understanding	0.21	1.4	n.s.	1.44	5.8	0.001
<i>Do you think Petros is doing the right thing by avoiding all dogs? Why?</i>						
Yes	0.18	1.3	n.s.	-1.33	-4.5	0.001
No	-0.19	-1.2	n.s.	1.15	4.4	0.001
<i>Why?</i>						
Fear	0.04	0.2	n.s.	-1.41	-4.8	0.001
Understanding	0.06	0.5	n.s.	0.80	3.6	0.001

The results also show that the intervention is a significant predictor of the answers in all but three items. The three items are the answer 'no' to the question: 'Do you think that Petros is doing the right thing

by wanting to retaliate against all dogs?’ and the answers ‘vague’ and ‘complex’ to the question ‘How do the kittens feel?’ In order to interpret the statistically significant intervention effects, it is necessary to transform the estimated η values back into the ‘success’ probabilities ϕ for each item. This is feasible with the following equation: $\phi = 1 / (1 + \exp(-\eta))$ (6). The values of η are equal to γ_{00} (pre intervention) and $\gamma_{00} + \gamma_{10}$ (post intervention). The direction of the changes in ϕ agreed with the hypothesized changes. For instance, in ‘How do the kittens feel?’ the probability of responding in the category of *vague emotion* decreased from 66.8% to 57.2% and the probability of responding in the category of *complex emotion* increased from 6.0% to 8.7%. Similarly, in the question ‘How does Riri feel?’ the probability of answering within the category of *vague emotion* decreased from 73.3% to 55.2% and the probability answering with a *complex emotion* response increased from 3.4% to 12.9%. In other words, in accordance with our hypothesis, the inadequate or insufficient description of affective states decreased and the identification of one or more affective states increased after the intervention.

In the question: ‘Does Riri resemble the other kittens?’ the probability of an affirmative answer increased from 20.7% to 45.3% and the probability of a negative answer decreased from 64.1% to 39.9%. It was observed, however, that notwithstanding the significant changes in these probabilities, the grade effect is also significant (and equal in size to the intervention effect). In particular, the probabilities of an affirmative answer are 32.1%, 46.0%, 60.6%, 73.5%, and 83.3% for the 1st, 2nd, 3rd, 4th, and 5th grade, respectively, and the probabilities of a negative answer are 49.0%, 34.1%, 21.8%, 13.0%, and 7.4% for the five grades, respectively. In other words, in accordance with our hypotheses, the identification of similarities increased after the intervention. This increase, however, needs to be approached with caution because identification of similarities also increased with grade.

In the question: ‘Do you think that Petros is doing the right thing by avoiding Lina?’ the probability of an affirmative answer decreased from 40.4% to 18.7%, and the probability of a negative answer increased from 50.5% to 74.1%. Regarding the justification of this answer, the probability of children responding within the category *fear* decreased from 47.8% to 20.1%, and the probability of answering within the category of *understanding* increased from 17.4% to 47.0%. Similarly, in the question: ‘Do you think that Petros is doing the right thing by avoiding all dogs?’ the probability of an affirmative answer decreased from 31.0% to 10.6% and the probability of a negative answer increased from 63.4% to 84.6%. In justifying their answer to this question, the probability of the children responding with the *fear* category decreased from 32.5% to 10.5% and the probability responding within the category of *understanding* increased from 31.0% to 50.0%. In accordance with our hypotheses, the avoidance and fearful behaviors decreased and the approach behaviors and empathic understanding increased.

In this kind of analysis, reliability describes the ratio of the true variance relative to the observed variance. According to Raudenbush and Bryk (2002), high reliability shows that there is large variability of the mean probabilities at level 2 or that the sample size is large. In most models tested here, reliabilities were 0.320 (true variance represents 32.0% of the total variance) or higher. The only exception is the reliability of the model testing the changes in the probabilities of the answer *complex* (to the question ‘How do the kittens

feel?"), which is relatively low (0.168). Although HLM corrects for low reliabilities, the results concerning this model should be approached with caution.

Discussion

Research on emotions and intercultural education in Cyprus (Zembylas, 2012) has highlighted the emotional ambivalence that children experience and the importance of children's 'emotional constructions' in how they perceive and, consequently, respond to different cultures. According to our analyses the participants' ability to recognize and identify emotions changed significantly in the post-intervention assessment phase. The ability to recognize the emotional consequences of behaviors form the basis for moral motivation in general and pro-social behaviors (Eisenberg & Miller, 1987; Hoffman, 1983). Likewise, these prerequisite social and emotional capabilities are considered fundamental in sensitizing children to difference, social justice, anti-violence and antiracism behaviours (Harris et al., 2007; Nikels et al., 2007). Qualitative evaluations conducted by the teachers indicated that holding the emotions activities at the start of the program allowed them to weave emotional expression into all the modules and classroom activities.

Significant changes in children's understanding of similarities and differences were evident in post-intervention assessments. Perhaps the role-plays, simulations and imaginary interactions with same-age cohorts from different cultures, religions, races, and abilities, as well as person-to-person contact in some of the activities, helped in reducing fear or 'stereotypic threat' and in increasing the projection of positive traits and affect (Crisp & Turner, 2009). The development of an empathic stance toward others who are initially considered to have stereotypical negative traits, has been documented in research using imaginative play to intervene and to ameliorate such negative attitudes and bias tendencies (Farr Darling, 2002). Hence, simulations, experiential role-play and critical reflection that aim to bolster empathic understandings and perspective taking, may be effective techniques to use in antiracial pedagogy endeavors. In addition, it is evident that positive interpersonal contact that is cooperative, rewarding, mutually beneficially and sanctioned by school authorities and parents, is also an effective method in reducing prejudice and discrimination (Berryman-Fink, 2006). Teachers' qualitative evaluations highlighted the effectiveness of the aforementioned techniques, and further elaborated how the classroom climate changed dramatically as students became more mindful of each other and of their differences.

The findings provide evidence that the students increased their ability to understand and articulate the function of generalization in prejudicial thinking. It appears that the experiential activities had a positive impact in that the students developed skills in critically appraising the socio-emotional mechanisms inherent in racism. As Thompson (1997) has emphasized, when such issues are brought into the classroom through art and performance, children are moved to reframe their understandings and deepen their experiences of how such difficult social-emotional issues are played out within relationships and contexts.

In summary, these findings contribute to the literature that explores the psychological and social impact of antiracism programming, a literature that is voluminous on theories and policies, but sparse when it comes to research on the impact of programming on school children (Walter 1999). It appears that the group discussions, dialogue and problem solving on perspective taking, naming emotions and identifying bias in

simulations and role plays, enhance children's understandings of complex social-emotional situations in what Farr Darling (2002) describes as constructing solutions in morally sensitive ways. The classroom encounters in understanding similarities and differences, group engagement in art and imaginative praxis, identifying and naming emotions, and perspective taking, aided children to come into contact in an empathic way with different others and to develop understandings on stereotypes, bias, and discrimination.

There are, however, several limitations in this study. The main outcome variables were the individual gains made by the program participants and there were no control group comparisons. The interpersonal skills and behaviors of the children were not assessed and this impedes our ability to fully evaluate the impact of the intervention. A caveat mentioned by several teachers in our study pertains to the tendency of some students to provide the 'socially desirable' answer during exercises. These comments suggest that teachers should be alert when they see a 'facile' change in students' behavior as it may reflect a skin-deep compliance to the demands of the situation. In addition, although the inferential analysis showed that at pre-intervention, grade effects were not apparent, there was no control group comparison in this study and this also diminishes the power of the findings. Similarly, the small number of participants, together with the fact that they were from 10 schools and different grades, limited the robustness of the findings, as well as the understandings as to what was happening at the classroom level.

Further investigation of how programming that sensitizes children to the underlying mechanisms of racism may be carried out in seeking to challenge and reduce prejudice, bias and discrimination. The evaluation of such programming through the use of both qualitative and quantitative methodology is also essential if teachers and other professionals working in schools are to gain knowledge of the theoretical underpinnings of antiracism programs, the intervention methods that can be applied, and the benefits that can be obtained by students, classrooms, schools and communities at large.

References

- Arnold, C. L. (1992). An introduction to hierarchical linear models. *Measurement and Evaluation in Counseling and Development*, 25 (2), 58-90.
- Aveling, N. (2007). Anti-racism in schools: a question of leadership? *Discourse: Studies in the Cultural Politics of Education*, 28(1), 69-85.
- Banks, J. A. (2006). Improving race relations in schools: from theory and research to practice. *Journal of Social Issues*, 62 (3), 607-614.
- Bartolo, P.A. (2010). Why school psychology for diversity? *School Psychology International*, 31(6), 631-643.
- Bartolo, P.A., Borg, M., Cefai, C., & Martinelli, V. (2010). School psychology for diversity: Editorial. *School Psychology International*, 31(6), 563-566.
- Benson, J., & Poliner, R. A. (1997). *Dialogue: Turning Controversy into Community*. Cambridge, MA: Educators for Social Responsibility.
- Berryman-Fink, C. (2006). Reducing prejudice on campus: the role of intergroup contact in diversity education. *College Student Journal*, 40(3), 511-516.

- Boyd, D., & Arnold, M. L. (2000). Teachers' beliefs, anti-racism and moral education: problems of intersection. *Journal of Moral Education, 29*(1), 23-45.
- Connolly, P., Fitzpatrick, S., Gallager, T., & Harris, P. (2006). Addressing diversity and inclusion in the early years in conflict-affected societies: a case study of the Media Initiative for Children- Northern Ireland. *International Journal of Early Years Education, 14*(3), 263-278.
- Crisp, R. J., & Turner, R. N. (2009). Can imagined interaction produce positive perceptions? Reducing prejudice through simulated social contact. *American Psychologist, 64*(4), 231-240.
- Daniels, B (2010). Developing inclusive policy and practice in diverse contexts: a South African experience. *School Psychology International, 31*(6), 631-643.
- Derman-Sparks, L., & Brunson-Phillips, C. (1997). *Teaching/learning anti-racism*. New York, NY: Teachers College Press.
- Eisenberg, N., & Miller, P. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*(1), 91-119.
- Eisner, M., & Malti, T. (2012). The future of research on evidenced-based developmental violence prevention in Europe. *International Journal of Conflict and Violence, 6*(2), 166-175.
- Farr Darling, L. (2002). Moles, porcupines, and children's moral reasoning: unexpected responses. *Early Years, 22*(2), 91-103.
- Forlin, C. (2010). The role of the school psychologist in inclusive education for ensuring quality learning for all learners. *School Psychology International, 31*(6), 617-630.
- Freire, P. (2005). *Education for Critical Consciousness*. New York: Continuum International Publishing Group.
- Garrod, A., Beal, C. B., Jaeger, W., Thomas, J., Davis, J., Leiser, N., & Hodzic, A. (2003). Culture, ethnic conflict and moral orientations in Bosnian children. *Journal of Moral Education, 32*(2), 131-150.
- Gasser, L., Malti, T., & Gutzwiller-Helfenfinger, E. (2012). Aggressive and nonaggressive children's moral judgments and moral emotion attributes in situations involving retaliation and unprovoked aggression. *Journal of Genetic Psychology, 173*(4), 417-439.
- Grant, C. A., & Sleeter, C. E. (1998). *Turning on learning: Five approaches for multicultural teaching plans for race, class, gender, and disability* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Harris, E., McFarland, J., Siebold, W., Aguilar, R., & Sarmiento, A. (2007). Universal prevention program outcomes. *Journal of School Violence, 6*(2), 75-91.
- Heron, J. (1996). Helping whole people learn. In D. Boud & N. Miller (Eds.), *Working with experience: animating learning* (pp. 75 - 91). London: Routledge.
- Hoffman, M. (1983). Affective and cognitive processes in moral internalization. In E. Higgins, D. Ruble & W. Hartup (Eds.), *Social cognition and social development: A sociocultural perspective* (pp. 236-274). New York: Cambridge University Press.
- Kim, M. K. (2012). Theoretically grounded guidelines for assessing learning progress: cognitive changes in ill-structured complex problem-solving contexts. *Education Technology Research and Development, 60*(4), 601-622.

- Koroni, M., Garagouni-Areou, F., Roussi-Vergou, C. J., Zafiropoulou, M., & Piperakis, S. M. (2008). The stigmatization of obesity in children: a survey in Greek elementary schools. *Appetite*, 52 (1), 241-244.
- Levy Paluck, E. (2006). Diversity training and intergroup contact: a call to action research. *Journal of Social Issues*, 62(3), 577-595.
- Malti, T., Killen, M., & Gasser, L. (2012). Social judgments and emotion attributions about exclusion in Switzerland. *Child Development*, 83(2), 697-711.
- Mattheou, D., Roussakis, Y., & Theocharis, D. (2006). Who carries the national flag? The politics of cultural identity in the increasingly multicultural Greek school. *European Education*, 37(4), 50-60.
- McCoach, B. (2010). Dealing with dependence (Part II): A gentle introduction to hierarchical linear modeling. *Gifted Child Quarterly*, 54 (3), 252-256.
- Motti-Stefanidi, F., Pavlopoulos, V., Obradovic, J., Dalla, M., Takis, N., Papathanassiou, A., & Masten, A. S. (2008). Immigration as a risk factor for adolescent adaptation in Greek urban schools. *European journal of developmental psychology*, 5(2), 233-261.
- Moutsios, A. (2003). The organizational nonreform of the Greek school. *European education*, 35(3), 60-72.
- Nezlek, J. B. & Zyzanski, L. E. (1998). Using hierarchical linear modeling to analyze grouped data. *Group Dynamics: Theory, Research, and Practice*, 2 (4), 313-320.
- Nikels, H. J., Mims, G. A., & Mims, M. J. (2007). Allies against hate: a school-based diversity sensitization training experience. *The Journal for Specialists in Group Work*, 32(2), 126-138.
- Pampel, F. C. (2000). *Logistic Regression: A Primer*. Thousand Oaks: SAGE.
- Postle, D. (1993). Putting the heart back in learning. In D. Boud, R. Cohen & D. Walker (Eds.), *Using Experience for Learning* (pp. 33-45). Buckingham: The Society for Research into Higher Education & Open University Press.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: applications and data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Sefa Dei, G. J. (2008). Schooling as community: race, schooling, and the education of African youth. *Journal of Black studies*, 38(3), 346-366.
- Selman, R. L. (2003). *The Promotion of Social Awareness. Powerful lessons from the partnership of developmental theory and classroom practice*. NY: Russel Sage Foundation.
- Thompson, A. (1997). For: anti-racism education. *Curriculum Inquiry*, 27(1), 7-44.
- Triliva, S. Anagnostopoulou, T., Hatzinikolaou, S., Chimienti, G., & Mastorakou, A. (2009). The development and evaluation of a program to sensitize Greek Grammar School students to issues relating to disability. *European Journal of Counselling Psychology*, 1(1), 3-17.
- Triliva, S., Anagnostopoulou, T., & Hatzinikolaou, S. (2008). *Ούτε καλύτερος, ούτε χειρότερος αλλά διαφορετικός [Neither better, nor worse, just different]*. Athens: Gutenberg.
- Walter, S. (1999). *Reducing prejudice and stereotyping in schools: multicultural education series*. Willston, VT: Teachers College Press.

- Woolgar, M., Steele, A., Steele, M., Yabsley, S., & Fonagy, P. (2001). Children's play narrative responses to hypothetical dilemmas and their awareness of moral emotions. *British Journal of Developmental Psychology, 19*(1), 115-128.
- Zembylas, M. (2012). The politics of fear and empathy: emotional ambivalence in 'host' children and youth discourses about migrants in Cyprus. *Intercultural Education 23*(3), 195-208.
- Zoniou-Sideri, A., & Vlachou, A. (2006). Greek teachers' belief systems about disability and inclusive education. *International Journal of Inclusive Education, 10* (4-5), 379-394.