The Effect of a School-Based Creative Expression Program on Immigrant and Refugee Children's Mental Health and Classroom Social Relationships:

A Cluster Randomized Trial in Elementary School

Ruth Kevers[1], Caroline Spaas[1], Ilse Derluyn[2], Sofie de Smet[1,3], Wim Van Den Noortgate[4], Hilde Colpin[5] and Lucia De Haene[1]

[1]Parenting and Special Education Research Unit, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium; [2]Centre for the Social Study of Migration and Refugees, Department of Social Work and Social Pedagogy, University of Ghent, Ghent, Belgium; [3]Studies in Performing Arts and Media, Faculty of Arts and Philosophy, University of Ghent, Ghent, Belgium; [4]Methodology of Educational Sciences Research Group & imec-ITEC, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium; [5]School Psychology and Development in Context Research Unit, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

Author Note

Correspondence concerning this article should be addressed to Caroline Spaas, Parenting and Special Education Research Unit, Faculty of Psychology and Educational Sciences, KU Leuven, Leopold Vanderkelenstraat 32 (box 3765), B-3000 Leuven, caroline.spaas@kuleuven.be

The Effect of a School-Based Creative Expression Program on Immigrant and Refugee Children's Mental Health and Classroom Social Relationships:

A Cluster Randomized Trial in Elementary School

Abstract

Objective This study evaluated the effects of a school-based creative expression program on mental health and classroom social relationships in elementary school children with a refugee and non-refugee migration background. It was hypothesized that children receiving the intervention would report less externalizing and internalizing problem behaviors, less post-traumatic functioning, and more positive classroom social relationships at post-test than children receiving education as usual, particularly for refugee children.

Methods Classes in three multiethnic Belgian elementary schools were randomly assigned to a creative intervention (7 classes, 68 students) or control condition (6 classes, 52 students). All participants (8-12 years old) had a migration background. Almost half (47%) were refugees, 53% were first to third generation non-refugee immigrants. Data collection included pre- and post-test assessment with children, parents and teachers. Multilevel analysis was used to assess outcomes.

Results Children in the intervention condition rated the classroom climate at post-test more positive than their control peers ($d = .33$). Children who received the intervention did not show less symptomatic functioning than children in the education-as-usual condition. However, post-hoc analysis by baseline severity showed that students with high baseline levels of post-traumatic stress reported less trauma symptoms at post-test in the intervention group than in the control group ($d = -.97$). This effect was moderated by children's refugee background,

indicating a differential effect in which refugee children show more reduction of trauma symptoms as compared to non-refugee immigrant children.

Conclusion The intervention supported classroom climate and alleviating post-traumatic stress in children with increased posttraumatic symptomatology.

Key Words

Refugee children; immigrant children; mental health; school-based intervention; creative expression workshops; classroom social relationships

Public Policy Relevance Statement

The school and the relationships that develop within the school context are central to refugee and immigrant children's mental health. This study indicates that a school-based creative expression program can support classroom climate and alleviating traumatic symptoms in refugee and immigrant elementary school children. Policy makers should consider incorporating a focus on fostering classroom safety and positive peer interactions as central points of concern in their efforts to increase refugee and immigrant students' well-being.

Introduction

With western multi-ethnic societies facing challenges of minorities' social integration, child mental health in immigrant communities is a scholarly topic with growing public health implications. As social and structural dynamics of discrimination, exclusion and polarization are consistently documented to negatively impact immigrant children's mental health (Priest et al., 2013; Rousseau et al., 2013; Spaas et al., 2021), scholarly literature shows a growing consensus that promoting well-being in immigrant children is inherently intertwined with aims of supporting cultural integration, social participation, and positive minority-majority group relationships (Measham et al., 2014; Özbek et al., 2015). A growing body of studies document the at-risk mental health functioning of immigrant children, in particular showing increased prevalence rates of internalizing problems (Derluyn et al., 2008; Dimitrova et al., 2016; Jackson et al., 2012; Kouider et al., 2014; Van Oort et al., 2007).

Within the broad group of immigrant children, refugee children are a particular group facing long-term cumulation of pervasive stressors. Although non-refugee and refugee immigrant children share stressors of resettlement, acculturative identity formation, minority-related discrimination, and potential exposure to community violence (Jaycox et al., 2002), refugee children's life-histories of collective violence and their coping with trauma-related sequelae in a context of exile and cultural change yield sensitivity to their specific mental health and service needs (Betancourt et al., 2017; Buchmüller et al., 2018). Indeed, refugee children's life-trajectories of forced displacement are characterized by traumatic pre-migratory stressors of war, atrocities, deprivation, loss or imprisonment of family members, and disruption of school trajectories, further compounded by cumulative stressors during flight, including separation from family members, threatening life conditions, and deprivation (Bolland, 2010; Derluyn et al., 2014). Once in the host country, refugee children face multiple stressors of cultural adaptation, residence insecurity, social isolation or experiences of discrimination

(Beiser & Hou, 2016; Hynie, 2018; Montgomery, 2011; Vervliet et al., 2014), often further complicated by the disruption of these children's family context in which cumulative distress may invoke family conflict and decreased parental availability in exile (De Haene et al., 2010a; 2013). Mental health studies indicate that this cumulation of disruptive life-experiences places refugee children at increased risk for mental health difficulties, with high prevalence of post-traumatic stress disorder (PTSD), depression and anxiety, and a range of broader internalizing and externalizing problems, such as psychosomatic complaints, re-enacting play, bereavement reactions, aggressive behavior, and concentration difficulties (e.g., Betancourt et al., 2012; Pacione et al., 2013; Yaylaci, 2018).

The Role of School-Based Mental Health Care

To promote refugee and immigrant children's mental health, scholarly work emphasizes the preventive and curative role of schools (de Wal Pastoor, 2015; Murray et al., 2010; Rousseau & Guzder, 2008). First, schools can operate as access point to or locus of mental health services, as refugee and non-refugee immigrant families generally underutilize regular services, although they face high levels of adversity and are at risk for developing psychosocial vulnerability (Bodvin et al., 2018; Ellis et al., 2010; Verhulp et al., 2017). Schools, as low-threshold spaces that are not imbued with mental health-related stigma, can operate as spaces of psychosocial prevention and support, by implementing effective mental health intervention (Beehler et al., 2012; Franco, 2018; Nadeau et al., 2014; Sullivan & Simonson, 2016). Second, school-based psychosocial interventions that engage with refugee and immigrant children's families may lower thresholds in family-school interaction and promote parents' involvement in children's school trajectories (Rousseau & Guzder, 2008). Third, school-based mental health practice can play an important role in promoting positive intercultural relationships and counteracting social dynamics of stigma and discrimination in multi-ethnic classrooms (Fazel,

2015). Studies have extensively demonstrated the primary importance of school-related variables in determining the long-term mental health of refugee and immigrant children and adolescents (e.g., Montgomery, 2011; Noam et al., 2014). Being an important part of the social context in diaspora, the school and the relationships that develop within this institutional context are central aspects for immigrant children's and their families adaptation and integration within the host community (De Haene et al., 2020). For the particular group of refugee children, studies document how peer relationships play a dual, protective or exacerbating role in post-trauma vulnerability. On the one hand, disrupted peer relationships and discrimination are found to predict long-term mental health problems in refugee children and adolescents, including persisting posttraumatic functioning (Keles et al., 2016; Montgomery, 2011; Montgomery & Foldspang, 2008; Vervliet et al., 2014). On the other hand, positive peer relationships have been documented to play a protective role, buffering against the development of mental health problems and mitigating the impact of diminished parental availability and broader social dynamics of stereotyping and discrimination (e.g., Fazel et al., 2012; Kia-Keating & Ellis, 2007; Montgomery, 2011). This body of research on the role of peer relationships in refugee and immigrant children's mental health indicates how schools can provide a vital forum for promoting belonging and social integration, counteracting broader social dynamics of discrimination, exclusion, and tension within minority-majority relationships (de Wal Pastoor, 2015; Kia-Keating et al., 2007; Walsh et al., 2010).

Initial Evidence-Base on School-Based Interventions for Refugee and Immigrant Children

Reflecting this increasing interest in school-based mental health interventions for refugee and immigrant children, a seminal body of intervention studies generates an initial evidence-base on different school-based intervention modalities. For refugee children, existing

school-based interventions are diverse, differing in scopes (e.g. prevention versus curation), targets (e.g., individual, classroom, or school level; focus on student, parent, or teacher), and treatment modalities (e.g., creative expression, verbal processing) (Bennouna et al., 2019; Tyrer & Fazel, 2014). School-based cognitive behavioral interventions (e.g., Ehntholt et al., 2005) generally show positive outcomes; multitier interventions (e.g., Ellis et al., 2013), which include several levels of service (ranging from preventive strategies to specialized mental health treatment), reveal mixed results. For the broader group of immigrant and refugee children, existing school-based interventions demonstrate promising results in the improvement of individual mental health through a range of treatment modalities, including mindfulness (e.g., Fung et al., 2019), cognitive behavioral therapy interventions (e.g., Barrett et al., 2003; Kataoka et al., 2003), comprehensive service models (e.g., Beehler et al., 2012), and creative expression programs (e.g., Rousseau et al., 2004). Rousseau and colleagues developed a series of creative expression interventions for refugee and immigrant children in different developmental phases (Lacroix et al., 2007; Rousseau et al., 2004; Rousseau et al., 2007). These semi-protocolled programs use creative processes designed to provide children with tools to express feelings and process emotions regarding migration and exile, and have been explored in different studies (e.g., Rousseau et al., 2005). Here, the existing evidence-base points to a reduction in post-traumatic, depressive, and anxiety symptoms associated with the implementation of creative arts-based programs, but more outcome studies are needed for more robust evidence (Tyrer & Fazel, 2014; Sullivan & Simonson, 2016).

Overall, research on the effectiveness of school-based interventions for refugee and immigrant children remains relatively scarce, with several methodological and analytic avenues that need to be addressed in strengthening this body of scholarly work. First, evaluation studies have mainly focused on individual symptomatic functioning in refugee and immigrant children, disregarding the impact of school-based mental health interventions on the quality of inter-

ethnic relations (Tyrer & Fazel, 2014). Especially the gap in studies exploring the impact of school-based intervention on inter-ethnic peer and school relationships is remarkable, given the emphasis on the role of school-based programs in supporting positive school well-being in refugee and immigrant children. Further, the existing evidence-base mainly consists of relatively small sample sizes ($N < 100$; Tyrer & Fazel, 2014), invoking the necessity of larger controlled studies. Lastly, a multi-informant approach is recommended to complement the use of self-report questionnaires in child mental health assessment, including family members, peers, and school staff in the assessment of a broader range of outcome variables, including classroom social relationships.

Hypotheses, Aims, and Objectives This project aims at furthering the evidence-base on the role of school-based intervention in promoting refugee and immigrant children's mental health, school well-being, and peer relationships. It entails an intervention study in which a creative arts-based expression program aiming at promoting positive mental health outcomes (Rousseau et al., 2004) is implemented with refugee and non-refugee immigrant children in multi-ethnic classes in elementary schools. The study assesses the intervention's effect on (i) mental health outcome variables and (ii) classroom social relationships.

First, we hypothesize the intervention is effective in improving mental health outcomes in refugee and immigrant children. As a Canadian study reported a decrease in internalizing and externalizing symptoms in refugee and non-refugee immigrant children in the intervention condition (Rousseau et al., 2005), we hypothesize children in the intervention group will report less internalizing and externalizing problem behavior and less post-traumatic symptomatology than those in the control group; and we expect this will particularly be the case for the refugee children in this group given the long-term cumulation of pervasive stressors impacting their development. Second, we hypothesize the intervention is effective in promoting positive

classroom social relationships as outcome variable. Previous research has not yet studied the possible effects of this creative arts-based program on school well-being and relationships. Given that the construction of a safe space and the acknowledgement and appreciation of diversity are two of the assumed working mechanisms of the program under study (Rousseau et al., 2004), it is plausible that the intervention might positively impact classroom safety and the quality of peer relationships.

Within the existing body of intervention studies on school-based mental health support for refugee and immigrant children, the inclusion of outcomes related to school well-being and social relationships as well as the multi-informant approach (children, parents, and teachers) is a unique contribution to the developing scholarly field.

Method

Participants

The target population for the intervention consisted of elementary school children with a refugee and non-refugee migration background. In the Belgian context, reception education for newcomer children between age 6-12 is embedded within regular elementary schools, where newcomer children are dispersed across classes and participate in the regular academic curriculum with native-born children. Given the irregular distribution of newcomer children across schools, sample size was maximized through purposive recruitment of three schools with a high number of newcomer children. Based on information provided by the regional Agency of Educational Services (AGODI), 30 schools with the highest numbers of newcomers were contacted by the research team at the start of school year 2017-2018 in order to obtain more information about the migration background of the student population and to inform them about the aims and contents of the research project. The three schools eventually taking part in the study were selected out of the eight schools that considered participation, based on the estimated

amount of refugee and immigrant students, their interest in the project, expected drop-outs of students throughout the year (e.g., due to repatriation after a negative decision or relocation after a positive decision in asylum procedures), and their geographical location. All three schools were located in the same Flemish urban setting, which increased homogeneity in participants' environment, and facilitated data collection and intervention implementation because of good accessibility, elements deemed important given the complexity of conducting research with newcomer and refugee communities (de Smet et al., 2021).

Within each school, only classes of the third to sixth grade (children aged 8-12) were eligible for participation in view of the target group of the intervention under study and the skills needed to complete self-report questionnaires. At the start of the intervention study, class teachers' consent for participation was ensured by the school principal of each selected school, with thirteen classes included in the study (Two Grade 3 and two Grade 5 classes in the first school 1; two Grade 3, two Grade 4, two Grade 5 and two Grade 6 classes in School 2; and one class in School 3). In the first two schools, with two participating classes per grade level, classes within each grade were randomly assigned to either the experimental or control group. In the third school, where only one class participated in the study, this class was assigned at random to the intervention group. As such, the overall randomization resulted in both a school and grade stratified cluster randomized trial with class as unit of randomization (Campbell et al., 2012), and a total of 13 classes participating in our study (see Figure 1).

Recruitment of children and parents took place between January-February 2018. In all participating classes all children had a migration background. This multi-ethnic classroom composition reflects a broader social segregation of refugee and immigrant families that is characteristic of particular Belgian schools, especially in urban settings (e.g., Pulinx et al., 2017). Obtaining informed consent from parents was approached as an active, relational process through subsequent steps, in order to give parents the opportunity to exercise human agency,

regain autonomy and engage as true partners in the research process (Hugman et al., 2011; De Haene et al., 2010b). In an initial step, the research team was present during parent-teacher conferences in order to establish first contacts with parents and children, providing brief, informal information about the research project. During these informal contacts, parents were invited to more elaborate information sessions about the project that in a second step would be organized in the school, both during and outside school hours. In these information sessions that were supported by interpreters, the aims and design of the intervention study were outlined; formal informed consent forms (available in most parents' mother tongues) were clarified; and, if required, research participation was further discussed individually. In a third step, written informed consent and information on background characteristics (e.g., country of origin, parental education) were gathered from those parents agreeing to their child's participation in the study. Recruitment and consent were thus perceived as iterative processes, building on the assumption that making agreements on research participation, particularly with vulnerable, hard-to-reach populations, requires a negotiation process directed at the development of a shared understanding of what research involves (Block et al., 2012; Mackenzie et al., 2007). Informed consent was obtained for 120 children (response rate of 43.80%). Despite the limited consent rate, the obtained sample size in this study is relatively large compared to existing studies on school-based mental health interventions for refugee and immigrant children. The existing evidence-base mainly consists of smaller sample sizes ($N < 100$; Tyrer & Fazel, 2014), which are accepted in this research field due to uniqueness of the study population and the urgency of the public health issue under study (Evans & Ildstad, 2001).

The participating children's age ranged from 8 to 12 years, with an average of 9.58 ($SD = 1.22$). 53.3% of participants were boys, 46.7% were girls. All participants thus had a migration background, with variation regarding time in resettlement and differing migration generations: around one third of children arrived less than one year ago (30.8%); one third had

been in the country between one and five years (33.3%); slightly more than one tenth arrived between five and ten years ago (12.8%); and almost one in four of the participants was born in Belgium (23.1%), with a second- or third-generation migration background, as the child or grandchild of immigrants. Participants were categorized as 'refugee' or 'non-refugee' immigrant children based on parents' answers on the background characteristics questionnaire. Life-trajectories of forced displacement induced by persecution, war, or violence were a precondition to be categorized as refugees. The country of origin (e.g., Syria, Afghanistan) and legal status (e.g., refugee, asylum seeker, subsidiary protection) indicated on the background characteristics questionnaire further informed the categorization as refugees. In case of indications that an additional check would be relevant, a further oral exploration with parents was undertaken to distinguish refugee children from students with non-refugee migration backgrounds (e.g., labor migration). In our sample, 47% of participants were categorized as refugees, 53% as non-refugee immigrants. The participant sample was heterogeneous in terms of countries of origin, with participant children being born in 32 different countries. Refugee children were mainly born in Syria (27.3%), Somalia (20%), Afghanistan (16.4%), and Iraq (10.9%); for non-refugee immigrant children main countries of birth were Belgium (33.9%), Spain (12.9%), The Netherlands (9.7%) and Morocco (6.5%). It is important to note that of these non-refugee immigrant students, 88.7% were first- or second-generation immigrants, several of them having parents who were born in different countries than themselves (e.g., 21.8% of non-refugee students' parents were born in Morocco, some of them having resided in other European countries before settling in Belgium). The other 11.3% comprised third-generation immigrants. In terms of legal status, almost one third of participating children (29.9%) had obtained Belgian nationality. The other children had a refugee status (27.4%), subsidiary protection status (4.2%), were still in asylum procedure (0.9%), had no papers (4.3%), or had another legal status (e.g., parents with European work permit; other European

nationality) (33.3%). The majority of participants' mothers and fathers had received basic education (50.5% and 56.2% respectively); others had completed higher studies (19.8% and 22.54%); and a considerable group reported not having been to school (29.7% and 21.3%).

Measures

In testing our hypotheses, we assessed the impact of the intervention on both the mental health outcome variables (i.e., externalizing and internalizing behavioral problems, post-traumatic stress problems) and the classroom social relationships variable.

For the assessment of mental health outcome variables of externalizing, internalizing behavioral problems, and post-traumatic stress problems, a multi-informant approach was implemented by using the self-, parent-, and teacher-versions of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), for children supplemented by the post-traumatic stress problems scale of the Youth Self Report (YSR; Achenbach & Rescorla, 2001, 2007).

The *SDQ instruments* are widely used and well-validated measures of child psychopathology (Goodman et al., 2000), with robust psychometric properties and extensive application in culturally diverse study settings (including translations in more than 40 languages) (Bourdon et al., 2005; Goodman, 2001). In addition, SDQ scores have been shown to be sensitive to intervention effects in the Flemish school context (e.g., Vancraeyveldt et al., 2015). The 25 items have to be scored on a three-point Likert-scale with 0 = 'not true', 1 = 'somewhat true', and 2 = 'certainly true'. By summing scores on relevant items, following scales can be computed: five specific subscales (emotional symptoms, conduct problems, hyperactivity, peer problems, and prosocial behavior), each with five items; two broader subscales (externalizing and internalizing problems); and a total difficulties score, indicating the risk of mental health morbidity, that is the sum of the emotional symptoms, conduct problems, hyperactivity, peer problems subscales. An impact supplement assesses the respondent's perception of mental health problems, chronicity, distress, and social impairment,

providing additional information on impairment. The self-report version is intended for use with 11- to 16-year-olds, yet has been validated for use with younger children as well: the simplicity of the measurement should allow children aged 8 years and older to report on their self-perceptions in line with the SDQ statements (Curvis et al., 2014). Parents, teachers, and children completed the 25 items of the questionnaire; parents and teachers also filled out the impact supplement that enquires about symptoms in terms of chronicity, distress, social impairment, and burden for others. In our study, internal consistency (Cronbach's α) for the total difficulties scale for data wave one and two respectively was .71 and .76 for self-report; .80 and .70 for parent report; and .79 and .82 for teacher report. Except for the self-reported internalizing problems (α (T1) = .54; α (T2) = .59), all three informants' externalizing (self-report α (T1) = .66; α (T2) = .68); parent report α (T1) = .74; α (T2) = .76); teacher report α (T1) = .82; α (T2) = .87) and internalizing (parent report α (T1) = .70; α (T2) = .59; teacher report α (T1) =. 73; α (T2) = .72) scale scores showed acceptable to good internal consistency. Several of the smaller subscales of the SDQ (e.g. emotional symptoms, peer problems, conduct problems) had unacceptable internal consistency (threshold of Cronbach's α at both waves below .60; Taber, 2018) and therefore no further analyses were performed on these specific subscales. Internal consistency of parents' impairment scale was also insufficient (α (T1) = .37; α (T2) = .74) and thus excluded from further analysis; teachers' impairment scale showed satisfactory reliability (α (T1) = .67; α (T2) = .71).

The *Youth Self Report* (YSR) is part of the widely used Achenbach System of Empirically Based Assessment (ASEBA). The ASEBA forms have been translated in over 100 languages and are used in epidemiological surveys, clinical assessment, outcome evaluations, research, and other purposes in many countries around the world (Achenbach & Rescorla, 2007). The YSR comprises 112 closed questions assessing behavioral, emotional, and social problems as well as positive qualities and was developed for use with 11- to 18-year-olds, yet

reliability and validity of the YSR scales scores for younger children (ages 7-10) have been examined as well, results demonstrating that younger youths are able to provide reliable reports (Ebesutani et al., 2011). For the purpose of the present study, only the 14 items of the post-traumatic stress problems (PTSP) scale were included in the self-report questionnaire administered to children. These items do not investigate exposure to traumatic events, but assess post-traumatic stress symptoms. Each item is rated on a three-point Likert scale ranging from 0 (not at all) to 2 (very often), resulting in a scale score ranging from 0 to 28 and also enabling the level of post-traumatic stress problems to be categorized as 'normal', 'borderline', or 'clinical'. The YSR-14 item scale measuring post-traumatic stress problems has been shown to be diagnostically accurate in predicting pediatric PTSD diagnosis; and children appear to provide valid response in self-reporting their post-traumatic stress problems compared to caregiver reports (You et al., 2017). The sample's Cronbach's α for data wave one and two respectively was .73 and .78.

For measuring the quality of classroom social relationships as outcome variable, two subscales of the Dutch Class Climate Scale (DCCS; Donkers & Vermulst, 2014) were included in the self-report questionnaires. The Class Climate subscale assesses children's perceptions of classroom peer interactions by measuring negative events that can occur among classmates (e.g., bullying behavior). The Quality of Mutual Peer Relationships Subscale measures students' satisfaction with the amount of connectedness and cohesion among classroom peers. Questions are formulated in the first person and assess, for example, whether the child feels they belong to the group. Each of these subscales consists of eight items and is scored on a four-point scale ranging from 1 (almost never) to 4 (often), resulting in two subscale scores. Standardization data were gathered among a large sample of students enrolled in different schooling systems in The Netherlands, resulting in satisfactory norms for general elementary schools (COTAN, 2014). Until today, no standardization research with children of a minority

ethnic background is available yet. Evidence for the validity of the subscales was reported by

Donkers and Vermulst (2014) and by Breeman and colleagues (2015). Cronbach's α in our

sample for data wave one and two respectively was .84 and .85 for the Class Climate subscale

and .79 and .82 for the Quality of Mutual Peer Relationships subscale.

Intervention

The creative arts-based expression program evaluated in this study is a classroom-based,

creative arts program in which children are invited to work with stories of (forced) migration

in both verbal and non-verbal modalities. The design of the intervention was strongly based on

the manualized program that was developed and studied by Rousseau and colleagues (2004).

This program was designed to support immigrant and refugee students in multiethnic

elementary schools in Montreal bridging the gap between past and present, home and school.

Workshop sessions were implemented in each of the seven classes assigned to the intervention

condition during regular school hours by a qualified creative arts therapist in conjunction with

the class teacher. The creative arts therapists were recruited by the research team through an

organization for creative therapy external to the school, and had no previous relation to either

the school, its staff or the children involved in the intervention. While the original program

consists of twelve weekly sessions of two hours each, the program of this study only involved

eight (two-hour) sessions because the practical organization of the school year in the different

participant classes did not allow for a higher dose of implementation. Each class in the

intervention condition received an equal intervention dosage. As in the original program, the

intervention in this study was composed of three consecutive parts, combining storytelling and

drawing in both individual and collective work. All sessions had a fixed, predictable structure,

involving a warming-up and closure activity enveloping the main activity (see intervention

manual; Erit, 2010). The first component of the program (week 1-2) was organized around

myths and tales belonging to divergent cultures, aiming to represent non-dominant traditions

and engaging with themes of migration and disruptive or traumatic life-experiences. In these first sessions, the creative arts therapist (who administered the program in cooperation with the class teacher) recounted a tale that dealt with themes of migration and culture, encouraged the children to empathize with it, and then invited them to draw and share about it. In the second component (week 3-6), children were asked to construct a story of migration of a character of their choice, depicting pre-migration, migration, and post-migration phases. Children were invited to draw and write about this migration journey and then talk about it with classmates. Unlike the original program, the initial construction of the character's story was developed collectively and not individually. Furthermore, drawing and writing was supplemented by brief theatrical exercises, in which children staged the migration journey, exploring it in another modality of expression. Overall, the range of symbolic and mythic referents mobilized in these activities was assumed to allow refugee and non-refugee immigrant children to express and account for their experiences while maintaining a safe distance to difficult and potentially threatening aspects such as loss, separation, and trauma (Rousseau et al., 2003). The third component of the program (week 7-8) aimed to bridge the gap between home and school by inviting children to bring in tales from their family and community contexts. As in the original program, children were now invited to bring in their identity more explicitly, by asking their parents for traditional tales from their homeland. The original program's focus on family myths and tales was in this study's intervention supplemented with the invitation to ask parents which games they often played in their childhood, which were then played in the subsequent session.

Procedure

In January 2018, participant classes within each school and grade were randomly assigned to either the experimental or the control condition. As a result, 68 children were assigned to the intervention and 52 to the control condition (see Figure 1). Students, parents, and teachers were non-blinded to allocation.

Data collection took place at two time points: before the start of the intervention and after intervention completion. Following the pre-test assessment of the mental health and classroom social relationships through the administration of questionnaires with all participant children, parents, and teachers in March 2018, the 8-week creative arts-based expression program was implemented in the intervention classes (March-May 2018) while the control groups received education as usual. After intervention completion, post-test assessment of children, parents, and teachers was administered in June 2018. Teachers completed questionnaires independently, with every classroom teacher completing one questionnaire for each participating child in their class (see Figure 1); data collection with children and parents was organized collectively (i.e., children of one or several classes together; parents together), assisted by the research team. During this collective questionnaire administration, children and parents received additional support by means of visualization of questions (pictograms), interpretation by qualified interpreters, translated questionnaires (if available in the participants' mother language), and standardized clarifications by the research team. Although questionnaires were completed in the same room, participants worked independently and socially desirable responses were discouraged as much as possible by ensuring participants that their answers would be treated confidentially. In case parents were not available to attend collective questionnaire administrations, they could complete them at home. Considerable parental drop-out occurred between both measurement points (see Figure 1). Despite several and repeated attempts, we did not succeed in inviting these parents back to school a second time to complete post-test questionnaires. Some parents we simply could not reach, others indicated that they did lacked the time or opportunity to participate a second time. Parents' withdrawal from the study at post-test assessment might be indicative of known difficulties in researching this hard-to-reach population (e.g., Fête et al., 2019), and could potentially be related to experienced time constraints of parents negotiating a myriad of different responsibilities in the

host country, or to negative experiences with our research at pre-test assessment (e.g., parents feeling like they were being asked sensitive questions; parental perceptions of an insufficiently transparent research framework).

Although standardized monitoring was not included in this intervention study, implementation quality of the creative program was secured in several ways. The two creative arts therapists implementing the workshops were trained as licensed creative therapists; both participated in regular supervisory sessions provided by the research team. Standardized implementation was maximized by the therapists' shared rehearsing and making of minor adjustments to the intervention manual that prescribed the course of the sessions; registration that each of the 8 sessions took place was ensured for each participant-class. Teachers in the intervention group were informed about the aims and contents of the intervention before the start of implementation and frequent consultation between teachers and creative arts therapists took place throughout its course.

Prior to commencing the study, approval of the research design was obtained from the university's Ethics Committee (Social and Societal Ethics Commission, KU Leuven, application G-2017 08 884).

Data Analysis

Quantitative data-analysis started with descriptive analyses of participants' sociodemographic characteristics and mental health, including Chi-Square tests of independence and non-parametric Mann-Whitney tests to determine possible baseline differences between the intervention and control group. Non-parametric Mann-Whitney tests were opted for because Shapiro-Wilk's W-tests and Q-Q plots indicated that the assumption of normality was not met for most background variables. Subsequently, multilevel analysis was used to assess the effects of the intervention. Visual analyses of the residuals of the outcome

variables did not reveal evidence that the assumption of normality was violated. Multilevel modeling was needed to take the correlation between scores of children of the same class into account (Peugh, 2010). Model parameters were estimated using the maximum likelihood procedure in the SPSS statistical package version 25. To test the significance of fixed effects, we used Wald tests, whereas for testing variance components, we compared the difference in model deviance values (-2 times the log likelihood) of a model with and a model without the variance component in question to a Chi² distribution with 1 degree of freedom to decide whether the model including the variance component(s) being tested fitted the data significantly better.

In a first step, unconditional models (i.e., intercept only models) were estimated for all outcome variables to investigate the amount of variance of the post-test scores at the two levels (i.e., child and class level). Intra class correlations ranged from .01 to .19 and design effects ranged from 1.08 to 2.56. Hence, for some outcome variables, design effects exceeded the critical value of 2 (Muthén & Satorra, 1995); for others, it did not, but for reasons of consistency two-level modeling was used for all analyses. The school level was not modeled because the inclusion of only three schools in the study design made a reliable estimation and meaningful interpretation of variance components at this level impossible. In a second step, several predictors were gradually added to the unconditional models: a dummy-coded intervention variable (0 = control group, 1 = intervention group); a continuous, grand mean-centered baseline variable (in order to remove the variance explained by the baseline score from the error variance in the estimate of the difference in outcome between intervention and control group, improving power to detect treatment effects); and a dummy-coded refugee background variable (0 = non-refugee, 1 = refugee). Finally we fitted models including an interaction between baseline and treatment group and refugee background and treatment group. Including these cross-level interactions allowed us to examine whether the relationship between treatment and

outcome was moderated by the baseline score or refugee background. Because of the small number of classes ($N = 13$), the power of our significance tests is relatively low, even if effect sizes are not very small. Therefore, we do not only pay attention to significant findings (with α = .05), but also to strong non-significant tendencies ($p < .10$; Spilt et al., 2016), yet without drawing firm conclusions. To decide which model best fitted the data, the deviance values (-2 times the log likelihood) for two nested models were used (Peugh, 2010). Specifically, the difference in the deviance statistics between the models was referenced to a Chi-square distribution with degrees of freedom equal to the difference in the number of estimated parameters between the models (Field, 2013). The effect size of the intervention effect was computed by calculating the difference in expected scores between control and intervention group (i.e., the parameter estimate of the intervention regression coefficient) divided by the unexplained standard deviation (i.e., the square root of the sum of the unexplained between-person and between-class variance). Interpretation of these effect sizes is similar to the interpretation of Cohen's d (Cohen, 1992; Van Den Noortgate & Onghena, 2008): .20 is considered small, .50 is considered medium, and .80 is considered large.

Results

Descriptive Statistics

We analyzed whether control and intervention group statistically differed on the demographic characteristics (as reported by the parents), using Chi-Square tests of independence. The two groups did not statistically differ on any of the background characteristics (i.c., sex, age, country of birth, legal status, maternal and paternal education) except time in Belgium ($\chi^2(3) = 8.33$; $p = .040$): for example, in the intervention group, there were almost twice as many participant children who were less than one year in Belgium at pre-test (37.88%) than in the control group (21.57%).

Mean scores and standard deviations on the outcome variables at pre-test and post-test, for the total sample as well for each condition separately, are reported in Table 1. In the total sample, the average score of post-traumatic stress problems at pre-test was 10.17. Clinical classification of the scores indicates that 30% of participant children scored in the borderline or clinical PTSD-range (Achenbach & Rescorla, 2007). The average score of self-reported total difficulties (SDQ) at pre-test was 11.66. This corresponds to 25% of participant children reporting a high or very high amount of difficulties (Green et al., 2005; YouthinMind, 2016). Here, teachers made a less alarming assessment than children, only locating 5% of participants within this range of (very) high symptomatic functioning. Cross-informant comparisons indicate that an important group (22.8%) of children reporting high or very high difficulties was not detected by their teacher (Kevers et al., under revision). The mean score for Class Climate at pre-test was 2.76, which means that on average, participants experienced the atmosphere in the class as negative (mean score corresponding to norm score between 'low' to 'below average') (Donkers & Vermulst, 2014). Thus, participants on average reported high levels of mental health problems and negative perceptions of classroom safety (Kevers et al., under revision).

The only outcome variable for which there was a significant difference between the control and intervention group at pre-test was Class Climate: a Mann-Whitney test indicated that Class Climate was rated more positively in the control group ($M = 3.01$) than in the intervention group ($M = 2.56$), $U = 1041$, $p < .000$. In the multilevel modeling, we control for this baseline difference by including the baseline score as covariate.

Model Estimations

Before estimating multilevel models, missing data were explored (Baraldi & Enders, 2010). Over all twelve post-test outcome variables, 10.8% of the scores were missing[1]. Little's

---

[1] To calculate the proportion of missing data, we divided the number of missing scores (156) by the expected number of scores for the 12 outcome variables on the post-test measurement occasion (120 * 12 = 1440).

(1988) Missing Completely At Random (MCAR) test indicated that there was no convincing evidence for the missing data not being completely at random, $\chi^2 = 5.292$, $df = 55$, $p = 1.00$. Specific measures, such as multiple imputation, were therefore not required to handle missing data, as missing values in this case would not flaw parameter estimates when using the maximum likelihood procedure (Black, 2011; Enders, 2017).

To examine the effects of the intervention, we estimated multilevel models using the same step-by-step approach for each of the twelve outcome variables: (1) fitting unconditional models, (2) including the level 2 predictor 'intervention status', (3) including level 1 predictors 'baseline' and 'refugee background', and (4) including cross-level interactions. Including 'time in Belgium' (cf. statistical difference between intervention and control group regarding 'time in Belgium', ($\chi^2(3) = 8.33$; $p = .040$)) to the list of level 1 predictors did not fundamentally change the results of the multilevel analyses: the same predictors and interactions remained (un)significant. Furthermore, for none of the outcome variables, 'time in Belgium' had a significant effect on the outcome scores. Therefore this background variable was not included in the list of level 1 predictors.

The most complex multilevel model (the one including cross-level interactions) for each outcome variable is presented in Tables 2-4. Taking the χ^2_{change} statistic into account, this most complex model including the entire set of predictors was not always the best fit for the data, yet the parameter estimates of the individual regression coefficients were on the same line in full and more reduced models. Baseline scores significantly predicted post-test scores for all outcome variables. A main effect of the creative arts-based intervention was found for the classroom social relationships outcome variable (Class Climate). Although the positive observed effect in the control group as well as the additional positive effect for the intervention group are non-significant in the most complex model including cross-level interactions (see Table 4), the more reduced model with covariates intervention status, baseline, and refugee

background points to a main intervention effect approaching significance (see Table 5; Model with level 1 predictors): taking baseline differences into account, the intervention children, compared to the control children, rated the climate in the classroom as more positive at post-test ($\gamma = .21$; $p = .098$). The effect size of this intervention effect was .33, indicating a small effect.

No main effect was found on the mental health outcome variables. We explored whether refugee background had an effect on the outcome scores. On self-reported total difficulties (SDQ) and post-traumatic stress problems (PTSP) at post-test, refugee children appeared to report significantly less symptomatology than non-refugee immigrant peers. The refugee-treatment interaction was significant for none of the outcome variables, indicating that the effect of the intervention did not depend on whether the student had a refugee background. The differences in self-reported total difficulties (SDQ) and post-traumatic stress problems (PTSP) at post-test between participants with a refugee and non-refugee background were reaffirmed in an additional sensitivity analysis, whereby we added baseline self-reported total difficulties (SDQ) and post-traumatic stress problems (PTSP), gender and time in resettlement as covariates to our multi-level model.

For some outcome variables, a significant baseline-treatment interaction was found, which suggests there were significant differences between the control and intervention group depending on baseline severity (Nunes et al., 2011). For parent-reported externalizing difficulties, the model including cross-level interactions appeared to fit the data best, and a significant baseline-treatment interaction was found ($\gamma = 0.35$; $p = .019$; see Table 3). Figure 2 illustrates this baseline-treatment interaction. As shown in Figure 2, post-test externalizing scores increased as baseline scores increased, and this appeared to be stronger the case in the intervention group than in the control group (steeper slope for the intervention group). Furthermore, the interaction graph indicates that especially for children with rather high

externalizing baseline scores, parents in the intervention group reported higher average externalizing difficulties at post-test than parents in the control group. For the following outcome variables, analyses also gave evidence for a baseline-treatment interaction, although not always statistically significant: teacher-reported impact score ($\gamma = 0.26$; $p = .061$; see Table 4); post-traumatic stress problems ($\gamma = -0.29$; $p = .107$; see Table 4); and quality of mutual peer relationships score ($\gamma = 0.31$; $p = .033$; see Table 4). However, for all three outcome variables, the deviance statistic indicated that the global model including these cross-level interactions did not improve the fit. For illustrative purposes, baseline-treatment interaction plots are included in Figure 2-5 to facilitate the interpretation of these interactions.

In order to distinguish possible differential intervention effects according to baseline severity, post-hoc analyses were performed for the post-traumatic stress problems outcome variable, for which model estimations pointed towards a baseline-treatment interaction, although not statistically significant. For the post-traumatic stress problems scale, the ASEBA manual provides cut-off values to categorize scores as normal, borderline, or clinical (Achenbach & Rescorla, 2007). Based on these cut-off values, a PTSP 'baseline profile' dummy variable was constructed, distinguishing a 'normal' and a 'borderline/clinical' subgroup. A Chi-Square test of independence indicated that the control and intervention group did not statistically differ on baseline profile ($\chi^2(1) = 1.463$; $p = .227$). Subsequently, step-by-step multilevel models including level 2 (intervention status) and level 1 (refugee background) predictors were estimated for the normal and borderline/clinical subset of participants separately (Field, 2013). Tables 6 and 7 present the results of these post-hoc analyses. A significant intervention effect depending on baseline severity was found: for the normal PTSD baseline profile group (see Table 6), there were no significant differences between post-test scores of control and intervention participants, while for the borderline/clinical PTSD baseline profile group (see Table 7), the intervention did appear to have a significant positive effect ($\gamma =$

-3.86; $p = .038$): children with a borderline/clinical baseline profile reported significantly less post-traumatic stress problems at post-test assessment in the intervention group than in the control group. The effect size of this significant intervention effect was -.97, indicating a large effect. The clinical significance of this finding is as follows: in the intervention group, 11.1% of participants evolved from the borderline/clinical range of symptomatology at baseline towards the normal range of post-traumatic stress problems at post-test, versus 7.8% in the control group. Furthermore, in the subgroup of children reporting borderline or clinical levels of post-traumatic stress problems at baseline assessment, there was a significant difference between refugee and non-refugee participants at post-test ($\gamma = -4.63$; $p = .002$), with refugee children reporting significantly less difficulties than their non-refugee peers.

Discussion

Main Findings

In this study, a cluster randomized design was used to investigate the impact of a school-based creative expression intervention on refugee and non-refugee immigrant children's mental health and classroom social relationships in elementary school. In addition, this study aimed to investigate whether the effects of the intervention would be moderated by the migration background of participant children (refugee vs. non-refugee immigrant background). ~~Compared to previous studies, a relatively larger sample of students was recruited and not only children's, but also teachers' and parental assessments were administered.~~

The results of the study indicate that the intervention played a role in improving classroom climate as well as in alleviating post-traumatic symptomatology. First, regarding the impact on the classroom social relationships outcome variable, our study was the first to investigate the impact of this creative expression program on classroom social relationships. Although the intervention was not associated with an improvement in the quality of peer relationships, an effect of the intervention on children's perceptions of the climate in the class

was found: intervention children experienced classroom peer interactions as significantly more positive than children in the control group. Given that the construction of a safe space was one of the assumed effective factors of the program under study (Rousseau et al., 2004), this effect was in line with the hypothesis. The magnitude of this effect should however not be overestimated: the effect size was small *(d = .33)* and both in the intervention and the control group, the classroom climate at post-test was still rated 'below average' according to the Dutch norms for elementary school children (Donkers & Vermulst, 2014). Here, it is important to acknowledge that this unsafe perception of classroom peer interactions may result from a combination of risk factors that many students in multiethnic classes share: family contexts characterized by an accumulation of resettlement stressors (e.g., Jaycox et al., 2002) as well as increased prevalence of post-traumatic functioning may account for increased arousal in students. Difficult peer relationships can be the result (e.g., being bullied; for a review, see Zych et al., 2017) and are likely to be associated with more unsafe perceptions of classroom peer interactions. In reverse, experiences related to unsafe peer interactions (such as bullying victimization and discrimination) have been found to predict increased PTSD symptomatology in refugee adolescents (Kira et al., 2014). Overall, it seems of pivotal importance to make the fostering of classroom safety and positive peer interactions a central focus in school-based interventions aiming to improve young refugees' and immigrants' wellbeing.

Second, regarding mental health outcomes, our study generates some preliminary evidence of the role of the intervention in improving mental health outcome. In contrast to what was hypothesized, intervention children compared to control children did not show less internalizing and externalizing symptoms, neither in self, parent, nor teacher assessment. In the Canadian context, implementation of the arts-based expression program, on which the intervention studied here was strongly based, led to a significant decrease in internalizing and externalizing symptoms (Rousseau et al., 2005). In our study, the absence of an effect on

internalizing and externalizing difficulties might be partly explained by the relatively low implementation dose of the intervention (eight sessions compared to 12 in the original intervention). Yet, on post-traumatic stress problems, the intervention did seem to have a significant impact: for those children reporting elevated levels of post-traumatic symptoms at pre-test, the intervention was associated with less post-traumatic stress problems at post-test than education-as-usual (baseline-treatment interaction). This finding suggests that for those children with high levels of post-traumatic symptomatology, even a limited number of sessions may have played a role in creating a holding environment for the containment of traumatic distress, supporting a reduction of posttraumatic symptoms. This finding generates an interest into further exploring the mechanisms underlying symptom reduction in this group-based intervention program for children experiencing increased traumatic stress: even when the quality of peer interactions remained generally low, specific intervention processes within the classroom context seem to have supported an improved sense of safety and continuity. A further understanding of particular working mechanisms and whether they are embedded within adult-child caregiving relationships or classroom peer relationships to the same degree, may allow to identify central intervention characteristics and further our understanding of school-based mental health promotion in this vulnerable group.

For externalizing difficulties as reported by parents, the differential increase for the intervention group as compared to the control group, especially for children with high pre-test scores, warrants further reflection. As this finding is not mirrored by teacher and self-reports, it generates questions on how children cope with the transfer of themes and dynamics expressed and experienced within the classroom-based intervention to the home context and parent-child interaction. Given the sensitivity of topics related to migration, traumatization, and current social condition (including discrimination) within parent-child communication that is often marked by avoidance of these topics (De Haene et al., 2020; Rousseau et al., 2013), this finding

raises the question to what extent children with a more acting-out profile have difficulties in conveying their lived experience of intervention themes within family interaction, invoking problem behavior within parent-child or family interaction. Alternatively, this differential increase of externalizing difficulties as reported by parents in the intervention group as compared to the control group, could potentially be explained by the fact that parents of children who participated in the intervention and who potentially shared about the intervention at home (de Smet et al., in preparation), became more sensitive in recognizing and labeling their children's difficulties as behavioral difficulties in the context of cumulative stressors.

In contrast to what was hypothesized, no moderation effect of refugee background was found in the primary analyses: the intervention had no differential effects for refugee versus non-refugee immigrant children. However, post-hoc analysis indicated that refugee background did moderate the impact of the intervention on post-traumatic stress problems in those participants with a severe PTSD baseline profile: in this subgroup, refugee students reported significantly less post-traumatic problems at post-test than their non-refugee immigrant peers. In general, refugee children in both control and intervention classes reported significantly less post-traumatic and combined externalizing and internalizing difficulties than their non-refugee peers. Here, the non-refugee students in our sample – who in fact represent a very heterogeneous group in terms of cultural, socioeconomic, and migratory background – appeared to be particularly vulnerable, not only confirming the at-risk mental health status of immigrant children but equally indicating their potential exposure to traumatic stressors of community or domestic violence (Jaycox et al., 2002). From this perspective of trauma exposure in non-refugee immigrant children, an important question generated by this study's findings is to what extent the thematic foci of the creative expression program could sufficiently relate to their suffering, because of the creative intervention's focus on migration- and trauma-related

stressors rather than current social conditions that may be more pressing for immigrant children with a non-refugee background.

Limitations, strengths, and future directions

Research in school contexts with immigrant populations often seem to be confronted with an accumulation of practical obstacles. Indeed, several practical issues and obstacles complicated the recruitment and implementation phase of this intervention study, which gave rise to several limitations. First, notwithstanding the intensive recruitment phase of this intervention, a limited number of schools and parents decided to participate in this study. Here it is important to denote that language barriers, social differences, and dynamics of distrust in relationship to official institutions may have complicated the process of obtaining parents' informed consent. Besides, schools with the highest numbers of newcomers also seemed to characterized by a highly diverse and socially vulnerable student population. These school-specific characteristics may generate particular day-to-day challenges for school actors that may hamper the perceived feasibility to initiate and conduct research collaboration. Second, in setting up the implementation of the intervention with the schools, the intervention dosage and creative art-based expression program had to be aligned with each school planning, which gave rise to a smaller number of sessions in this study.

Notwithstanding these practical obstacles encountered in the research process, the intervention study foregrounds several important and innovatory strengths compared to the existing body of research on school-based mental health interventions for refugee youth. A first strength of this intervention study is the relatively large obtained sample size in comparison to previous studies on school-based mental health interventions for refugee populations ($N > 100$; Tyrer & Fazel, 2014). At the same time, the high number of refugee and non-refugee immigrant children in participating schools may limit the generalizability of our study's findings to school contexts with a similar multi-ethnic school population, though refugee and non-refugee

immigrant children's school trajectories may be quite frequently characterized by such multi-ethnic school composition in an urban setting (Pulinx et al., 2017). Second, responding to the increasing call of scholars to reinstall agency and autonomy at the side of participants in research procedures (Block et al., 2012; De Haene et al., 2010b; Mackenzie et al., 2007), an active iterative, collaborative process of obtaining and (re)negotiating informed consent characterizes the recruitment and data collection phases of this intervention study. In addition, the continuous collaboration with professional interpreters throughout this study in the data collection phase, which included the administration of questionnaires and focus groups with participating children and parents, aimed to facilitate such a process of a shared understanding and negotiation of research participation. Third, data collection involved the engagement of multiple informants by the administration of questionnaires, enabling the inclusion of the perspective of the child, parent and teacher in school-based mental health support for refugee and non-refugee immigrant children. Fourth, an important innovatory potential of this intervention study lies in the inclusion of outcomes related to school well-being and peer relationships in classrooms, which expands the dominant focus on the assessment of symptomatology within the existing body of studies of school-based interventions studies. Fifth, the implementation of the intervention involved a continuous process of monitoring and therapeutic reflection that was guided by multiple researchers. Indeed, regular supervisory meetings took place between the creative arts therapists and the research team in order to prepare, discuss and evaluate the enfoldment of the workshops before, during and after the implementation phase. This monitoring procedure could however be further strengthened in future studies by, for example, the inclusion of standardized checklists and feedback questionnaires for the involved creative arts therapists. Furthermore, future research could also comprise a more longitudinal focus to scrutinize the long-term impact of school-based mental health interventions for immigrant youth. However, the presence of a relatively high drop-out

rate of the student population in schools under study, potentially attributed to a relatively high rate of relocations for immigrant and refugee populations, may hamper a longitudinal follow-up in future studies. Last, as our findings foregrounded the particular vulnerability of non-refugee immigrant students, future research is urged to respond to the need to foster the development of adequate psychosocial care for immigrant children.

Conclusion

This cluster randomized intervention study on the effects of a school-based creative expression program on mental health and classroom social relationships in elementary school refugee and non-refugee immigrant children, provides evidence on the potential of school-based intervention to support classroom climate and alleviate post-traumatic stress in children with increased levels of posttraumatic symptomatology. Through its experimental design, multi-informant approach and findings, this study provides novel answers to the need for more robust evidence on the effectiveness of school-based interventions for refugee and immigrant children (Tyrer & Fazel, 2014; Sullivan & Simonson, 2016). The study furthermore sheds particular light on the potential value of interventions targeting school-based relational dynamics as a key aspect in supporting the mental health of children with migration backgrounds.

This study's findings might be of particular importance in the current COVID-condition, as emerging research underscores the importance of better supporting refugee and immigrant school children, who appear particularly vulnerable to the impact of the pandemic and related policies (e.g., quarantine, school closures) on their school trajectories, access to adequate care and, ultimately, their well-being (Langer Primdahl et al., 2021).

References

Achenbach, T. M., & Rescorla, L. A. (2001) *Manual for the ASEBA school-age forms & profiles.* University of Vermont, Burlington.

Achenbach, T. M., & Rescorla, L. A. (2007) *Multicultural supplement of the manual for the ASEBA school-age forms and profiles.* University of Vermont, Burlington.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*(1), 5-37.

Barrett, P. M., Sonderegger, R., & Xenos, S. (2003). Using FRIENDS to combat anxiety and adjustment problems among young migrants to Australia: A national trial. *Clinical Child Psychology and Psychiatry*, *8*(2), 241-260.

Beehler, S., Birman, D., & Campbell, R. (2012). The effectiveness of cultural adjustment and trauma services (CATS): Generating practice-based evidence on a comprehensive, school-based mental health intervention for immigrant youth. *American Journal of Community Psychology, 50*, 15–68.

Beiser, M., & Hou, F. (2016). Mental health effects of premigration trauma and postmigration discrimination on refugee youth in Canada. *The Journal of Nervous and Mental Disease, 204*(6), 464–470.

Bennouna, C., Khauli, N., Basir, M., Allaf, C., Wessells, M., & Stark, L. (2019). School-based programs for supporting the mental health and psychosocial wellbeing of adolescent forced migrants in high-income countries: A scoping review. *Social Science & Medicine*, *239*, 112558.

Betancourt, T. S., Newnham, E. A., Birman, D., Lee, R., Ellis, B. H., & Layne, C. M. (2017). Comparing trauma exposure, mental health needs, and service utilization across clinical

samples of refugee, immigrant, and US-origin children. *Journal of Traumatic Stress, 30*(3), 209-218.

Betancourt, T. S., Newnham, E. A., Layne, C. M., Soeun, K., Steinberg, A. M., Ellis, H., & Birman, D. (2012). Trauma history and psychopathology in war-affected refugee children referred for trauma-related mental health services in the United States. *Journal of Traumatic Stress, 25*, 682–690.

Bhugra, D., Gupta, S., Schouler-Ocak, M., Graeff-Calliess, I., Deakin, N. A., Qureshi, A., ... & Till, A. (2014). EPA guidance mental health care of migrants. *European Psychiatry*, *29*(2), 107-115.

Black, A. C., Harel, O., & McCoach, D.B. (2011). Missing data techniques for multilevel data: Implications of model misspecification. *Journal of Applied Statistics*, *38*(9), 1845-1865.

Block, K., Warr, D., Gibbs, L., & Riggs, E. (2012). Addressing ethical and methodological challenges in research with refugee-background young people: reflections from the field. *Journal of Refugee Studies*, *26*(1), 69-87.

Bodvin, K., Verschueren, K., De Haene, L., & Struyf, E. (2018). Social inequality in education and the use of extramural support services: Access and parental experiences in disadvantaged families. *European Journal of Psychology of Education*, *33*(2), 215-233.

Bolland, K. (2010). *Children on the move. A report on children of Afghan origin moving to Western countries.* New York: Unicef.

Bourdon, K., Goodman, R., Rae, D., Simpson, G., & Koretz, D. (2005). The Strengths and Difficulties Questionnaire: US normative data and psychometric properties. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(6), 557–564.

Breeman, L. D., Wubbels, T., Van Lier, P. A. C., Verhulst, F. C., Van der Ende, J., Maras, A., ... & Tick, N. T. (2015). Teacher characteristics, social classroom relationships, and

children's social, emotional, and behavioral classroom adjustment in special education. *Journal of School Psychology*, *53*(1), 87-103.

Buchmüller, T., Lembcke, H., Busch, J., Kumsta, R., & Leyendecker, B. (2018). Exploring mental health status and syndrome patterns among young refugee children in Germany. *Frontiers in Psychiatry*, *9*, 212.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). Consort 2010 statement: extension to cluster randomised trials. *British Medical Journal*, *345*, e5661.

Ceri, V., Özlü-Erkilic, Z., Özer, Ü., Kadak, T., Winkler, D., Dogangün, B., & Akkaya-Kalayci, T. (2017). Mental health problems of second generation children and adolescents with migration background. *International Journal of Psychiatry in Clinical Practice*, *21*(2), 142-147.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Curvis, W., McNulty, S., & Qualter, P. (2014). The validation of the self-report Strengths and Difficulties Questionnaire for use by 10-year-old children in the UK. *British Journal of Clinical Psychology, 53*, 131-137.

De Haene, L., Adriaenssens, P., Deruddere, N., & Rober, P. (2020). Trauma narration in family therapy with refugees: Working between silence and story in supporting a meaningful engagement with family trauma history. In L. De Haene & C. Rousseau (Eds.), *Working with refugee families: Trauma and exile in family relationships* (pp. 148-171). Cambridge, UK: Cambridge University Press.

De Haene, L., Dalgaard, N. T., Montgomery, E., Grietens, H., & Verschueren, K. (2013). Attachment narratives in refugee children: Interrater reliability and qualitative analysis in pilot findings from a two-site study. *Journal of Traumatic Stress, 26,* 413-417.

De Haene, L., Grietens, H., & Verschueren, K. (2010a). Adult attachment in the context of refugee traumatisation: The impact of organized violence and forced separation on parental states of mind regarding attachment. *Attachment & Human Development, 12*, 249-264.

De Haene, L., Grietens, H., & Verschueren, K. (2010b). Holding harm: Narrative methods in mental health research on refugee trauma. *Qualitative Health Research*, *20*(12), 1664–1676.

de Smet, S., Rousseau, C., Deruddere, N., Kevers, R., Spaas, C., Missotten, L.C., Stalpaert, C., & De Haene, L. (2021). A contextualized perspective on research participation in collaborative refugee research: A multi-site exploration of relational dynamics in collaborative research. *Cultural Diversity & Ethnic Minority Psychology*, 1-13.

de Wal Pastoor, L. (2015). The mediational role of schools in supporting psychosocial transitions among unaccompanied young refugees upon resettlement in Norway. *International Journal of Educational Development, 41,* 245–254.

Derluyn, I., Broekaert, E., & Schuyten, G. (2008). Emotional and behavioural problems in migrant adolescents in Belgium. *European Child & Adolescent Psychiatry, 17*(1), 54-62.

Derluyn, I., Mels, C., Watters, C. & Broekaert, E. (2014). 'We are all the same, coz exist only one earth, why the BORDER EXIST'. Messages of migrants on their way. *Journal of Refugee Studies, 27*, 1-20.

Dimitrova, R., Chasiotis, A., & Van De Vijver, F. (2016). Adjustment outcomes of immigrant children and youth in Europe: A meta-analysis. *European Psychologist*, *21*(2), 150–162.

Donkers, A., & Vermulst, A.A. (2014). *Klimaatschaal: Handleiding [Class Climate Scale: Manual].* Uden, the Netherlands: Orthopraktijk Ad Donkers & Ad Vermulst.

Ebesutani, C., Bernstein, A., Martinez, J. I., Chorpita, B. F., & Weisz, J. R. (2011). The Youth Self Report: Applicability and validity across younger and older youths. *Journal of Clinical Child & Adolescent Psychology*, *40*(2), 338-346.

Ehntholt, K. A., Smith, P. A., & Yule, W. (2005). School-based cognitive-behavioural therapy group intervention for refugee children who have experienced war-related trauma. *Clinical Child Psychology and Psychiatry*, *10*, 235–250.

Ellis, B. H., Lincoln, A. K., Charney, M. E., Ford-Paz, R., Benson, M., & Strunin, L. (2010). Mental health service utilization of Somali adolescents: Religion, community, and school as gateways to healing. *Transcultural Psychiatry*, *47*(5), 789-811.

Ellis, B.H., Miller, A.B., Abdi; S., Barrett, C., Blood, E.A., & Betancourt, T.S. (2013). Multi-tier mental health program for refugee youth. *Journal of Consulting and Clinical Psychology, 81*, 129-140.

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour research and therapy*, *98*, 4-18.

Érit (L'équipe de recherche et d'intervention transculturelles) (2010). *Art et contes. Manuel de formation – Atelier d'expression créatrice.* Montréal, Canada.

Evans, C. H., & Ildstad, S. T. (Eds.). (2001). *Small clinical trials: Issues and challenges.* Washington, DC: NAP.

Fazel, M. (2015). A moment of change: Facilitating refugee children's mental health in UK schools. *International Journal of Educational Development, 41,* 255-261.

Fazel, M., Reed, R., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in high-income countries: risk and protective factors. *Lancet, 379*, 266-282.

Fête, M., Aho, J., Benoit, M. et al. (2019). Barriers and recruitment strategies for precarious status migrants in Montreal, Canada. *BMC Med Res Methodology 19*(41).

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Franco, D. (2018). Trauma without borders: The necessity for school-based interventions in treating unaccompanied refugee minors. *Child and Adolescent Social Work Journal*, *35*(6), 551–565.

Fung, J., Kim, J. J., Jin, J., Chen, G., Bear, L., & Lau, A. S. (2019). A randomized trial evaluating school-based mindfulness intervention for ethnic minority youth: exploring mediators and moderators of intervention effects. *Journal of Abnormal Child Psychology*, *47*(1), 1-19.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337–1345.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *British Journal of Psychiatry, 177*, 534–539.

Green, H., McGinnity, A., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of children and young people in Great Britain, 2004.* Hampshire, UK: Palgrave Macmillan.

Hugman, R., Bartolomei, L., & Pittaway, E. (2011). Human agency and the meaning of informed consent: Reflections on research with refugees. *Journal of Refugee Studies*, *24*(4), 655–671.

Hynie, M. (2018). The social determinants of refugee mental health in the post-migration context: A critical review. *The Canadian Journal of Psychiatry, 63*(5), 297-303.

Jackson, M. I., Kiernan, K., & McLanahan, S. (2012). Immigrant–native differences in child health: Does maternal education narrow or widen the gap? *Child Development, 83*(5), 1501-1509.

Jaycox, L. H., Stein, B. D., Kataoka, S. H., Wong, M., Fink, A., Escudero, P. I. A., & Zaragoza, C. (2002). Violence exposure, posttraumatic stress disorder, and depressive symptoms among recent immigrant schoolchildren. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(9), 1104-1110.

Kataoka, S. H., Stein, B. D., Jaycox, L. H., Wong, M., Escudero, P., Tu, W., ... & Fink, A. (2003). A school-based mental health program for traumatized Latino immigrant children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(3), 311-318.

Keles, S., Friborg, O., Idsøe, T., Sirin, S., & Oppedal, B. (2016). Depression among unaccompanied minor refugees: the relative contribution of general and acculturation-specific daily hassles. *Ethnicity & Health*, *21*(3), 300–317.

Kevers, R., Spaas, C., Colpin, H., Van Den Noortgate, W., de Smet, S., Derluyn, I., & De Haene, L. (under revision). Mental health problems in refugee and immigrant elementary school children in Flanders, Belgium.

Kia-Keating, M., & Ellis, B. H. (2007). Belonging and connection to school in resettlement: Young refugees, school belonging, and psychosocial adjustment. *Clinical Child Psychology and Psychiatry*, *12*(1), 29-43.

Kira, I. A., Lewandowski, L., Ashby, J. S., Somers, C., Chiodo, L., & Odenat, L. (2014). Does bullying victimization suppress IQ? The effects of bullying victimization on IQ in Iraqi and African American adolescents: A traumatology perspective. *Journal of Aggression, Maltreatment & Trauma*, *23*(5), 431-453.

Kouider, E. B., Koglin, U., & Petermann, F. (2014). Emotional and behavioral problems in migrant children and adolescents in Europe: A systematic review. *European Child & Adolescent Psychiatry, 23*(6), 373-391.

Lacroix, L., Rousseau, C., Gauthier, M. F., Singh, A., Giguère, N., & Lemzoudi, Y. (2007). Immigrant and refugee preschoolers' sandplay representations of the tsunami. *The Arts in Psychotherapy*, *34*(2), 99-113.

Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198–1202.

Mackenzie, C., McDowell, C., & Pittaway, E. (2007). Beyond 'do no harm': The challenge of constructing ethical relationships in refugee research. *Journal of Refugee Studies*, *20*(2), 299-319.

Measham, T., Guzder, J., Rousseau, C., Pacione, L., Blais-McPherson, M., & Nadeau, L. (2014). Refugee children and their families: Supporting psychological well-being and positive adaptation following migration. *Current Problems in Pediatric and Adolescent Health Care, 44*, 208–215.

Montgomery, E. (2008). Long-term effects of organized violence on young Middle Eastern refugees' mental health. *Social Science and Medicine, 67*, 1596-603.

Montgomery, E. (2011). Trauma, exile and mental health in young refugees. *Acta Psychiatrica Scandinavica, 124*, 1-46.

Montgomery, E., & Foldspang, A. (2008). Discrimination, mental problems and social adaptation in young refugees. *European Journal of Public Health, 18,* 156-61.

Murray, K.E., Davidson, G.R., & Schweitzer, R.D. (2010). Review of refugee mental health interventions following resettlement: Best practices and recommendations. *American Journal of Orthopsychiatry, 80*, 576–85.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. In P. V. Marsden (Ed.), *Sociological methodology* (pp. 267−316). Blackwell.

Nadeau, L., Rousseau, C., & Measham, T. (2014). Addressing cultural diversity through collaborative care. In *Cultural consultation* (pp. 203-221). Springer.

Noam, G., Oppedal, B., Idsoe, T., & Panjwani, N. (2014). Mental health problems and school

    outcomes among immigrant and non-immigrant early adolescents in Norway. *School Mental*

    Health, *6*(4), 279-293.

Nunes, E. V., Pavlicova, M., Hu, M. C., Campbell, A. N., Miele, G., Hien, D., & Klein, D. F.

    (2011). Baseline matters: The importance of covariation for baseline severity in the analysis

    of clinical trials. *The American Journal of Drug and Alcohol Abuse*, *37*(5), 446-452.

Onderwijs Vlaanderen (n.d.). Anderstalige nieuwkomers.

    https://onderwijs.vlaanderen.be/nl/anderstalige-nieuwkomers

Özbek, E., Bongers, I. L., Lobbestael, J., & van Nieuwenhuizen, C. (2015). Turkish and

    Moroccan young adults in the Netherlands: The relationship between acculturation and

    psychological problems. *Journal of Immigrant and Minority Health*, *17*(6), 1687-1696.

Pacione, L., Measham, T., & Rousseau, C. (2013). Refugee children: Mental health and

    effective interventions. *Current Psychiatry Reports, 15*, 341-349.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School*

    Psychology, *48*(1), 85-112.

Priest, N., Paradies, Y., Trenerry, B., Truong, M., Karlsen, S., & Kelly, Y. (2013). A systematic

    review of studies examining the relationship between reported racism and health and

    wellbeing for children and young people. *Social Science & Medicine, 95*, 115-127.

Pulinx, R., Kemper, R., Verswijvel, K., De Bruyne, L., Vanthournout, G., Baeten, M., ... &

    Vandommele, G. (2017). Cartografie en analyse van het onthaalonderwijs voor anderstalige

    nieuwkomers en OKAN-leerlingen: het onthaalonderwijs voor anderstalige nieuwkomers en

    OKAN-leerlingen in kaart brengen en kritisch analyseren.

Rousseau, C., Benoit, M., Gauthier, M., Lacroix, L., Alain, N., Viger Rojas, M., . . . Bourassa,

    D. (2007). Classroom drama therapy program for immigrant and refugee adolescents: A pilot

    study. *Clinical Child Psychology and Psychiatry*, *12*, 451–465.

Rousseau, C., Drapeau, A., Lacroix, L., Bagilishya, D., & Heusch, N. (2005). Evaluation of a classroom program of creative expression workshops for refugee and immigrant children. *Journal of Child Psychology and Psychiatry*, *46*(2), 180-185.

Rousseau, C., Ferradji, T., Mekki-Berrada, A., & Jamil, U. (2013). North African Muslim immigrant families in Canada giving meaning to and coping with the war on terror. *Journal of Immigrant & Refugee Studies*, *11*(2), 136-156.

Rousseau, C., & Guzder, J. (2008). School-based prevention programs for refugee children. *Child and Adolescent Psychiatric Clinics of North America, 17*, 533-549.

Rousseau, C., Jamil, U., Bhui, K., & Boudjarane, M. (2015). Consequences of 9/11 and the war on terror on children's and young adult's mental health: A systematic review of the past 10 years. *Clinical Child Psychology and Psychiatry, 20*(2), 173-193.

Rousseau, C., Lacroix, L., Bagilishya, D., & Heusch, N. (2003). Working with myths: Creative expression workshops for immigrant and refugee children in a school setting. *Art Therapy, 20*(1), 3-10.

Rousseau, C., Singh, A., Lacroix, L., Bagilishya, D., & Measham, T. (2004). Creative expression workshops for immigrant and refugee children. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*(2), 235-238.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods, 12*(2), 347–367.

Spaas, C., Verelst, A., Devlieger, I., Aalto, S., Andersen, A.J., Durbeej, N., Hilden, P.K., Kankaapää, R., Langer Primdahl, N., Opaas, M., Osman, F., Peltonen, K., Sarkadi, A., Skovdal, M., Jervelund, S.S., Soye, E., Watters, C., Derluyn, I., Colpin, H., & De Haene, L. (2021). Mental health of refugee and non-refugee newcomer youth in European secondary

education: The role of family separation, daily material stressors and discrimination in resettlement. *Journal of Youth & Adolescence, 51*, 848-870.

Spilt, J. L., Leflot, G., Onghena, P., & Colpin, H. (2016). Use of praise and reprimands as critical ingredients of teacher behavior management: Effects on children's development in the context of a teacher-mediated classroom intervention. *Prevention Science*, *17*(6), 732-742.

Sullivan, A. L., & Simonson, G. R. (2016). A systematic review of school-based social-emotional interventions for refugee and war-traumatized youth. *Review of Educational Research*, *86*(2), 503-530.

Taber, K.S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273-1296.

Tyrer, R.A., & Fazel, M. (2014). School- and community-based interventions for refugee and asylum-seeking children: A systematic review. *PLOS ONE, 9*, e89359.

UNHCR (2019). *Global trends: Forced displacement in 2018.* Retrieved from: https://www.unhcr.org/5d08d7ee7.pdf

Vancraeyveldt, C., Verschueren, K., Wouters, S., Van Craeyevelt, S., Van Den Noortgate, W., & Colpin, H. (2015). Improving teacher-child relationship quality and teacher-rated behavioral adjustment amongst externalizing preschoolers: Effects of a two-component intervention. *Journal of Abnormal Child Psychology, 43*, 243-257.

Van Den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention*, *2*(3), 142-151.

Van Oort, F. V., Joung, I. M., Mackenbach, J. P., Verhulst, F. C., Bengi-Arslan, L., Crijnen, A. A., & Van Der Ende, J. (2007). Development of ethnic disparities in internalizing and

externalizing problems from adolescence into young adulthood. *Journal of Child Psychology and Psychiatry, 48*(2), 176-184.

Verhulp, E. E., Stevens, G. W., Van de Schoot, R., & Vollebergh, W. A. (2013). Understanding ethnic differences in mental health service use for adolescents' internalizing problems: the role of emotional problem identification. *European Child & Adolescent Psychiatry*, *22*(7), 413-421.

Vervliet, M., Lammertyn, J., Broekaert, E. & Derluyn, I. (2014). Longitudinal follow-up of the mental health of unaccompanied refugee minors. *European Journal of Child & Adolescent Psychiatry, 23*, 5, 337-346.

Walsh, S. D., Harel-Fisch, Y., & Fogel-Grinvald, H. (2010). Parents, teachers and peer relations as predictors of risk behaviors and mental well-being among immigrant and Israeli born adolescents. *Social Science & Medicine*, *70*(7), 976-984.

Yaylaci, F. T. (2018). Trauma and resilient functioning among Syrian refugee children. *Development and Psychopathology, 30*(5), 1923-1936. doi: 10.1017/S0954579418001293

You, D. S., Youngstrom, E. A., Feeny, N. C., Youngstrom, J. K., & Findling, R. L. (2017). Comparing the diagnostic accuracy of five instruments for detecting posttraumatic stress disorder in youth. *Journal of Clinical Child & Adolescent Psychology*, *46*(4), 511-522.

YouthinMind (2016). Scoring the Strengths and Difficulties Questionnaire. Retrieved on August 23d 2019 from https://www.sdqinfo.com/py/sdqinfo/c0.py

Zych, I., Farrington, D., Llorent, V., & Ttofi, M. (2017). *Protecting children against bullying and its consequences* (Springer Briefs in Psychology). Springer.

Figure 1

Flowchart of Classes and Participants in the Cluster Randomized Trial

ENROLLMENT　　　　　　　　　　**Selection of participant schools**
September-December 2017　　　　　Schools assessed for eligibility (N = 30)
　　　　　　　　　　　　　　　　　Schools interested in participation (N = 8)
　　　　　　　　　　　　　　　　　Schools selected for participation (N = 3)

School 1　　　　　　　　　School 2　　　　　　　　　School 3

Participating classes (N = 4)　Participating classes (N = 8)　Participating classes (N = 1)
Grade 3 (N = 2)　　　　　　　Grade 3 (N = 2)
Grade 5 (N = 2)　　　　　　　Grade 4 (N = 2)
　　　　　　　　　　　　　　Grade 5 (N = 2)
　　　　　　　　　　　　　　Grade 6 (N = 2)

RANDOM ASSIGNMENT　　　**Allocation of participant classes**
January 2018

Classes allocated to **INTERVENTION** (N = 7)　　Classes allocated to **CONTROL** (N = 6)
Child participants (N = 68)　　　　　　　　　　Child participants (N = 52)
[Total child participant sample: N = 120; Excluded N = 154 (no parental consent)]

BASELINE MEASUREMENT　　　**Multi-informant assessment**
March 2018

　　　　Children (N = 65)　　　　　　　　　　　　Children (N = 52)
　　　　Parents (N = 65)　　　　　　　　　　　　Parents (N = 50)
　　　　Teachers (N = 66)　　　　　　　　　　　Teachers (N = 51)

FOLLOW-UP MEASUREMENT　　　**Multi-informant assessment**
June 2018

　　　　Children (N = 65)　　　　　　　　　　　　Children (N = 52)
　　　　Parents (N = 43)　　　　　　　　　　　　Parents (N = 38)
　　　　Teachers (N = 67)　　　　　　　　　　　Teachers (N = 51)

Table 1

Means and Standard Deviations of the Outcome Variables for the Total Sample (N = 120) and for the

Intervention (n = 68) and Control Group (n = 52) Respectively

Scale	Total Group	Intervention Group	Control Group	U (p)
SDQ-self report				
Total problems-T1	11.66 (5.46)	10.91 (5.43)	12.60 (5.41)	1369 (0.078)
Total problems-T2	12.00 (5.96)	11.70 (5.81)	12.37 (6.16)	
Externalizing-T1	5.16 (3.30)	4.82 (3.30)	5.60 (3.27)	1491.5 (0.274)
Externalizing-T2	5.71 (3.54)	5.75 (3.52)	5.67 (3.61)	
SDQ-parent report				
Total problems-T1	10.30 (6.21)	10.08 (5.96)	10.58 (6.57)	1578.5 (0.793)
Total problems-T2	8.85 (5.66)	9.07 (6.05)	8.61 (5.25)	
Externalizing-T1	5.22 (3.73)	5.31 (3.64)	5.10 (3.88)	1727 (0.698)
Externalizing-T2	4.47 (3.50)	4.84 (4.05)	4.05 (2.76)	
Internalizing-T1	5.03 (3.55)	4.77 (3.43)	5.38 (3.71)	1469.5 (0.378)
Internalizing-T2	4.38 (3.21)	4.23 (3.24)	4.55 (3.20)	
SDQ-teacher report				
Total problems-T1	6.92 (5.10)	7.30 (5.26)	6.43 (4.90)	1868 (0.308)
Total problems-T2	6.64 (5.17)	6.84 (5.57)	6.37 (4.63)	
Externalizing-T1	4.13 (3.80)	4.52 (4.01)	3.63 (3.49)	1889.5 (0.252)
Externalizing-T2	3.87 (4.00)	4.34 (4.23)	3.25 (3.61)	
Internalizing-T1	2.79 (2.88)	2.79 (2.78)	2.80 (3.03)	1714.5 (0.860)
Internalizing-T2	2.76 (2.76)	2.49 (2.71)	3.12 (2.81)	
Impact score-T1	0.59 (1.15)	0.67 (1.21)	0.49 (1.08)	1844 (0.269)
Impact score-T2	0.57 (1.04)	0.69 (1.16)	0.39 (0.83)	
YSR				
PTSP-T1	10.17 (4.99)	9.52 (5.00)	10.98 (4.90)	1386 (0.095)
PTSP-T2	10.36 (5.42)	10.06 (5.39)	10.73 (5.49)	
Climate Scale				
Class climate-T1	2.76 (0.74)	2.56 (0.67)	3.01 (0.76)	**1041 (0.000)**
Class climate-T2	2.80 (0.76)	2.78 (0.72)	2.82 (0.81)	
Peer relationships-T1	3.14 (0.64)	3.12 (0.64)	3.17 (0.65)	1589 (0.579)
Peer relationships-T2	3.12 (0.66)	3.08 (0.73)	3.17 (0.56)	

Table 2

Parameter Estimates (and Corresponding Standard Errors) for the Multilevel Model Including Cross-Level Interactions for SDQ – Total Difficulties (Self-, Parent-, and Teacher-Report) and SDQ – Externalizing Difficulties (Self-Report)

Parameter	SDQ-total-self n=115	SDQ-total-parent n=81	SDQ-total-teacher n=117	SDQ-extern-self n=115
Fixed effects				
Intercept	12.66 (0.90)***	10.03 (0.94)***	7.00 (0.84)***	5.69 (0.53)***
Intervention	-0.03 (1.22)	-0.11 (1.38)	-0.33 (1.14)	0.35 (0.72)
Baseline	0.75 (0.11)***	0.71 (0.12)***	0.64 (0.10)***	0.70 (0.12)***
Refugee	-2.27 (1.17)°	-1.16 (1.30)	0.07 (1.03)	-0.74 (0.77)
Intervention * Baseline	-0.04 (0.15)	0.02 (0.16)	0.16 (0.13)	-0.02 (0.16)
Intervention * Refugee	1.61 (1.60)	1.40 (1.83)	0.18 (1.37)	0.62 (1.04)
Variances				
Intercept (level 2)	0.99 (1.22)	1.14 (1.95)	1.36 (1.07)	0.09 (0.35)
Residual (level 1)	16.29 (2.31)***	13.83 (2.52)***	11.12 (1.57)***	7.17 (1.01)***
Parameters	8	8	8	8
ICC^2	0.06	0.08	0.11	0.01
Design effect^3	1.49	1.66	1.91	1.08

°$p < .10$
*$p < .05$
**$p < .01$
***$p < .001$

---

[2] ICC = intraclass correlation. The ICC can be defined both as the proportion of outcome score variation that occurs across classes (i.e., level-2 units) and as the expected correlation between the outcome scores of two students (i.e., level-1 units) from the same class (Peugh, 2010). The ICC can be calculated as follows: ICC = level 2 variance / (level 2 variance + level 1 variance).

[3] The design effect quantifies the effect of independence violations on standard error estimates and is an estimate of the multiplier that needs to be applied to standard errors to correct for the negative bias that results from nested data (Peugh, 2010). Some researchers believe that design effect estimates greater than 2.0 indicate a need for multilevel analysis (e.g., Muthén & Satorra, 1995). The design effect can be computed as follows: Design effect = 1 + (n_c-1)*ICC. [n_c = average number of students per class (120/13 = 9.23)]

Table 3

Parameter Estimates (and Corresponding Standard Errors) for the Multilevel Models Including Cross-Level Interactions for SDQ – Internalizing and Externalizing Difficulties (Parent- and Teacher-Report)

Parameter	SDQ-intern-parent n=81	SDQ-intern-teacher n=117	SDQ-extern-parent n=81	SDQ-extern-teacher n=118
Fixed effects				
Intercept	5.06 (0.48)***	3.24 (0.56)***	4.83 (0.56)***	3.98 (0.55)***
Intervention	-0.06 (0.71)	-0.51 (0.76)	-0.29 (0.82)	-0.03 (0.74)
Baseline	0.69 (0.12)***	0.69 (0.09)***	0.57 (0.11)***	0.78 (0.10)***
Refugee	-1.08 (0.78)	0.08 (0.55)	-0.07 (0.73)	-0.25 (0.73)
Intervention * Baseline	-0.04 (0.16)	0.01 (0.12)	0.35 (0.15)*	0.02 (0.13)
Intervention * Refugee	0.40 (1.06)	0.06 (0.74)	1.45 (1.04)	0.50 (0.97)
Variances				
Intercept (level 2)	0.02 (0.53)	1.07 (0.65)	0.54 (0.76)	0.34 (0.39)
Residual (level 1)	4.98 (0.95)***	3.09 (0.44)***	4.43 (0.82)***	5.86 (0.82)***
Parameters	8	8	8	8
ICC	0.004	0.26	0.11	0.05
Design effect	1.03	3.14	1.91	1.41

$°p < .10$
$*p < .05$
$**p < .01$
$***p < .001$

Table 4

Parameter Estimates (and Corresponding Standard Errors) for the Multilevel Models Including Cross-Level Interactions for SDQ – Impact Score (Teacher-Report), Post-Traumatic Stress Problems, Quality of Mutual Peer Relationships, and Class Climate

| Parameter	SDQ-impact-teacher n=113	PTSP n=114	Quality of mutual peer relationships n=115	Class climate n=115
Fixed effects				
Intercept	0.39 (0.18)*	11.61 (0.93)***	3.17 (0.13)***	2.66 (0.12)***
Intervention	0.20 (0.24)	-1.01 (1.26)	-0.13 (0.17)	0.12 (0.17)
Baseline	0.38 (0.11)**	0.68 (0.13)***	0.41 (0.10)***	0.61 (0.12)***
Refugee	0.25 (0.25)	-2.60 (1.26)*	0.01 (0.14)	0.02 (0.18)
Intervention * Baseline	0.26 (0.14)°	-0.29 (0.18)°	0.31 (0.14)*	-0.15 (0.17)
Intervention * Refugee	-0.09 (0.32)	1.76 (1.69)	0.03 (0.19)	0.19 (0.24)
Variances				
Intercept (level 2)	0.03 (0.04)	0.88 (1.74)	0.04 (0.03)	0.00[a]
Residual (level 1)	0.62 (0.09)***	18.44 (2.75)***	0.21 (0.03)***	0.40 (0.05)***
Parameters	8	8	8	8
ICC	0.05	0.05	0.16	0
Design effect	1.41	1.41	2.32	1

°*p* < .10
p < .05
**p* < .01
***p* < .001
[a] This covariance parameter is redundant. The test statistic cannot be computed.

Table 5

Parameter Estimates (and Corresponding Standard Errors) for All Estimated Multilevel Models (Using a Step-by-Step Approach) for Climate Scale – Class Climate

Parameter	Unconditional model	Model with level 2 predictor (intervention)	Model with level 1 predictors (control variables)	Model with cross-level interactions
Fixed effects				
Intercept	2.82 (0.08)***	2.84 (0.13)***	2.63 (0.11)***	2.66 (0.12)***
Intervention		-0.03 (0.17)	0.21 (0.13)°	0.12 (0.17)
Baseline			0.54 (0.09)***	0.61 (0.12)***
Refugee			0.13 (0.12)	0.02 (0.18)
Intervention * Baseline				-0.15 (0.17)
Intervention * Refugee				0.19 (0.24)
Random effects				
Intercept (level 2)	0.02 (0.03)	0.02 (0.03)	0.00ª	0.00ª
Residual (level 1)	0.55 (0.08)***	0.55 (0.08)***	0.41 (0.05)***	0.40 (0.05)***
Deviance (-2LL)	261.128	261.097	217.640	216.244
Parameters	3	4	6	8

Note. ICC = 0.04; design effect = 1.33

°$p < .10$

*$p < .05$

**$p < .01$

***$p < .001$

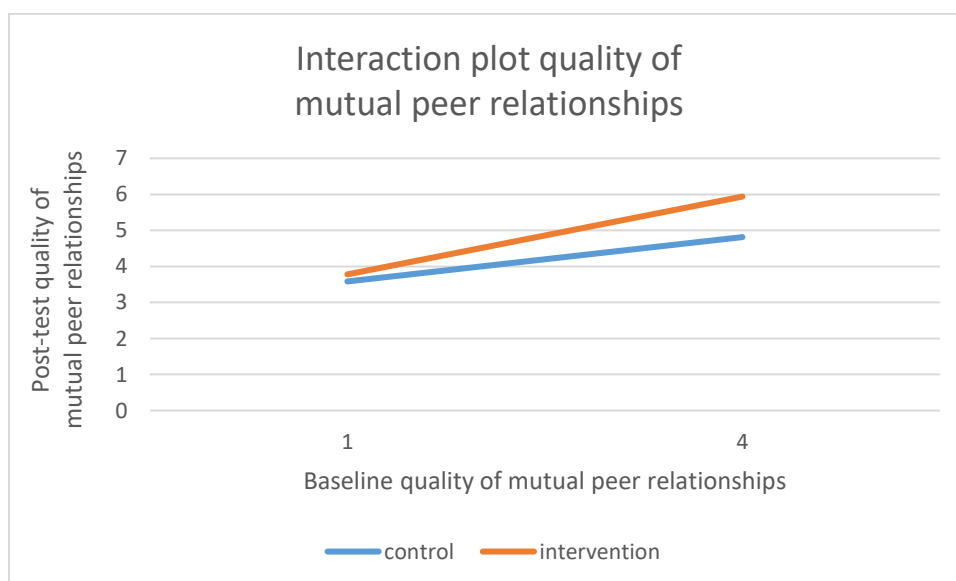ª This covariance parameter is redundant. The test statistic cannot be computed.

Figure 2

Interaction Plot of SDQ – Externalizing Difficulties (Parent-Report)
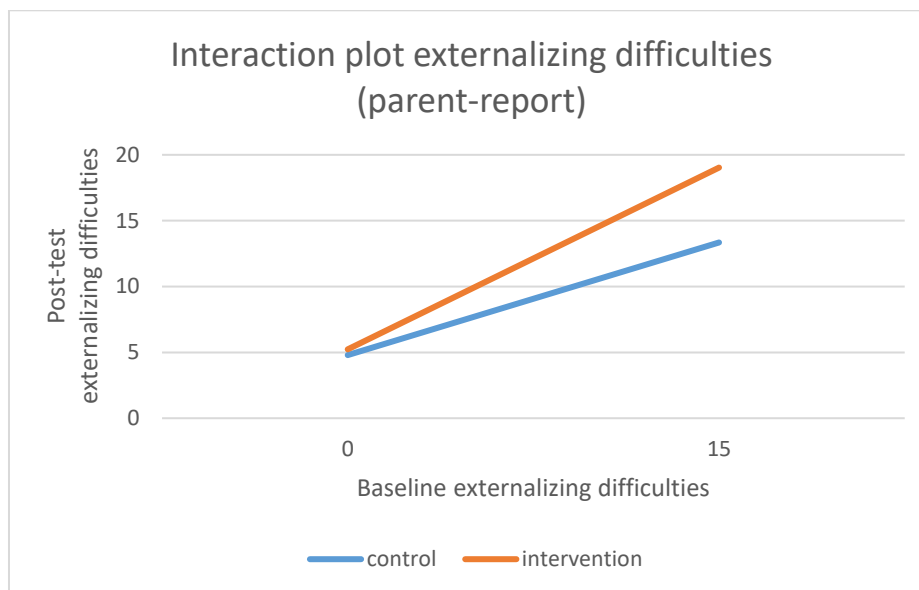
Figure 3

Interaction Plot of SDQ – Impact Score (Teacher-Report)
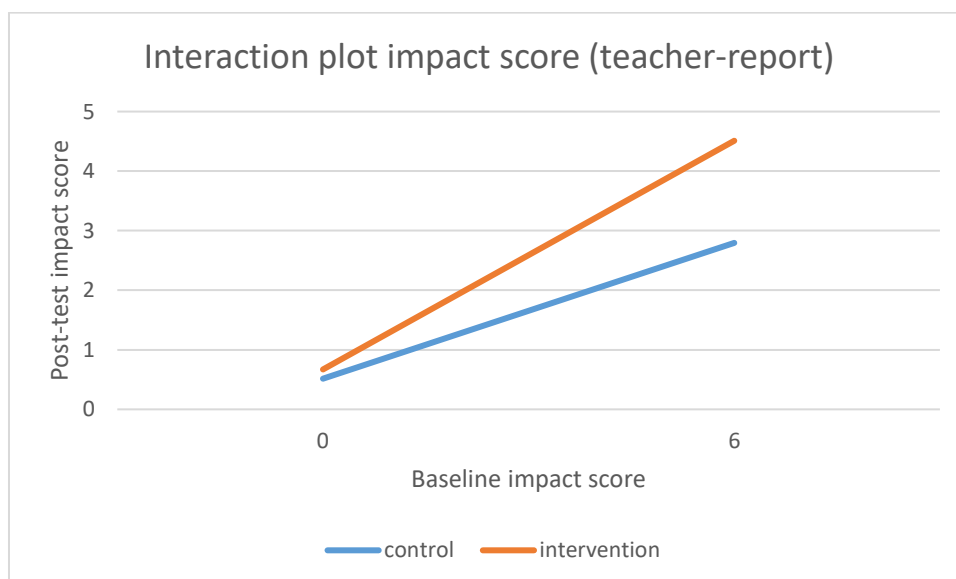
Figure 4

Interaction Plot of Post-Traumatic Stress Problems
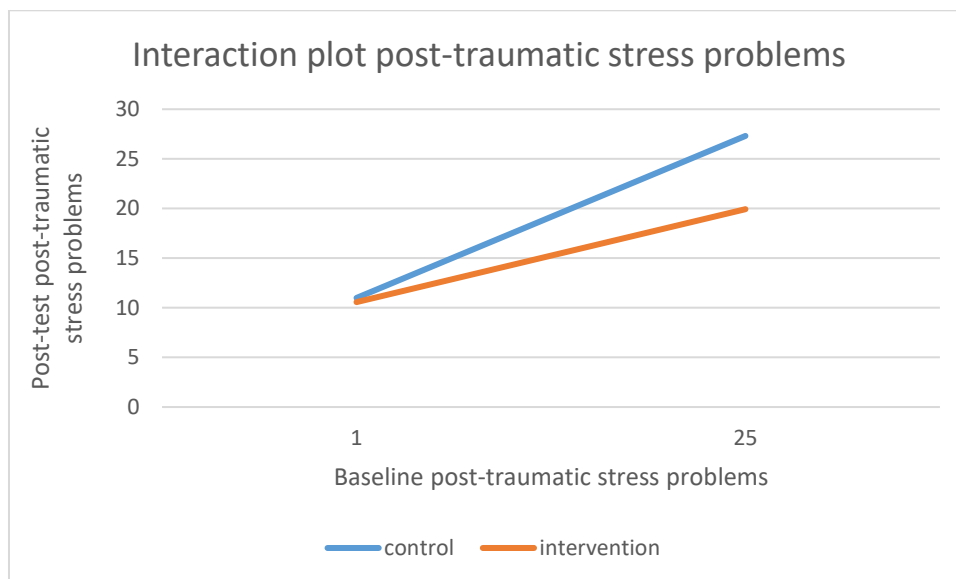
Figure 5

Interaction Plot of Quality of Mutual Peer Relationships

Table 6

Parameter Estimates (and Corresponding Standard Errors) for the Multilevel Models for Youth Self

Report – Post-Traumatic Stress Problems for the Normal Baseline Group (n = 81)

Parameter	Unconditional model	Model with level 2 predictor (intervention)	Model with level 1 predictors (control variables)
Fixed effects			
Intercept	8.90 (0.74)***	8.36 (1.09)***	9.12 (1.29)***
Intervention		0.95 (1.45)	0.74 (1.60)
Refugee			-1.12 (1.06)
Random effects			
Intercept (level 2)	3.24 (3.03)	2.91 (2.94)	4.31 (3.80)
Residual (level 1)	19.69 (3.46)***	19.75 (3.49)***	18.75 (3.42)***
Deviance (-2LL)	468.071	467.657	454.654
Parameters	3	4	5

Note. ICC = 0.14; design effect = 2.16

°$p < .10$

*$p < .05$

**$p < .01$

***$p < .001$

Table 7

Parameter Estimates (and Corresponding Standard Errors) for the Multilevel Models for Youth Self

Report – Post-Traumatic Stress Problems for the Borderline/Clinical Baseline Group (n = 36)

Parameter	Unconditional model	Model with level 2 predictor (intervention)	Model with level 1 predictors (control variables)
Fixed effects			
Intercept	13.91 (0.85)***	15.39 (1.14)***	17.32 (1.19)***
Intervention		-3.04 (1.63)° (p=.071)	-3.86 (1.54)* (p=.038)
Refugee			-4.63 (1.35)** (p=.002)
Random effects			
Intercept (level 2)	0.00[a]	0.00[a]	1.44 (3.05)
Residual (level 1)	25.56 (6.11)***	23.26 (5.56)***	14.24 (4.16)**
Deviance (-2LL)	212.767	209.464	189.673
Parameters	3	4	5

Note. ICC = 0; design effect = 1

°*p* < .10

*p < .05

**p < .01

***p < .001

[a] This covariance parameter is redundant. The test statistic cannot be computed.