



JOURNAL ON EDUCATION IN EMERGENCIES

Special Issue on Psychosocial Support and Social and Emotional Learning in Emergencies

EDITORIAL NOTE

Ragnhild Dybdahl and James Williams

EiE RESEARCH ARTICLES

Developing and Validating the International Social and Emotional Learning Assessment: Evidence from a Pilot Test with Syrian Refugee Children in Iraq

Nikhil D'Sa and Allyson Krupar

Teachers' Observations of Learners' Social and Emotional Learning: Psychometric Evidence for Program Evaluation in Education in Emergencies

Ha Yeon Kim, Kalina Gjicali, Zezhen Wu, and Carly Tubbs Dolan

Creating a Tool to Measure Children's Wellbeing: A PSS Intervention in South Sudan

Moses Olayemi, Melissa Tucker, Mamour Choul, Tom Purekal, Arlene Benitez, Wendy Wheaton, and Jennifer DeBoer

How Do We Know If Teachers Are Well? The Wellbeing Holistic Assessment for Teachers Tool

Fernanda Soares, Nina Menezes Cunha, and Paul Frisoli

Evaluating the 3Cs Program for Caregivers of Young Children Affected by the Armed Conflict in Colombia

Lina María González Ballesteros, José M. Flores, Ana María Ortiz Hoyos, Amalia Londoño Tobón, Sascha Hein, Felipe Bolívar Rincon, Oscar Gómez, and Liliana Angélica Ponguta

How Family Relationships Predict the Effectiveness of a Psychosocial Group Intervention among War-Affected Children

Raija-Leena Punamäki, Kirsi Peltonen, Marwan Diab, and Samir R. Qouta

EiE FIELD NOTES

Using a Participatory Approach to Create SEL Programming: The Case of Ahlan Simsim

Shanna Kohn, Kim Foulds, Charlotte Cole, Mackenzie Matthews, and Laila Hussein

Developing a Culturally Relevant Measure of Resilience for War-Affected Adolescents in Eastern Ukraine

Sergiy Bogdanov, Andriy Girnyk, Vira Chernobrovkina, Volodymyr Chernobrovkin, Alexander Vinogradov, Kateryna Harbar, Yuliya Kovalevskaia, Oksana Basenko, Irina Ivanyuk, Kimberly Hook, and Mike Wessells

Developing the Group Facilitation Assessment of Competencies Tool for Group-Based Mental Health and Psychosocial Support Interventions in Humanitarian and Low-Resource Settings

Gloria A. Pedersen, Manaswi Sangraula, Pragya Shrestha, Pooja Lakshmin, Alison Schafer, Renasha Ghimire, Nagendra P. Luitel, Mark J. D. Jordans, and Brandon A. Kohrt

BOOK REVIEWS

Can Big Bird Fight Terrorism? Children's Television and Globalized Multicultural Education by Naomi A. Moland
Kate Lapham

NISSEM Global Briefs: Educating for the Social, the Emotional and the Sustainable edited by Andy Smart, Margaret Sinclair, Aaron Benavot, Jean Bernard, Colette Chabbott, S. Garnett Russell, and James Williams
Solfrid Raknes

COMMENTARY

How the Education in Emergencies Field Can Help the United States Respond to COVID-19
Rebecca Winthrop and Helen Shwe Hadani

JOURNAL ON EDUCATION IN EMERGENCIES

Editor-in-Chief:

Dana Burde, PhD, New York University

Lead Editors for the Special Issue:

Ragnhild Dybdahl, PhD, University of Bergen

James Williams, EdD, George Washington University

Senior Managing Editor:

Heddy Lahmann, PhD, New York University

Deputy Managing Editor:

Nathan Thompson, New York University

Editorial Board:

Carine Allaf, PhD, Qatar Foundation International

Ragnhild Dybdahl, PhD, University of Bergen

Mark Ginsburg, PhD, University of Maryland

Augustino Ting Mayai, PhD, Sudd Institute

Ruth Naylor, PhD, Education Development Trust

S. Garnett Russell, PhD, Teachers College, Columbia University

Sweta Shah, PhD, International Early Childhood Development Consultant

James Williams, EdD, George Washington University

Editorial Assistants:

Sneha Bolisetty, New York University

Samantha Colón, New York University

Deanna Pittmann, New York University

Claudia Segura, New York University

Ruqaiyah Zarook, New York University

Technical Reviewers:

Amanda Blewitt, New York University

Emily Dunlop, New York University

Haider Fancy, New York University

João Souto Maior, New York University

Copy Editor:

Dody Riggs

Designer:

Patrick McCarthy

ISSN 2518-6833

Open-source publication; distribution free of charge; available in soft copy only

Copyright © 2021, Inter-agency Network for Education in Emergencies (INEE).

Inter-agency Network for Education in Emergencies

c/o International Rescue Committee

122 East 42nd Street, 14th floor

New York, NY 10168

<https://inee.org/journal>

journal@inee.org

New York University

International Education Program

246 Greene Street, Suite 300

New York, NY 10003

JOURNAL ON EDUCATION IN EMERGENCIES

VOLUME 7, NUMBER 2

DECEMBER 2021

Special Issue on Psychosocial Support and Social and Emotional Learning in Emergencies

EDITORIAL NOTE

Ragnhild Dybdahl and James Williams.....5

EiE RESEARCH ARTICLES

Developing and Validating the International
Social and Emotional Learning Assessment:
Evidence from a Pilot Test with Syrian Refugee
Children in Iraq
Nikhil D'Sa and Allyson Krupar.....20

Teachers' Observations of Learners' Social and
Emotional Learning: Psychometric Evidence
for Program Evaluation in Education in
Emergencies
*Ha Yeon Kim, Kalina Gjicali, Zezhen Wu, and
Carly Tubbs Dolan*57

Creating a Tool to Measure Children's
Wellbeing: A PSS Intervention in South Sudan
*Moses Olayemi, Melissa Tucker, Mamour Choul,
Tom Purekal, Arlene Benitez, Wendy Wheaton,
and Jennifer DeBoer*..... 104

How Do We Know If Teachers Are Well? The
Wellbeing Holistic Assessment for Teachers
Tool
*Fernanda Soares, Nina Menezes Cunha, and
Paul Frisoli* 152

Evaluating the 3Cs Program for Caregivers
of Young Children Affected by the Armed
Conflict in Colombia
*Lina María González Ballesteros,
José M. Flores, Ana María Ortiz Hoyos, Amalia
Londoño Tobón, Sascha Hein, Felipe Bolívar
Rincon, Oscar Gómez, and Liliana Angélica
Ponguta*..... 212

How Family Relationships Predict the
Effectiveness of a Psychosocial Group
Intervention among War-Affected Children
*Raija-Leena Punamäki, Kirsi Peltonen, Marwan
Diab, and Samir R. Qouta*253

EiE FIELD NOTES

Using a Participatory Approach to Create SEL
Programming: The Case of *Ahlan Simsim*
*Shanna Kohn, Kim Foulds, Charlotte Cole,
Mackenzie Matthews, and Laila Hussein*.....289

Developing a Culturally Relevant Measure of
Resilience for War-Affected Adolescents in
Eastern Ukraine
*Sergiy Bogdanov, Andriy Girnyk, Vira
Chernobrovkina, Volodymyr Chernobrovkin,
Alexander Vinogradov, Kateryna Harbar, Yuliya
Kovalevska, Oksana Basenko, Irina Ivanyuk,
Kimberly Hook, and Mike Wessells* 311

Developing the Group Facilitation Assessment
of Competencies Tool for Group-Based Mental
Health and Psychosocial Support Interventions
in Humanitarian and Low-Resource Settings
*Gloria A. Pedersen, Manaswi Sangraula, Pragya
Shrestha, Pooja Lakshmin, Alison Schafer,
Renasha Ghimire, Nagendra P. Luitel, Mark
J. D. Jordans, and Brandon A. Kohrt*335

BOOK REVIEWS

*Can Big Bird Fight Terrorism? Children's
Television and Globalized Multicultural
Education* by Naomi A. Moland
Kate Lapham.....377

*NISSEM Global Briefs: Educating for the Social,
the Emotional and the Sustainable* edited by
Andy Smart, Margaret Sinclair, Aaron Benavot,
Jean Bernard, Colette Chabbott, S. Garnett
Russell, and James Williams
Solfrid Raknes..... 381

COMMENTARY

How the Education in Emergencies Field Can
Help the United States Respond to COVID-19
Rebecca Winthrop and Helen Shwe Hadani..384

CONTRIBUTORS 391

EDITORIAL NOTE

RAGNHILD DYBDAHL AND JAMES WILLIAMS¹

Emergencies—including conflict, disasters, and forced migration—have negative effects on the psychosocial development of individuals, as well as on the short- and long-term development of local and global communities. Children and youth are particularly vulnerable in these circumstances because of the exposure to adversity, and because they miss out on education, play, nurturing care, and other important childhood experiences. The malnutrition, violence, and grief children experience are associated with neurodevelopmental outcomes and cognitive changes that are likely to affect their learning (Herringa 2017; Shonkoff and Garner 2012; Goenjian et al. 2005; Moore and Varela 2010; Charlson et al. 2019). Moreover, it is not only students who are affected by these mental health conditions but their teachers, parents, caregivers, and community members.

It is widely believed that education can play a central role in the protection and promotion of children’s psychosocial wellbeing in multiple ways. First, education appears to have positive effects on psychosocial wellbeing in and of itself. In addition, educational settings are often staging points for the provision of psychosocial support (Bosqui and Marshoud 2018). Second, schools and other educational settings can provide some stability by offering children safety, predictability, and a sense of accomplishment, dignity, and hope. Third, efforts to bring schooling and mental health interventions together, as embodied by social and emotional learning (SEL) curricula and instructional practices, can help all students acquire the attitudes and skills they need to manage and regulate their complex and difficult emotions, build prosocial attitudes, learn empathy and awareness of others, and develop conflict-resolution skills (Mahoney, Durlak, and Weissberg 2018). Equipping students with these skills in emergency contexts may help reduce the risk of conflict in fragile environments, while simultaneously helping children manage challenging situations.

Finally, advances in cognitive science further highlight the critical nature of the social and emotional dimensions of learning (Jones and Khan 2017). Beyond the cognitive content knowledge children are expected to acquire, positive experiences in educational settings can help children learn to live and thrive—by themselves and with others.

¹ Ragnhild Dybdahl and James Williams served as lead editors for this special issue of the *Journal on Education in Emergencies* and contributed equally to its development and production. Their names are listed alphabetically.

Until recently, the social and emotional dimensions of schooling have been largely informal, and often peripheral to what is considered the primary cognitive purpose of education. And yet, human development researchers from a number of fields increasingly argue that learning and thinking are inherently social and emotional, and that emotional and psychological engagement are inherent in and necessary for learning: “Quite literally, it is neurobiologically impossible to think deeply about or remember information about which one has had no emotion because the healthy brain does not waste energy processing information that does not matter to the individual” (National Academies of Sciences, Engineering, and Medicine 2018, 29). Scientist Mary Helen Immordino-Yang makes a similar claim in Smart et al. (2019): “What I argue—drawing on psychological, anthropological, and even biological perspectives—is that the very nature of human biology is social . . . There is no such thing as non-social thought: your values are derived from and situated in the cultural and temporal context in which you live” (288). This body of research has led to increasing emphasis on the noncognitive dimensions of education, particularly in the industrialized world.

Despite widespread recognition of the importance of the noncognitive dimensions of education, a consensus on terminology and conceptual grounding has yet to be developed. A number of foundational conceptual frameworks have been proposed, most prominently the CASEL framework noted above (Mahoney et al. 2018; see also CASEL 2020a, 2020b), typologies of key 21st-century competencies (OECD 2005), as well as syntheses of the existing research on learning, including the psychosocial and social and emotional components (Jones 2018; National Academies 2018; INEE 2016, 2018; IASC 2017; UNICEF 2015; IFRC 2009; USAID 2019; Bub and Dalrymple 2020). Classification schemes vary widely in their value orientations and research bases, ranging from individual psychologically focused conceptions, such as that of CASEL, to socially oriented ideas of social and emotional skills, such as those of Sustainable Development Goal Target 4.7, to formulations grounded in the cultivation of human capital, such as that of the Organisation for Economic Co-operation and Development (2005). More recent resilience-enhancing interventions often focus on social-environmental interventions and target factors in the social and material environments in order to strengthen the community and schools by providing, for example, support and tools for teachers and parents (Miller et al. 2021). There has been a call for more conceptual clarity in the field of mental health and psychosocial support (MHPSS), which suggests that putting more focus on the causal models that guide decisions on what interventions are appropriate in different conditions and populations may strengthen responses and avoid doing harm (Miller et al. 2021).

Importantly, notwithstanding the lack of consensus in terminology, recent efforts have been made to operationalize promising research findings. For example, recent guidance from the Inter-agency Network on Education in Emergencies (2018) outlines approaches for incorporating psychosocial support (PSS) and social and emotional learning (SEL) into formal and nonformal education responses to emergencies.² In addition, a recent toolkit (MHPSS.net 2021) on MHPSS and education in emergencies (EiE) has made more than 160 resources, tools, and key documents about standards and guidelines available to practitioners in the field.

Despite rapid growth in the body of research on PSS/SEL and efforts to put these findings to use, large gaps in the evidence and the implementation of tools remain. For example, the research does not make it clear how children's and youth's psychosocial wellbeing intersects with emergencies and education. Moreover, there is limited awareness of the available approaches to promoting PSS/SEL or of the appropriate use of these tools.

Thus, many EiE practitioners are working to develop and implement PSS/SEL programming while simultaneously working with researchers to collect evidence on particular measures, instruments, and programs. To this end, the Inter-agency Network on Education in Emergencies has commissioned a measurement library in which it can store tools and measures developed by researchers and practitioners working on PSS/SEL in emergency contexts (see <https://inee.org/measurement-library>).³ Reviewing and making this evidence public is an important part of building knowledge for the field. A substantial body of research has been carried out in high-income countries and in nonemergency contexts, far less in low-income nations and conflict-affected contexts.

This special issue, which contributes to the evidence on PSS/SEL in emergencies, provides a snapshot of strategies and tools developed and used to understand the status of wellbeing and psychosocial support and the effectiveness of programming. It contains six research articles, three field notes, two book reviews, and one commentary. The authors who contributed to the issue work at 30 institutions based in more than 12 countries.

2 Though differing in origins and emphasis, much of the research in the field applies to both concepts, in which case we refer to the collective work in this area as psychosocial support and social and emotional learning, or PSS/SEL.

3 To represent the stages of development of different measures, the measurement library uses images of an olive tree—a seedling to represent measures in their initial stages; a sapling to denote measures with some good evidence; and a mature tree to indicate measures with well-documented validity and reliability for the intended purpose and ready for use.

Assessment and measurement are at the center of much current research on PSS/SEL. This is reflected in this issue as well, where most of the articles wrestle in one way or another with how to conceptualize and appropriately measure the aspects of PSS and SEL they target. As is made clear in this issue, it is challenging to develop appropriate measurement instruments for children and young people affected by conflict. There are tradeoffs between developing ideal measurement processes in the laboratory, confirming their feasibility in the field, and meeting the immediate needs of children and their teachers for effective programming. Local agents, often teachers, must be able to administer and use measurement instruments to assess the wellbeing of their students, and that of other teachers and children's caregivers. Program field staff members need to be able to monitor and evaluate programs, whereas researchers and humanitarian and development agencies need research evidence in order to develop an understanding of what works, with whom, and where.

Complicating these issues are questions related to contextualization: To what extent are the constructs developed in the West, often in the US or UK, universal? To what extent do such constructs capture the understanding of PSS/SEL that is important to children, families, and societies in non-Western contexts? To what extent is understanding of self-efficacy or prosocial behavior universal? Research from developing contexts suggests that traditional agricultural societies put greater value on group responsibility and solidarity than on individual traits (Jukes et al. 2018; Jeong 2019), yet within such societies, teachers and parents may attach different values to different attributes. Parents may focus on respect for authority, whereas teachers may emphasize curiosity and other values that lend themselves to new ways of thinking that are needed in the larger world (see, e.g., Jukes et al. 2018). On the other hand, non-Western researchers such as Kagitcibasi (2005) criticize a simple collectivist versus individualistic understanding and thus are developing more nuanced alternative conceptualizations.

We are pleased that such theories and research from beyond the Global North are incorporated into the articles in this issue, along with questions about context and the processes of contextualization. Indeed, industrialized societies vary in how they weigh and value different SEL components. Japanese educators, for example, may put more emphasis on prosocial identification with the collective than on individual achievement.

We identify several themes running through this issue. First, a number of authors discuss the development of PSS/SEL measures to monitor and evaluate program effectiveness. Nikhit D'Sa and Allyson Krupar begin their article, "Developing and Validating the International Social and Emotional Learning Assessment: Evidence from a Pilot Test with Syrian Refugee Children in Iraq," by noting the lack of low- or no-cost instruments that can be adapted to different situations to collect data on children and young people affected by crisis and conflict, often in low-resource settings. To help remedy this situation, the authors tested the validity and reliability of the International Social and Emotional Learning Assessment (ISELA) for Syrian refugee children ages 6-12 who are living in Iraq. The ISELA's performance- and scenario-based measures of self-concept, stress management, perseverance, empathy, and conflict resolution can be used reliably by groups of assessors. Of particular note in the article is the authors' thoughtful discussion of tradeoffs between psychometric rigor and feasibility under field conditions.

In their article, "Teachers' Observations of Learners' Social and Emotional Learning: Psychometric Evidence for Program Evaluation in Education in Emergencies," Ha Yeon Kim, Kalina Gjicali, Zezhen Wu, and Carly Tubbs Dolan describe how they developed and tested the Teachers' Observation of Learners' Social Emotional Learning (TOOLSEL) with a sample of 3,661 displaced Syrian children enrolled in Lebanese public schools, and with those taking part in a nonformal remedial program. The TOOLSEL is a questionnaire for teachers about children's classroom behavior that is intended to assess social, emotional, behavioral, and cognitive competencies among primary school-age children in fragile, conflict-affected settings. The authors provide a detailed discussion of the feasible, reliable, and valid use of the instrument, along with cautions against its misuse. Recognizing that the TOOLSEL was developed using knowledge and tools from high-resource, nonconflict settings, the authors propose adaptations like their own as an intermediate step to take when conditions do not permit the development of SEL measures rooted in a full, participant-informed co-construction of knowledge research process.

Next, in "Creating a Tool to Measure Children's Wellbeing: A PSS Intervention in South Sudan," Moses Olayemi, Melissa Tucker, Mamour Choul, Tom Purekal, Arlene Benitez, Wendy Wheaton, and Jennifer DeBoer report on their development of an instrument to measure student wellbeing in South Sudan. They created the instrument to help evaluate the impact of a psychosocial support program offered by local teachers, who were trained to facilitate PSS activities in child-friendly spaces for 560,000 primary school children. The authors detail their process of developing and adapting the instrument, during which they drew

from questionnaires that were well-tested and widely used, but in very different contexts. They found that the three core domains identified by experts (emotional wellbeing, social wellbeing, and resilience) were important in the context of South Sudan, but with important nuances. Wellbeing, for example, was understood in terms of social relationships and individually. Resilience in particular was modified to a self-regulation factor. The authors suggest that all such instruments be tested before they are used in the field in order to assess their fit with local meanings and contexts; that they be revised and retested based on the findings; and that local leaders be engaged to play a leading role in the adaptation process.

Continuing on the measurement theme, Fernanda Soares, Nina Menezes Cunha, and Paul Frisoli discuss the development of the Wellbeing Holistic Assessment for Teachers (WHAT) tool to measure teacher wellbeing in their article, titled “How Do We Know If Teachers Are Well? The Wellbeing Holistic Assessment for Teachers Tool.” The WHAT tool, which uses self-reported data from 1,659 Salvadoran teachers, is a combination of four commonly used measures that were translated into Spanish and adapted to local conditions. The authors conclude that the WHAT tool can be used to measure the wellbeing of teachers in the Salvadoran context in terms of emotion regulation, perceived stress, emotional exhaustion, and classroom management self-efficacy; however, the instrument has not been tested for program-evaluation purposes. The authors recommend that the WHAT tool undergo “a rigorous contextual adaptation process,” including “translation, back translation, cognitive interviewing, and pilot testing,” all while bearing in mind “the normative nature of teacher wellbeing” (183).

In their article, “Evaluating the 3Cs Program for Caregivers of Young Children Affected by the Armed Conflict in Colombia,” Lina María González Ballesteros, José M. Flores, Ana María Ortiz Hoyos, Amalia Londoño Tobón, Sascha Hein, Felipe Bolívar Rincon, Oscar Gómez, and Liliana Angélica Ponguta describe the development and evaluation of a resilience wellbeing promotion intervention for caregivers of young children who are enrolled in home-based and institutional early childhood development centers in Colombia. The intervention combines several psychosocial intervention approaches, the application of community-participatory research principles, and the utilization of early childhood development settings as an entry point for implementation. An important contribution of the intervention is the combined pragmatic and applied approaches, with a contextual and theoretical framework that builds on Bronfenbrenner’s ecological model. The authors emphasize risk and protective factors at the macro and meso levels, and at the micro level in the target communities.

In “How Family Relationships Predict the Effectiveness of a Psychosocial Group Intervention among War-Affected Children,” Raija-Leena Punamäki, Kirsi Peltonen, Marwan Diab, and Samir R. Qouta focus similarly on the important role caregivers and family members play in supporting children’s wellbeing. Their study integrates systems theory, attachment theory, and resilience theory, which resulted in a more profound understanding of the different buffering effects of various family types and has important practical implications. This study seriously addresses the cultural context of Palestine to deepen our understanding of why and how psychosocial support can affect children differently, based on families’ resources, support, and emotional patterns. Despite the considerable attention given in the literature to social support and the crucial role of supportive relationships, this article describes a rare example of research that examines the moderating impact of attachment, parenting, and sibling relationships on the effectiveness of the help given children traumatized by war. It is also a rare empirical study of a family systems approach.

In our first field note, titled “Using a Participatory Approach to Create SEL Programming: The Case of *Ahlan Simsim*,” Shanna Kohn, Kim Foulds, Charlotte Cole, Mackenzie Matthews, and Laila Hussein argue for the critical importance of participatory and trauma-informed approaches to designing SEL content for children affected by conflict and trauma. They detail the processes used to create *Ahlan Simsim*, a Sesame Street television program for children in the Middle East that is designed to bring early learning to children and families affected by the Syrian crisis through the media and direct services. The bottom-up development process involved communities and local child-development experts in Iraq, Jordan, Syria, and Lebanon in refining the program focus area and creating locally relevant, trauma-informed content that draws from the SEL strategies most appropriate and effective for audiences in the Syrian response area. Through a series of steps designed to examine the social and emotional landscapes of children and their caregivers, the program developers created a framework of common emotions that children often had difficulty naming, and related coping strategies. The authors claim that the bottom-up development process, which drew from the children’s existing knowledge to avoid retraumatizing them, was essential to meeting the particular needs of these children.

The field note by Sergiy Bogdanov, Andriy Girnyk, Vira Chernobrovkina, Volodymyr Chernobrovkin, Alexander Vinogradov, Kateryna Harbar, Yuliya Kovalevskaya, Oksana Basenko, Irina Ivanyuk, Kimberly Hook, and Mike Wessells, titled “Developing a Culturally Relevant Measure of Resilience for War-Affected Adolescents in Eastern Ukraine,” describes the development and the

psychometric properties of the first measure of resilience specifically created for war-affected adolescents in Eastern Ukraine. The article describes this important new instrument for measuring resilience, and the theoretical and methodological rigor and innovation it demonstrates. The authors designed a mixed methods study that used a systematic qualitative data analysis and triangulation to identify local concepts of resilience, which subsequently informed the development of an instrument to measure resilience. This inductively developed concept of resilience in Ukraine was further operationalized through a set of questionnaire items, which were validated using exploratory structural equation modeling. The instrument is a brief, reliable, and valid measure of resilience factors on different socioecological levels. Interestingly, some of the findings also provide suggestions for interventions, such as PSS programs in Ukraine, that could build more open school ecosystems that engage parents as active actors in the education process.

In their field note, “Developing the Group Facilitation Assessment of Competencies Tool for Group-Based Mental Health and Psychosocial Support Interventions in Humanitarian and Low-Resource Settings,” Gloria A. Pedersen, Manaswi Sangraula, Pragya Shrestha, Pooja Lakshmin, Alison Schafer, Renasha Ghimire, Nagendra P. Luitel, Mark J. D. Jordans, and Brandon A. Kohrt describe the development of a tool to assess the group facilitation competencies of the individuals who provide MHPSS services. Group-based services are useful in low-resource environments and where technical expertise is limited, with the further advantage of enhancing participants’ social support, empathy, and collective problem-solving. Aimed at adult facilitators, the Group Facilitation Assessment of Competencies Tool, or GroupACT, is a structured observational tool that assesses group facilitation competencies during standardized role-plays with actor clients, or with actual clients during the delivery of group sessions. These facilitation competencies include developing and reviewing group ground rules, facilitating participation among all group members, fostering empathy among members, encouraging collaborative problem-solving, addressing barriers to attendance, time management, and ensuring group confidentiality. The authors provide suggestions on using the tool to provide group-based MHPSS services in the health, education, protection, and other humanitarian sectors.

Turning to the book reviews, in a fascinating review of *Can Big Bird Fight Terrorism? Children’s Television and Globalized Multicultural Education*, Naomi A. Moland’s provocative book based on the innovative Sesame Street initiative in Nigeria, Kate Lapham brings unusual clarity to issues the book’s author raises about the educational power of media, and television in particular; the inherent challenges of teaching multiculturalism in culturally divided societies; and the

role external actors can play in local and national issues, among others. The review and the book both provide richly nuanced discussions of matters of place, of self and other, and the role of education in identity-based conflict.

Solfrid Raknes reviewed the *NISSEM Global Briefs: Educating for the Social, the Emotional and the Sustainable*, edited by Andy Smart, Margaret Sinclair, Aaron Benavot, Jean Bernard, Colette Chabbott, S. Garnett Russell, and James Williams. Raknes points to the book's comprehensive coverage of SEL and the inspiration the volume can offer as it addresses how to educate children to have the skills needed to achieve the 17 UN Sustainable Development Goals: skills for life, skills for the 21st century, and the skills needed in a modern, unequal, globalized, and polluted world. It should be noted that NISSEM has now published two additional global briefs (see NISSEM.org).

In their timely commentary, "How the Education in Emergencies Field Can Help the United States Respond to COVID-19," Rebecca Winthrop and Helen Shwe Hadani see the COVID-19 pandemic as bringing EiE to the developed world, in particular the United States. Looking at the US experience of the pandemic and its educational response, the writers see both successes and failures in the US response, lessons for future emergencies in a world where disease pays little attention to GDP, and innovations that can be shared. Their observations provide an important commentary on the times in which this special issue is appearing.

Looking back and writing as we near the end of 2021, we note that the COVID-19 pandemic has strongly highlighted the psychosocial, social, and socioemotional nature of the school experience for children and their families. Even in the most advantaged areas of the world that are free of conflict and have had sufficient resources and infrastructure to continue schooling during the lockdowns, it is clear that children and their families depend to a far greater extent than many had thought—both socially and emotionally—on school. School closings and restrictions on public gatherings have taken a high toll on the learning and social experiences of young people everywhere, even those lucky enough to be in school or learning online. But many children are not so lucky. Pandemic-related barriers have limited almost all children's exposure to school and have kept many children—temporarily at least—out altogether. Much has rightfully been made of the pandemic's exacerbation of existing disparities in the provision of, access to, and quality of learning. Yet prior to the pandemic, children and young people in emergencies and those affected by conflict were even more likely than others to be out of school, to be in psychosocial distress, and to be in particular need of a curriculum based in PSS/SEL. All of these challenges have grown worse during the pandemic.

In these ways, this special issue underscores the critical importance of providing education in a time of global emergency and has deeply sharpened our appreciation of how reliant children, their families and communities, and the world as a whole are on education and schooling to provide psychosocial anchors, social and emotional connections and meaning, relationships, and learning.

ACKNOWLEDGMENTS

We would like to express our deep gratitude to all who made this special issue possible: Editor-in-Chief Dana Burde and her team, notably Heddy Lahmann, Senior Managing Editor, and Nathan Thompson, Deputy Managing Editor. Particular thanks go to the editorial assistants who guided us through the process of developing this issue: Samantha Colon, Deanna Pittman, and Ruqaiyah Zarook. We thank *JEiE* Editorial Board members Carine Allaf, Mark Ginsburg, Augustino Ting, Mayai, Ruth Naylor, Susan Garnett Russell, and Sweta Shah for this opportunity and for their ongoing thought leadership in building evidence and scholarship in *EiE*.

Funding from the Porticus Foundation, the LEGO Foundation, and the Open Society Foundations made the special issue possible. We are deeply grateful. The journal is hosted by New York University's Department of Applied Statistics, Social Science, and Humanities; the International Education Program; and the Center for Practice and Research at the Intersection of Information, Society, and Methodology (PRIISM). Without their critical administrative services, meeting space, library services, and more, we would be unable to do this work. Dean Brooks, Sonja Anderson, Peter Transburg, Sarah Montgomery, Lindsey Fraser, and others at INEE play essential roles in the journal's nurture and development. We thank Dody Riggs, our copy editor, and Patrick McCarthy, our designer, for their hard work, as well as our technical reviewers, Amanda Blewitt, Emily Dunlop, Haider Fancy, and João Souto Maior. We acknowledge and thank the many anonymous peer reviewers who give generously of their time and expertise, which has been especially challenging during the COVID-19 pandemic. We thank the researchers and practitioners who carried out the important work featured here, which substantially advances the psychosocial support and social and emotional learning of children and young people living in emergency contexts.

Finally, enormous thanks to our many anonymous volunteer reviewers for *JEiE* volumes 5-7, which were published from 2019 to 2021: J. Lawrence Aber, Frances Aboud, Aya Aboul Hosn, Mai Abu Moghli, Oula Abu-Amsha, Camilla Addey,

Julee Allen, Harald Bakkelund, Lesley Bartlett, Lucy Bassett, Jorge Baxter, Brenda Bell, Lyndsay Bird, Amanda Blewitt, Elizabeth Buckner, Albane Buriel, Jihae Cha, Mary Ann Chacko, Samira Chatila, Marjorie Chinen, Jaehwa Choi, Vidur Chopra, Sara Clarke-Habibi, Elisheva Cohen, Valerie de Marinis, Nada Elattar, Danielle Falk, Anna Farrell, Paul Frisoli, Sandra García, Esther Goh, Morten Greaves, Jacqueline Hayden, Anne Inger Helmen Borge, Peter Hyll-Larsen, Margaret Irving, Magdalena Janus, Sarah Kabay, Amy Kapit, Romilla Karnati, Jo Kelcey, Melissa Kelly, Kathryn Kigera, Elisabeth King, Julia Lerch, Paula Liliana Mantilla-Blanco, Naomi Moland, Christine Monaghan, Pamela Montalbano, Martha Moreno, Ashley Nemiro, Ana Nieto, Mario Novelli, Mara Ohorodnik, Julia Paulson, Gloria A. Pedersen, Benjamin Piper, Raija-Leena Punamäki, Elisabeth Rice, Kathryn Schwartz, Katie-Jay Scott, Daniel David Shephard, Michael Shipler, Carole Ann Spreen, Nicolas Stahelin, Gita Steiner-Khamsi, Nora Sveaass, Rebecca Tarlau, Beth Tuckwiller, Mari Ullmann, Eduardo Vergara, Carolina Vergara Lamarre, Rachel Wahl, Marcy Whitebook, Hakim Mohandas Amani Williams, and Travis Wright.

REFERENCES

- Bosqui, Tania Josiane, and Bassam Marshoud. 2018. "Mechanisms of Change for Interventions Aimed at Improving the Wellbeing, Mental Health and Resilience of Children and Adolescents Affected by War and Armed Conflict: A Systematic Review of Reviews." *Conflict and Health* 12 (1): 15: <https://doi.org/10.1186/s13031-018-0153-1>.
- Bub, Kristen, and Kelsey Dalrymple. 2020. *Best Practices on Effective SEL/Soft Skills Interventions in Distance Learning*. Washington, DC: USAID.
- Charlson, Fiona, Mark van Ommeren, Abraham Flaxman, Joseph Cornett, Harvey Whiteford, and Shekhar Saxena. 2019. "New WHO Prevalence Estimates of Mental Disorders in Conflict Settings: A Systematic Review and Meta-Analysis." *The Lancet* 394 (10194): 240-48. [https://doi.org/10.1016/S0140-6736\(19\)30934-1](https://doi.org/10.1016/S0140-6736(19)30934-1).
- CASEL (Collaborative for Academic, Social, and Emotional Learning). 2020a. "Reunite, Renew, and Thrive: Social and Emotional Learning (SEL) Roadmap for Reopening School." <https://casel.org/casel-gateway-sel-roadmap-for-reopening>.

- CASEL (Collaborative for Academic, Social, and Emotional Learning). 2020b. "What Is Social and Emotional Learning?" <https://casel.org/what-is-sel>.
- Goenjian, Armen K., David Walling, Alan M. Steinberg, Ida Karayan, Louis M. Najarian, and Robert Pynoos. 2005. "A Prospective Study of Posttraumatic Stress and Depressive Reactions among Treated and Untreated Adolescents 5 Years after a Catastrophic Disaster." *American Journal of Psychiatry* 162 (12): 2302-08. <https://doi.org/10.1176/appi.ajp.162.12.2302>.
- Herringa, Ryan J. 2017. "Trauma, PTSD, and the Developing Brain." *Current Psychiatry Reports* 19 (10): 69. <https://doi.org/10.1007/s11920-017-0825-3>.
- IFRC (International Federation of Red Cross and Red Crescent Societies). 2009. *Psychosocial Interventions: A Handbook*. Copenhagen: IFRC. https://pscentre.org/wp-content/uploads/2018/02/PSI-Handbook_EN_July10.pdf.
- IASC (Inter-agency Standing Committee). 2017. *IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings*. Geneva: IASC.
- INEE (Inter-agency Network on Education in Emergencies). 2016. *INEE Background Paper on Psychosocial Support and Social and Emotional Learning for Children and Youth in Emergency Settings*. New York: INEE. https://inee.org/system/files/resources/INEE_PSS-SEL_Background_Paper_ENG_v5.3.pdf.
- INEE (Inter-agency Network on Education in Emergencies). 2018. *Guidance Note: Psychosocial Support – Facilitating Psychosocial Wellbeing and Social and Emotional Learning*. New York: INEE. https://inee.org/system/files/resources/INEE_Guidance_Note_on_Psychosocial_Support_ENG_v2.pdf
- Jeong, Jissun. 2019. *Locally Valued Children's Social, Emotional, and Behavioral Competencies and Skills in Fantekwa, Ghana*. Unpublished manuscript. Washington DC: George Washington University, Graduate School of Education and Human Development.
- Jones, Stephanie. M. 2018. *Frontiers in the Science and Practice of Social-Emotional Learning in Preschools and Schools*. Cambridge, MA: Harvard Graduate School of Education. <http://helpingfamilieschange.org/wp-content/uploads/2018/04/Frontiers-in-the-science-and-practice-of-social-emotional-learning-in-preschools-and-schools-Stephanie-Jones.pdf>.

- Jones, Stephanie M., and Jennifer Kahn. 2017. *The Evidence Base for How We Learn: Supporting Students' Social, Emotional, and Academic Development. Consensus Statements of Evidence from the Council of Distinguished Scientists National Commission on Social, Emotional, and Academic Development.* Washington, DC: Aspen Institute. http://nationathope.org/wp-content/uploads/final_cds-evidence-base.pdf.
- Jukes, Matthew, Prosper Gabrieli, Nkanileka Loti Mgonda, Florentina Nsolezi, Grace Jeremiah, Jovina Tibenda, and Kristen L. Bub. 2018. "Respect is an Investment': Community Perceptions of Social and Emotional Competences in Early Childhood from Mtwara, Tanzania." *Global Education Review* 5 (2): 160-88.
- Kagiticbasi, Cigdem 2005. "Autonomy and Relatedness in Cultural Context: Implications for Self and Family." *Journal of Cross-Cultural Psychology* 36 (4): 403-22. <https://doi.org/10.1177/0022022105275959>.
- Mahoney, Joseph L., Joseph A. Durlak, and Roger P. Weissberg. 2018. "An Update on Social and Emotional Learning Outcome Research." *Phi Delta Kappan* 100 (4): 18-23. <https://doi.org/10.1177/0031721718815668>.
- MHPSS.net. 2021. "MHPSS and EiE Toolkit." <https://www.mhpss.net/toolkit/mhpss-and-eie>.
- Miller, Ken E., Mark J. D. Jordans, Wietse A. Tol, and Ananda Galappatti, Ananada. 2021. "A Call for Greater Conceptual Clarity in the Field of Mental Health and Psychosocial Support in Humanitarian Settings." *Epidemiology and Psychiatric Sciences* 30: 1-8. <https://doi.org/10.1017/S2045796020001110>.
- Moore, Kathryn, and R. Enrique Varela. 2010. "Correlates of Long-Term Posttraumatic Stress Symptoms in Children following Hurricane Katrina." *Child Psychiatry and Human Development* 41 (2): 239-50. <https://doi.org/10.1007/s10578-009-0165-6>.
- National Academies of Sciences, Engineering, and Medicine. 2018. *How People Learn II: Learners, Contexts, and Cultures.* Washington, DC: National Academies Press. <https://www.nap.edu/catalog/24783/how-people-learn-ii-learners-contexts-and-cultures>.

- OECD (Organisation for Economic Co-operation and Development). 2005. *The Definition and Selection of Key Competencies: Executive Summary*. Paris: OECD.
- Shonkoff, Jack P., and Andrew Garner. 2012. "The Lifelong Effects of Early Childhood Adversity and Toxic Stress." *Pediatrics* 129 (1): e232-46. <https://doi.org/10.1542/peds.2011-2663>.
- Smart, Andy, Margaret Sinclair, Aaron Benavot, Jean Bernard, Colette Chabbott, S. Garnett Russell, and James Williams. 2019. *NISSEM Global Briefs: Educating for the Social, the Emotional and the Sustainable*. <https://nissem.org/NGB1>.
- UNICEF (United Nations Children's Fund). 2015. *Developing Social and Emotional Learning (SEL) in Schools: Guidance for UNICEF country and regional offices*. Northampton, UK The University of Northampton, Northampton Centre for Learning Behaviour.
- USAID (United States Agency for International Development). 2019. *Social and Emotional Learning and Soft Skills USAID Policy Brief*. Washington, DC: USAID. <https://www.edu-links.org/resources/social-and-emotional-learning-and-soft-skills-usaid-education-policy-brief>.



Moving Minds Alliance works to scale-up the financing, policies, and leadership needed to effectively support young children and families affected by crisis and displacement everywhere. Originally established in 2017 by a group of philanthropic foundations, today Moving Minds Alliance is a multistakeholder partnership of 20+ organizations combining programmatic, funding, and research expertise to support prioritization of the youngest refugees and their caregivers. This publication was generously funded by Porticus, the LEGO Foundation, and the Open Society Foundations.

DEVELOPING AND VALIDATING THE INTERNATIONAL SOCIAL AND EMOTIONAL LEARNING ASSESSMENT: EVIDENCE FROM A PILOT TEST WITH SYRIAN REFUGEE CHILDREN IN IRAQ

NIKHIT D'SA AND ALLYSON KRUPAR

ABSTRACT

The growing focus on social and emotional learning for children of primary grade age in conflict-affected and fragile contexts necessitates an understanding of the effects these programs have. However, the dearth of valid and reliable measures of social and emotional learning skills in low-resource and crisis contexts has constrained the generation of this evidence. The few tools that have robust psychometric properties were developed for use in high-resource contexts; they often have usage costs, limit adaptation, and focus on adults as respondents. To address this gap, we developed the International Social and Emotional Learning Assessment (ISELA), an adaptable, cost-free, open-source, performance-based measure of self-concept, stress management, perseverance, empathy, and conflict resolution in children between ages 6 and 12. In this study, we focused on establishing the validity and reliability of the ISELA when used with Syrian refugee children in Iraq. We tested the latent structure, criterion validity, internal consistency reliability, and interrater reliability of the ISELA with 620 Syrian children. We were able to establish a theoretically grounded factor structure for all of the skills except perseverance. The ISELA can be used reliably by groups of assessors (Krippendorf's $\alpha > .86$) with strong internal consistency (KR-20 $> .70$). Our findings for criterion validity were promising but preliminary; grade and exposure to interpersonal threats demonstrated a positive association with social and emotional learning skills.

Received December 26, 2019; revised February 2, 2021; accepted August 2, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

INTRODUCTION

Forced displacement due to armed conflict can have deleterious effects on children's psychological and social development, often resulting in issues such as anxiety, depression, interpersonal violence, and posttraumatic stress disorder (Morgos, Worden, and Gupta 2007; Smith et al. 2002; Upadhyay, Srivastava, and Paul 2017). Even children who do not directly experience violence and loss due to forced displacement can be negatively affected by interpersonal aggression, lack of resources, and overcrowded living arrangements (Miller and Rasmussen 2010). Recent developments in neuropsychology reveal that children exposed to severe adversity often experience a physiological stress response that alters how their brains process information, which can adversely affect their ability to learn and thrive (Anda et al. 2011). However, there is preliminary and promising evidence that children who experience adversity can display remarkable psychosocial resilience and academic achievement when provided with opportunities for social and emotional learning (SEL), which is children's ability to understand and manage emotions, feel and show empathy, and develop positive relationships (CASEL 2015; Ungar et al. 2017; Winthrop and Kirk 2008). This has resulted in a proliferation of programs (McNatt et al. 2018) in conflict-affected and fragile states (CAFS) that focus on SEL. However, limited knowledge is available on the impact of these programs and only a few valid and reliable tools measure the SEL skills of children in CAFS. To address this gap, we developed the International Social and Emotional Learning Assessment (ISELA), an adaptable, cost-free, open-source, performance-based measure of self-concept, stress management, perseverance, empathy, and conflict resolution in children between ages 6 and 12. In this paper, we discuss the ISELA's psychometric properties when used with Syrian Kurdish refugee children living in Iraq.

SEL IN THE CONTEXT OF FORCED DISPLACEMENT

SEL programs have had a demonstrable positive impact on children's academic achievement (Durlak et al. 2011), and they also have been linked to reduced student attrition (Wang et al. 2016), conduct problems, and emotional distress in school-age children (Payton et al. 2008; Taylor et al. 2017). More notably, SEL programs are especially effective in protecting children who have experienced severe adversity and have limited access to other resources (Greenberg et al. 2017; Jones, Greenberg, and Crowley 2015). However, most SEL program impact studies have been conducted in high-resource contexts. In previous reviews and meta-analyses of SEL interventions (Durlak et al. 2011; Puerta, Valerio, and Bernal

2016; Sklad et al. 2012), only 7 percent to 12 percent of the studies were conducted outside of North America and Europe, and none focused on CAFS. This gap in our knowledge about the effects of SEL and other education programs has led the education in emergencies sector to call for more robust and rigorous research that is focused on identifying the impact these programs have on children's short- and long-term outcomes (Burde et al. 2017).

Rigorous efforts have been made recently to study the impact of SEL programs in CAFS. These studies have demonstrated three main challenges. First, they have found that the programs have a limited impact on children's SEL skills. For example, a rigorous cluster-randomized trial of the Learning to Read in a Healing Classroom program in the Democratic Republic of the Congo found that its statistically or practically meaningful impact on children's SEL skills was limited (Aber et al. 2017; Torrente et al. 2019). Even when programs have had a demonstrated impact on children's psychosocial development, the effects have been limited. An evaluation of the Writing for Recovery program in Gaza found that it had a small effect on depression symptoms but no other mental health outcomes (Lange-Nielsen et al. 2012). Second, most of the studies that found a demonstrated impact have been exploratory. For example, the evaluation of a psychoeducation program in Gaza was focused on only sixteen classrooms in four schools (Qouta et al. 2012).

Third, SEL impact studies have focused primarily on distal outcomes related to psychopathology by measuring anxiety, depression symptoms, posttraumatic stress symptoms, and psychological distress. The theory of change of most SEL programs is that these distal psychopathological outcomes are affected by children's more proximal SEL skills, such as self-awareness and perspective-taking (Torrente et al. 2019). However, because of the availability of well-validated tools that focus on children's mental health, research has focused largely on mental health outcomes and used them as a proxy for children's SEL skills. One measure that is often used in studies of children's SEL in CAFS is the Strengths and Difficulties Questionnaire (SDQ), an assessment that has been validated in several CAFS (Woerner et al. 2004). The SDQ was designed primarily as an assessment of children's mental health (Goodman and Goodman 2009) and often is used to screen children into programs of varying intensity (Goodman 1997). However, its use as a measure of children's SEL means that the research is not focused on the children's actual SEL skills; if a tool was not designed to fulfill its intended purpose, programs can make erroneous decisions about its effectiveness (D'Sa 2019b).

CHALLENGES IN MEASURING SEL

One challenge of responding to the need for further evidence on the effects of SEL programs in CAFS is the dearth of valid, reliable, and feasible measures. If researchers want to focus on measuring the SEL skills of children in CAFS, they face the reality that a majority of the SEL tools were developed in high-resource contexts. A recent review (Halle and Darling-Churchill 2016) noted that, of the 75 SEL measures included, most were developed in North America and only one-third had been used in a language other than English. Child development is mediated by local cultural and social norms (Super and Harkness 1986); thus, it differs according to the setting. Henrich and colleagues (Henrich, Heine, and Norenzayan 2010) gathered findings from across a range of disciplines to demonstrate how samples predominantly drawn from Western, educated, industrialized, rich, and democratic societies are used erroneously to make universal claims. This includes claims about the psychosocial skills targeted by SEL programs. By assuming that attributes of healthy child development look the same across diverse contexts and measuring them within these Western paradigms, we could be incorrectly pathologizing locally normative aspects of development. Or, conversely, we may be failing to identify signs of distress among children in CAFS.

An additional concern is that only a handful of SEL tools developed in high-resource contexts are valid—that is, they measure the intended construct—and reliable—they measure the same construct with consistency. Halle and Darling-Churchill (2016) found that only 6 of 75 measures they reviewed met their validity and reliability criteria. While burgeoning efforts are being made to develop a library of psychometrically rigorous SEL tools that can be used in CAFS (Ferrás, Weiss-Yagoda, and Dolan 2019), there is still a dearth of well-validated measures. This leaves researchers who are studying SEL programs in CAFS with two options. First, adapt and iteratively test the validity and reliability of a measure that was developed in a high-resource context, and attempt to align the constructs and skills with the program being evaluated. Second, develop a custom measure from scratch. Both options are challenging because they require a considerable investment of time and resources (D'Sa 2019b). Moreover, researchers who decide to adapt and iteratively test the validity and reliability of a measure that was developed in a high-resource context must also consider the fact that SEL tools that demonstrate strong validity and reliability (Halle and Darling-Churchill 2016) often require a subscription or must be purchased, and most have restrictive copyrights that do not allow users to freely contextualize the assessment.

Another important limitation is that few validated tools collect SEL information directly from children. The respondents for most well-validated measures of SEL skills for children in primary school grades are caregivers or teachers. In previous reviews of SEL measures (Denham, Ji, and Hamre 2010; Halle and Darling-Churchill 2016), all the measures identified as valid and reliable used caregivers' or teachers' responses to assess children's skills. While caregivers and teachers can provide important information about children's SEL skills, they often offer differing information. Caregivers do not have the opportunity to compare the child in their care to a large sample of children, and teachers, who often have a limited perspective on children in nonacademic settings, do not always know children well at the start of a school year when baseline or pretest data are collected (Darling-Churchill and Lippman 2016). The difficulty in obtaining valid and reliable reports from teachers is further exacerbated in CAFS, where teacher retention is particularly hard to maintain (Ring and West 2015). While child-reported measures of SEL skills do exist, they have not been established as broadly valid and reliable.

To conduct more rigorous research on the effects SEL programs have on the learning and development of children living in CAFS, we must design robust studies that use assessments that measure the proximal SEL skills of children reliably and validly. It would be especially beneficial if these tools were designed to collect data directly from children in CAFS and allowed users to adapt and contextualize the open-source tools as needed.

DEVELOPING A NEW MEASURE OF SEL SKILLS

In mid-2015, Save the Children was looking for measures to help them understand the effects its growing portfolio of SEL programs were having on primary school-age children in low-resource and crisis contexts. After reviewing the available tools and being faced with the challenges described above, the research team embarked on the process of developing an SEL measure. First, we conducted an internal review to determine which skills we should assess. This review focused not only on Save the Children's programs but also on the competency requirements in national education plans (e.g., Chirwa and Naidoo 2014; Ministry of Education and Sports 2010), guidance on developing SEL standards (e.g., Zinsler, Weissberg, and Dusenbury 2013), and the extant literature on children's SEL. This led us to identify five skills—self-concept, stress management, perseverance, empathy, and conflict resolution. We also identified other skills, such as emotion recognition, emotion regulation, and growth mindset, but we did not include these in the

measurement initiative because they were not an explicit part of Save the Children’s programs for children ages 6-12.

We subsequently reviewed the compendium of SEL measures from high-resource contexts (Denham et al. 2010; Haggerty et al. 2010) to determine possible adaptations and extensions for a new tool. Between 2016 and 2019, we tested several different types and modalities of questions for children—Likert-type scales, visual analog scales, vignettes, and performance-based items—in the Democratic Republic of the Congo, Egypt, Haiti, Iraq, Jordan, Kenya, Mexico, South Sudan, Tanzania, Syria, Thailand, and Uganda (INEE and EASEL 2020). We used three criteria during this iterative process:

- Content validity: Did the skills make sense to program staff and did they think the items satisfactorily captured the dimensions of the skills in their context?
- Internal consistency reliability: Did items measuring the same skill do so in a consistent manner?
- Feasibility: Was the measure easy to adapt and use, and were the results comprehensible for program staff?

As an example, the first version we developed (tested in Mexico and Thailand) included a mix of Likert-type questions and situation-judgment tasks. The Likert-type questions did not have strong internal consistency, and staff members noted that the children struggled with the response categories (“Not at all like me” to “A lot like me”). Indeed, there is evidence that using Likert-type questions may not be appropriate with young children, as they require extensive adaptation, testing, and pretraining (Mellor and Moore 2014; Royeen 1985). The situation-judgment tasks, which were designed to measure empathy, used Selman’s (2003) perspective-taking stage theory to rate children’s ranking of responses to stories of interpersonal conflict. However, program staff members noted that these items were hard to adapt and were not contextually relevant. This iterative design and testing done between 2016 and 2019 led to the development of the ISELA, which describes the development of five SEL skills in children ages 6 to 12: self-concept, stress management, perseverance, empathy, and conflict resolution. The ISELA was designed to be a subscription-free tool that could be adapted and used in program monitoring and evaluation efforts in low-resource and crisis contexts (D’Sa 2019a).

METHODS

Iraq hosts about 250,000 forcibly displaced Syrians, primarily those who identify as ethnically Kurdish (Durable Solutions Platform 2019). The prolonged conflict in Syria has resulted in children being exposed to severe violence and deprivation, which often has led to social, emotional, and behavioral issues. In a study of displaced Syrian children living in Iraq, about half of the children interviewed told of witnessing the violent deaths of loved ones; 90 percent of these children reported feelings of distress because of this loss (Brophy 2017). The refugee children's caregivers attributed this increase in their fearful and nervous behavior, as well as increased aggression, abuse of substances as a way of coping, and developing speech impediments, on their exposure to violent forced displacement (McDonald et al. 2017). In the latter half of 2018, in response to growing pressure to provide services for displaced children in Iraq, Save the Children started the second phase of a school-based education and child-protection project in Iraq's Dohuk, Ninewa, and Diyala provinces. Our study, which was embedded in the primary schools that were part of this project, focused on establishing the validity and reliability of the ISELA when used with Syrian refugee children living in Iraq. We answered the following research questions:

- Construct validity: How well are the observed variables predicted by the underlying SEL constructs?
- Criterion validity: To what extent are the SEL skills associated with variables that we theorize they should have a strong relationship with?
- Internal consistency reliability: How well do the items measuring each SEL skill relate to other items measuring that skill?
- Interrater reliability: What is the level of agreement in scoring items between different assessors?

RESEARCH DESIGN

Initial discussions with the program staff revealed that most of the Syrian refugee children who were to participate in this study were fluent in Syrian Kurdish but not Arabic. The Syrian Kurdish dialect these children spoke often was not written; even the youth and adult assessors had a hard time reading the script. Hence, we decided to translate the ISELA into colloquial Mesopotamian Arabic and to have 20 bilingual assessors administer the tool to children in Syrian Kurdish.

This in situ translation could have biased or altered responses if the assessors asked questions using different wording, so to limit such variations, the assessors debated the Kurdish translation during the training and reached a consensus on wording. During the data collection in February 2019, a trained team leader supervised each five-person data-collection team. This study was approved by the ethics review committee at Save the Children. The interviewers obtained informed consent from caregivers and assent from the children before data were collected.

SAMPLE

This study focused on the ten schools in Dohuk that only served Syrian refugee children. We used a root mean square error of approximation (RMSEA) sample size calculator (MacCallum, Browne, and Sugawara 1996; Preacher and Coffman 2006) to estimate the required sample, given the degrees of freedom for each level-one and level-two factor analysis model. Based on our assumptions (alpha=.05, power=.80, RMSEA for alternative distribution=.05, and RMSEA for null distribution=.02), we needed to collect data from a minimum of 500 children. In each of the ten schools, eight boys and eight girls were randomly sampled from each classroom for grades 1-4. After removing children who did not consent and because of the small class sizes in some schools (where enumerators assessed all children), we ended up with a sample of 620 children. In Table 1 we illustrate the demographic characteristics of our sample.

Table 1: Demographic Characteristics of Syrian Refugee Children from Dohuk, Iraq, Included in the ISELA Psychometric Pilot

Demographics	n=620
Female	49%
Age (average years)	9
Grade 1	26%
Grade 2	25%
Grade 3	25%
Grade 4	24%
Kurdish mother tongue	92%
Arabic mother tongue	13%
Multilingual	5%
SES (average of 9 household items)	6
Risks (average of 7)	1.2
Interpersonal threats (average of 10)	3.5

INSTRUMENT

The ISELA measures five SEL skills—self-concept, stress management, perseverance, empathy, and conflict resolution—with information drawn from six subtasks.

Self-concept is children's beliefs in their skills and abilities to meet present and future goals (Bandura 2006). In the first version of the ISELA (tested in 2016), we created a Likert-type questionnaire to measure children's growing capacity for independence and confidence in their daily activities, using Bandura's (2006) guidance for developing self-concept measures. However, these Likert-type items demonstrated weak internal consistency. Moreover, colleagues working with refugee children in Thailand and Jordan noted that their work was focused on helping children have a future vision for themselves because that was a significant challenge for the children they worked with. Given the vast literature on children's general and domain-specific self-concept and self-efficacy (Bong and Skaalvik 2003), we grounded the ISELA's measurement of self-concept on children's future orientation (Markus and Nurius 1986); that is, their vision of the "selves we imagine ourselves becoming in the future, the selves we hope to become, the selves we are afraid we may become, and the selves we fully expect we will become" (Oyserman and Fryberg 2006, 4). During the administration of the ISELA, children are asked to draw and reflect on two future selves and to identify a barrier to and support for each (six items: J3-J8). Assessors' scores reflect whether the participants can imagine a future and articulate what would support them or stop them from attaining that future.

Stress management is the conscious use of personal skills and resources to reduce the effects of chronic stress and/or acute adversity (Rutter 1981). Since most of Save the Children's SEL programs focused on teaching children explicit strategies to manage stress (like belly breathing or counting to ten), the subtask (three items: D1-D3) asks children to identify strategies they use to calm down if angry or upset. Assessors' scores reflect whether the child can accurately define up to three appropriate, nondestructive stress-management strategies. If the child cannot identify one appropriate strategy, they are not asked for additional strategies.

Perseverance refers to a child's ability to stay on task despite setbacks and the task being hard to complete (Von Culin, Tsukayama, and Duckworth 2014). In the context of SEL skills, it refers to the child's ability to motivate himself/herself to continue engaging in hard and complex social relationships. The perseverance subtask on the ISELA (four items: F1-F4) asks children to draw four increasingly

difficult geometric figures using their nondominant hand. At 20 and 40 seconds into each of the four drawing activities, the assessor asks the child if they want to stop and move on to the next subtask. Assessors' scores reflect not the accuracy of the child's drawing but their ability to persist with each drawing for 60 seconds; if a child asks to stop, they are not shown subsequent geometric figures.

Empathy is the ability to consider the perspective of other individuals, understand their emotional reactions, and integrate that into socially desirable interactions (Selman 2003). The empathy subtask (10 items: E1-E10) first asks participants to recognize sadness (E1) and anger (E6) in pictures of two children. E1 and E6 were adapted from the Assessment of Children's Emotion Skills (Schultz, Izard, and Bear 2004). The participants are then asked to describe four things (E2, E3, E7, E8) they could do to make the sad/angry child feel better—an adaptation of the International Development and Early Learning Assessment (Pisani, Borisova, and Dowd 2018)—before being asked to interpret the intentions of the child whose action (e.g., pushing or spilling water) caused the original child to feel sad/angry. The last four items (E4, E5, E9, E10) were developed to measure hostile attributional bias, which is children's tendency to attribute hostile intent to ambiguous provocations. We theorize that these four items measure conflict resolution, since children who attribute hostile intent to ambiguous provocations tend to negotiate aggressive resolutions to interpersonal conflict (Dodge et al. 2015).

Conflict resolution refers to the strategies and methods children use to peacefully negotiate interpersonal disputes with their peers (Lemerise and Arsenio 2000). Conflict resolution is assessed through children's interpretation of an interpersonal conflict vignette (four items: G1-G4), along with the four items from the empathy subtask. Participants are asked to name two things they would do if they were playing with a toy and another child asked to play with the same toy. They are then asked to name two things they would do if the other child took the toy from them without asking and started playing with it. These conflict-resolution subtask items were adapted from the International Development and Early Learning Assessment (Pisani et al. 2018) and the Challenging Situations Task (Denham et al. 2013).

Relationships focus on the individuals who are part of the child's social network. The relationship subtask does not measure a specific SEL skill, but items from this subtask are theorized to load onto the measurement of four of the five SEL skills described above. During prior conversations about the ISELA's content validity, program colleagues noted that the items from the SEL skill subtasks (described above) did not capture the skill adequately. Adding items from the

relationship subtask enabled us to broaden the measurable dimensions of these skills. In the relationship's subtask, participants are asked to describe the size of their social network, including family, peers, and community adults. For each type of individual in their social network, participants are asked if they talk to someone when they are sad (stress management), can ask for help when working on something difficult (perseverance), know when a person in their social network is feeling sad (empathy), and can ask for help with a problem with a peer (conflict resolution).

ADAPTATION OF THE ISELA

All questions in the ISELA have a binary response option: correct/incorrect or appropriate/inappropriate. Participants provide open-ended responses, and the assessors mark the responses as appropriate or inappropriate while collecting the data. This in situ coding was done for two reasons. First, it improved the feasibility of collecting data. In early versions of the ISELA, we collected actual participant responses for some questions and coded them post hoc. However, by investing in assessor training, we were able to reduce the time and resources needed to score the assessment, thereby improving feasibility. Second, given the normative nature of SEL skills, the open-ended response coding improves the adaptability of the ISELA. In each new context, the research team develops a list of socially and contextually appropriate and inappropriate response options. In Iraq, the program team from Save the Children developed the first list of appropriate and inappropriate response options for the stress-management, empathy, and conflict-resolution subtasks. The assessors, who were adults from the same community as the children, added to this list during training. After a one-day field test with children (which was not included in the final sample), the assessors refined the list.

Given the high rate of exposure to adversity among forcibly displaced Syrian children, we included questions about the following eight risk factors:

1. Have you ever had to work to earn money to support your family?
2. Have you ever had to miss school for longer than one month?
3. Have you ever had to leave your home because it was not safe?
4. Have you ever lived in a home where an adult regularly did drugs?
5. Have you ever lived in a home where people shouted or yelled?

6. Have you ever lived in a home where people pushed, slapped, or threw something?
7. Have you ever gone hungry because there was not enough food?
8. Have you ever had a family member leave home for more than six months?

Next, to capture the nature of the interpersonal threats they may have experienced, we asked children whether any of the following had occurred at their school in the previous week:

1. Did you feel afraid?
2. Did you feel afraid on your way to school?
3. Did a child tease another child?
4. Did a child leave out another child?
5. Did anyone say something mean?
6. Did children get into a fight?
7. Did anyone throw something to hurt another?
8. Did an adult scream or yell?
9. Did an adult hit or kick anyone?
10. Did an adult threaten to hurt a child?

DATA ANALYSIS

FACTOR ANALYSIS

To validate the structure of the five SEL constructs—self-concept, stress management, perseverance, empathy, and conflict resolution—we first fit an exploratory factor analysis (EFA) model (in Mplus 7.4) using a geomin rotation on a random half of the data, specifying that all observed variables were categorical. After observing the eigenvalues and fit statistics for the different EFA models

(e.g., one factor, two factor), we chose the model that had the strongest fit. We confirmed our EFA on the other random half of the data, fitting a confirmatory factor analysis model for each construct separately (weighted least squares means and variance estimator appropriate for categorical data). We observed the fit statistics to see how well the specified model fit the data (Schreiber et al. 2006).

CONVERGENT AND DISCRIMINANT VALIDITY

We used grade, experience of interpersonal threats, and exposure to risk factors to establish convergent and discriminant validity. We used grade instead of age because 10 percent of the children did not provide their age, and age and grade were strongly correlated ($r=0.75$). We hypothesized that grade would have a statistically and practically meaningful positive relationship with the five SEL skills. We also hypothesized that children with higher empathy and conflict-resolution scores would have the interpersonal skills to identify more interpersonal threats at their school. This hypothesis was based on Lemerise and Arsenio's (2000) integrated model of social information processing. They explain that the "intensity with which children experience emotions and their skill at regulating emotion will influence what is noticed and the meaning attributed to the situation" (113). Children with stronger SEL skills likely notice more of the interpersonal threats around them and encode these experiences in their memory to use when processing future social cues. Alternatively, we expected the risk factors the children faced to be negatively associated with their SEL skills (Anda et al. 2011; Betancourt et al. 2013).

To understand this convergent and discriminant validity, we fit a multilevel model for each SEL skill with grade, index of interpersonal threats, and index of risk factors as our predictors. We controlled for the differential effects of gender and household wealth and clustered the standard errors at the school level. The items we use in the ISELA artificially censor children's responses at the higher and lower ends of the scale. For example, in the empathy subtask, we asked children to recognize sadness/anger and what they would do to help another child feel better. The developmental trajectory of empathy likely extends to more foundational emotion-recognition skills, as well as more advanced third-order perspective-taking. Since the subtasks in the ISELA cannot capture this entire developmental trajectory due to time and resource constraints, we may be artificially censoring

children at the lower and higher ends of the scale for each skill. To deal with this censoring, we used a Tobit regression in fitting the multilevel models. Tobit blends ordinary least squares with a probit function to deal with the censoring of data (Stewart 2013).

INTERNAL CONSISTENCY RELIABILITY

We used the Kuder-Richardson-20 (KR-20) statistic, a variation of the Cronbach's alpha statistic designed specifically for use with dichotomous items, to understand how reliably the individual items under each ISELA construct measure that skill. The KR-20 statistic ranges from 0 to 1; values between 0.70 and 0.95 suggest strong internal consistency (Streiner 2003).

INTERRATER RELIABILITY

For the first boy and first girl from each grade assessed at each school (n=99), the assessors worked in pairs, with one assessor conducting the assessment and scoring while the other listened and scored. These data were analyzed using Krippendorff's alpha statistic, with values of 0.8-1 representing a strong level of agreement between the assessors (McHugh 2012).

RESULTS

FACTOR ANALYSIS

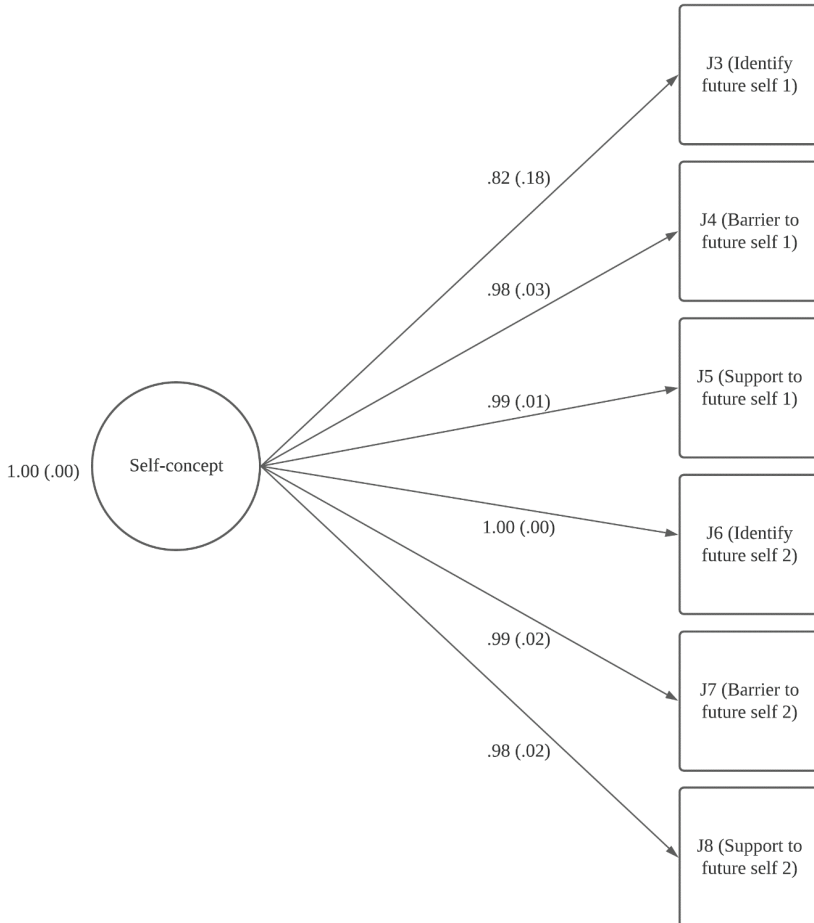
We fit a separate measurement model for each SEL skill. In Table 2 we provide an overview of the statistics we used to judge how well each model fit the data. In all five models, the comparative fit index and Tucker-Lewis index were above our acceptance threshold of 0.95 (Schreiber et al. 2006), which suggests that the models fit the data well when compared to a baseline model where all the paths were uncorrelated. The absolute fit indices—root mean square error of approximation and standardized root mean square residual—for all measurement models were lower than the prespecified thresholds of 0.06 and 0.08 (Schreiber et al. 2006), which suggests that our models fit the data well. Below we provide illustrations for each of the five measurement models and describe the structure of the loadings for each latent construct.

Table 2: Goodness of Fit Statistics from Confirmatory Factor Analysis Models Fit with Random Half of Data from Iraq

		Self-Concept	Stress Management	Perseverance	Empathy	Conflict Resolution
Split-half sample	n	309	309	309	309	309
Observed variables	x	6	4	3	5	6
Chi-square test of model fit	χ^2 (df, p-value)	8.93 (9, 0.44)	0.32 (2, 0.85)	0 (0, <0.001)	1.95 (4, 0.75)	11.93 (7, 0.10)
Comparative fit index	CFI	1	1	1	1	0.99
Tucker-Lewis index	TLI	1	1	1	1	0.97
Root mean square error of approximation	RMSEA (CI)	<0.001 (0-0.06)	<0.001 (0-0.06)	<0.001 (0-0)	<0.001 (0-0.06)	0.05 (0-0.09)
Standardized root mean square residual	SRMR	0.03	0.01	<0.001	0.02	0.06
Weighted root mean square residual	WRMR	0.22	0.12	0.01	0.23	0.56

Self-concept. In Figure 1, we illustrate the loading of the latent self-concept factor on the six observed variables from the self-concept subtask. We did not modify the model or any of its paths. We found that the latent self-concept construct predicted the six observed variables with almost equal loadings on all predicted pathways.

Figure 1: Standardized Loadings for Measurement Model for Latent Self-Concept Construct

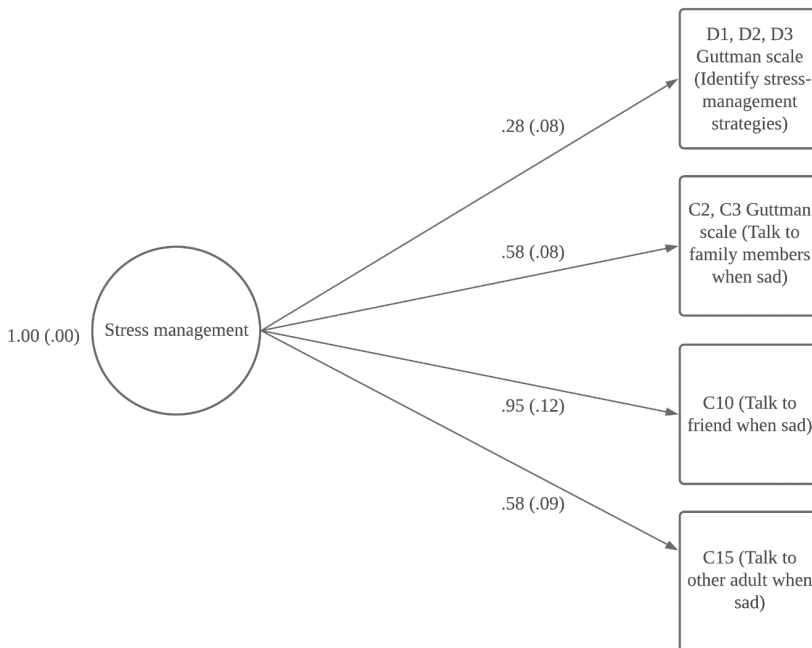


Note: All illustrated paths are statistically significant at $p < 0.05$.

Stress management. We hypothesized a single latent factor predicting items from the stress-management subtask and the relationship subtask (participants reach out to their social network when sad). We formed a Guttman scale (summed scale ranging from 0 to 3) with the three stress-management subtask questions because they were hierarchical: participants unable to identify an appropriate stress-

management strategy were not asked for subsequent strategies. Since we included these items in the same measurement model with other items that were from a different subtask and had a different response structure, we had to account for the hierarchical dependency. We also created a summed scale (range 0-2) with the first two items from the relationship questions, since participants who could not identify a family member to talk to when sad were not asked about a second family member. In our final model (see Figure 2), the stress-management latent construct predicted the four observed variables; the model fit the data well and was aligned with our initial hypothesis. The latent factor most strongly predicted participants' response to whether they talked to a friend when sad.

Figure 2: Standardized Loadings for Measurement Model for Latent Stress-Management Construct

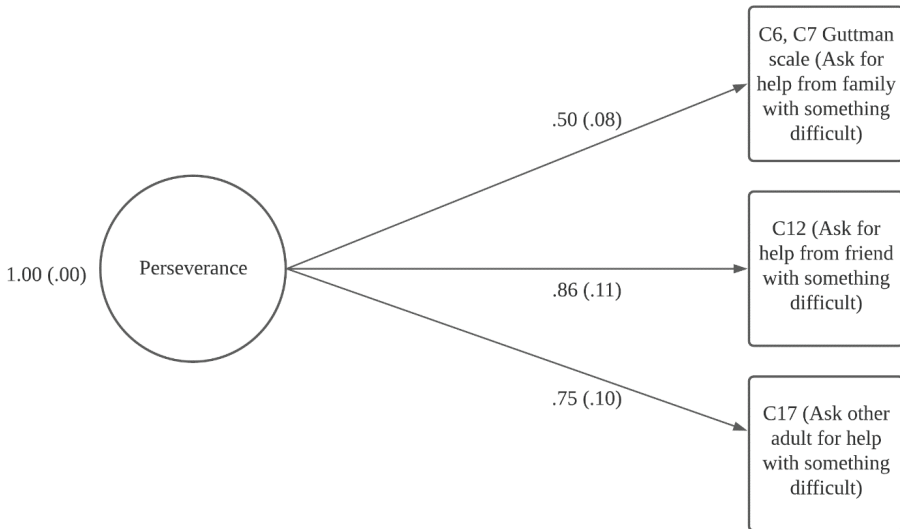


Note: All illustrated paths are statistically significant at $p < 0.05$.

Perseverance. We hypothesized that a single latent factor would predict five observed variables. First was a summed scale (range 0-4) of questions drawn from the perseverance subtask. Second was a summed scale (range 0-2) about which family members participants sought help from when working on something difficult. The final three dichotomous items included two relationship questions: whether the child approached a friend or community adult for help when working on something difficult, and whether the child chose to do the self-concept drawing.

The EFA model suggested that there were two latent factors, one predicting the three relationship items and the other predicting the drawing items from the perseverance and self-concept subtasks. We were unable to confirm this two-factor model, since one of the factors would have two observed variables and the model would not be identified. Additionally, the loadings for the two drawing items from the EFA were negative and did not fit our theory. Hence, we fit a single latent perseverance factor, predicted by the three observed relationship variables (see Figure 3). This model was fully saturated and precluded interpretation of the model fit. However, all factor loadings were strong, positive, and significant.

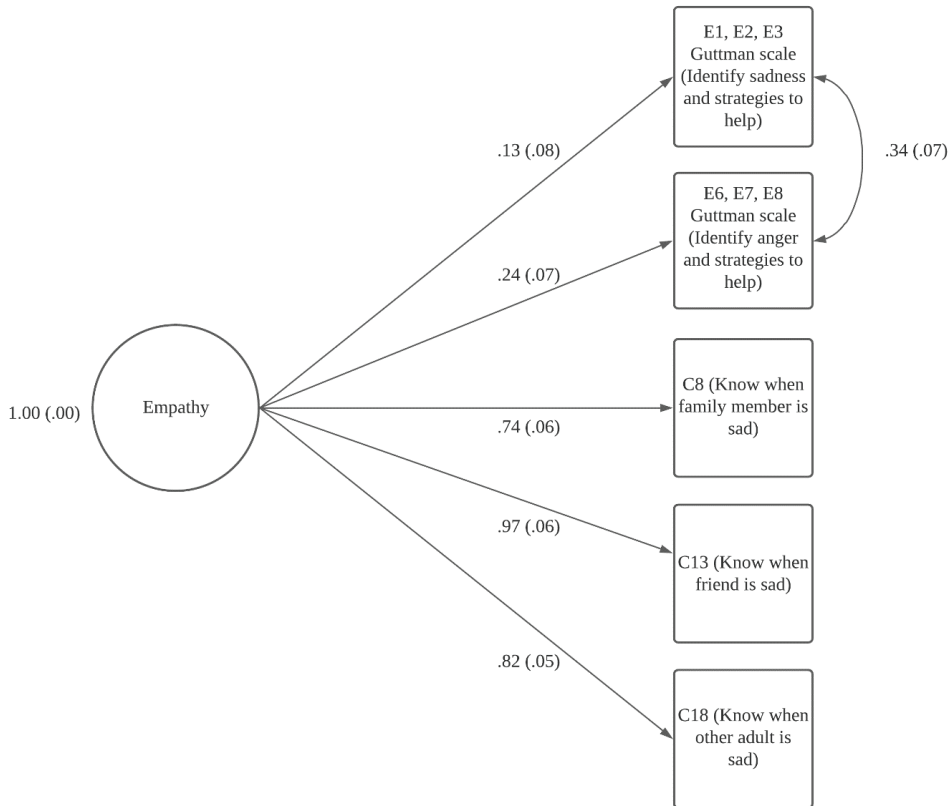
Figure 3: Standardized Loadings for Measurement Model for Latent Perseverance Construct



Note: All illustrated paths are statistically significant at $p < 0.05$.

Empathy. We theorized a single latent factor that predicted five observed variables. The first two variables were summed scales (range 0-3) from the empathy subtask: (1) recognize sadness and identify two things to do to help a sad child feel better, and (2) recognize anger and identify two things to help an angry child calm down. The last three items were from the relationship subtask: recognize when a family member, friend, or community member is sad. The final model (see Figure 4) fit the data well and was aligned with our initial theory. The one modification we made was to allow the residual variances on the two summed scales from the empathy subtask to covary. The latent factor most strongly predicted participants' response to whether or not they could identify when someone in their social network was sad.

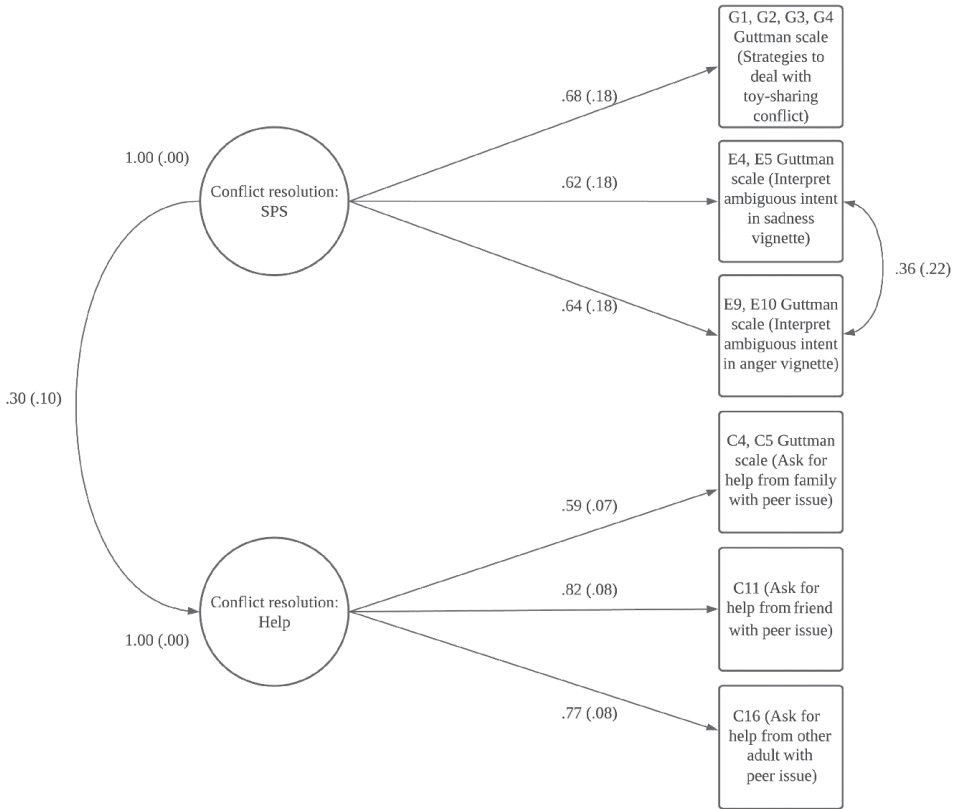
Figure 4: Standardized Loadings for Measurement Model for Latent Empathy Construct



Note: All illustrated paths are statistically significant at $p < 0.05$.

Conflict resolution. For conflict resolution, we theorized a two-factor model. The first latent factor—social problem-solving (SPS)—would be predicted by three summed-scale variables. The first (range 0-4) was a summation of the four items from the conflict-resolution subtask. The second and third were summations (range 0-2) from the empathy subtask: whether the participant attributed benign intent to ambiguous provocations in the sadness and the anger vignettes. The second latent factor—help—would be predicted by three observed variables from the relationship subtask: whether participants asked a family member, friend, or community member for help when resolving a peer issue. We expected the two latent constructs—SPS and help—to have a strong and statistically significant correlation, since they were measuring facets of conflict resolution. We made one modification to allow the residuals of the two summed variables from the empathy subtask to covary. This modified model (see Figure 5) fit the data well and all the paths were statistically and practically meaningful. The two latent factors—SPS and help—had a strong and positive correlation of 0.30.

Figure 5: Standardized Loadings for Measurement and Structural Model for Latent Conflict-Resolution Construct



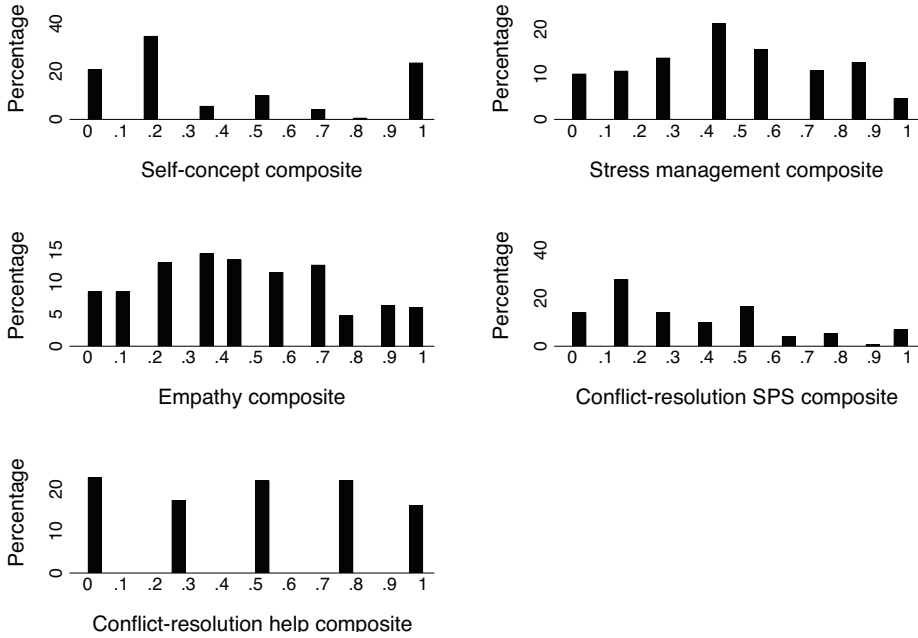
Note: All illustrated paths are statistically significant at $p < 0.05$.

CREATING COMPOSITE SCORES

Because the perseverance confirmatory factor analysis model was fully saturated, we were unable to interpret its fit and validate its structure; thus, we did not have sufficient evidence to proceed with further testing of the validity and reliability of this construct. Having established the validity of the internal structure of the other four SEL skills, the next step was to create composite scores to use in further analysis. One method was to use the confirmatory factor analysis loadings as weights for each item. However, given that the ISELA is used by teams in CAFS with limited analytic resources, we followed the more commonly applied research practice: create a sum for each skill and divide it by the total number of items. In so doing, we assume that all items are equally weighted in the composite score. We found a very strong and positive correlation between the unit-weighted and loading-weighted composites, which ranged from 0.83 for empathy to 0.99 for self-concept.

In Figure 6, we present five distribution histograms. The distributions for stress management and empathy were close to normal. The self-concept and help constructs had skewed distributions, with approximately 45 percent of participants scoring at the two extremes.

Figure 6: Distribution of Composite Scores for the SEL Skills



The five constructs had moderately strong, positive correlations with each other (see Table 3), ranging from 0.09 between self-concept and help to 0.53 between empathy and SPS. Overall, the correlations suggest that, while these constructs are related to each other, they still help capture different skills in children’s social and emotional development.

Table 3: Correlations of the Five Unit-Weighted SEL Composites

	Self-Concept	Stress Management	Empathy	Conflict Resolution: SPS	Conflict Resolution: Help
Self-concept	1				
Stress management	0.21	1			
Empathy	0.26	0.47	1		
Conflict resolution: SPS	0.35	0.35	0.53	1	
Conflict resolution: Help	0.09	0.47	0.27	0.28	1

CONVERGENT AND DISCRIMINANT VALIDITY

In Table 4 we present the results of fitting the five Tobit models to establish the convergent validity with grade and interpersonal threats, and the discriminant validity with exposure to risk factors. Our sample size was 576 children, due to missing data across the covariates in the model.

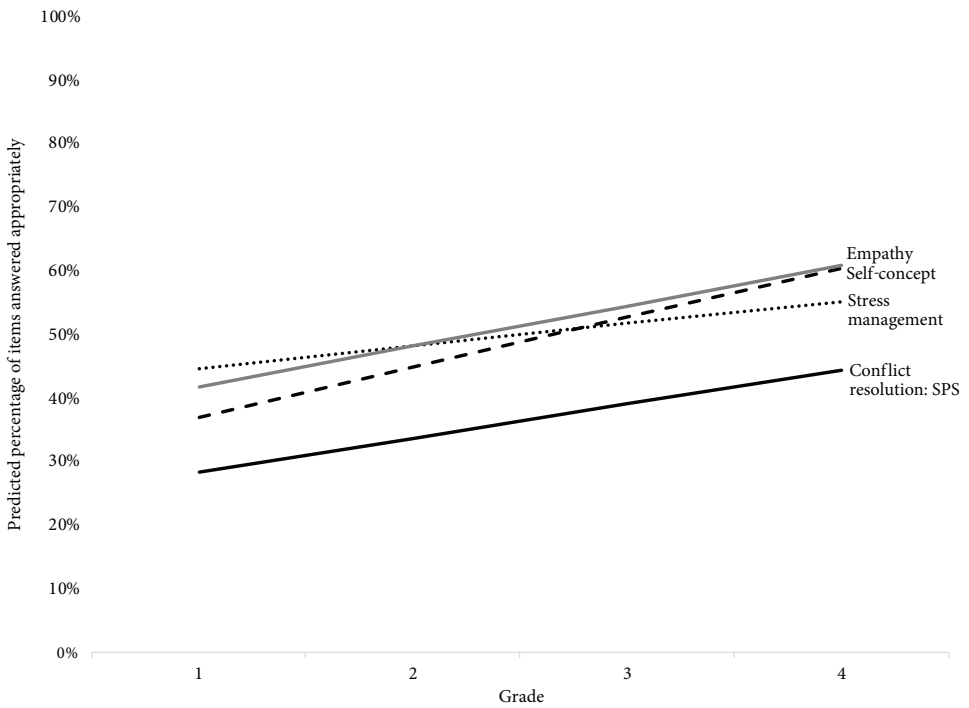
Table 4: Tobit Model Estimates for Predicted Social and Emotional Skill Scores

	Self-Concept	Stress Management	Empathy	Conflict Resolution: SPS	Conflict Resolution: Help
Grade	0.0779** [0.03]	0.0354* [0.02]	0.0635*** [0.01]	0.0532** [0.02]	0.0526 [0.03]
Gender	0.00625 [0.05]	-0.0468 [0.03]	-0.0403~ [0.02]	0.0114 [0.03]	0.00846 [0.04]
Socio-economic status	-0.173 [0.31]	0.104 [0.10]	0.230~ [0.14]	0.227* [0.11]	0.422 [0.30]
Number of risk factors	0.00186 [0.02]	-0.0111 [0.02]	0.00714 [0.02]	-0.0239 [0.02]	-0.0146 [0.02]
Number of threats	0.0228 [0.02]	0.0294*** [0.00]	0.0312*** [0.01]	0.0246** [0.01]	0.0454*** [0.01]
Intercept	0.321 [0.30]	0.276** [0.09]	0.105 [0.12]	0.0192 [0.12]	-0.0413 [0.22]
Sigma	0.612*** [0.08]	0.310*** [0.01]	0.305*** [0.02]	0.322*** [0.04]	0.537*** [0.03]
df	5	5	5	5	5
n	576	576	576	576	576

Note: Standard errors clustered at the school level. Standard errors in brackets.
p<0.10, * p<0.05, ** p<0.01, *** p<0.001

As theorized, we found a statistically and practically meaningful positive relationship between grade and four of the SEL skills; the child's grade was not predictive of their score on the help construct. However, for the other four constructs, a one-grade difference was associated with a 4- to 8-percentage-point change in the number of items children answered appropriately (see Figure 7).

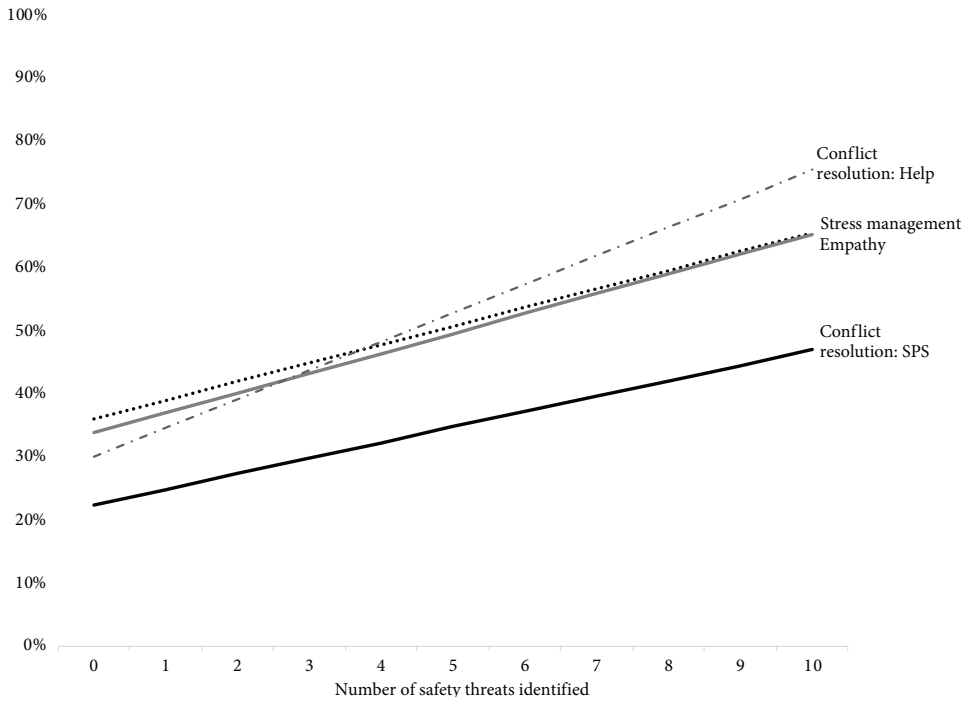
Figure 7: Predicted Social and Emotional Skill Scores by Grade, Controlling for Gender, Socioeconomic Status, Exposure to Risk Factor, and Interpersonal Threats



Note: n=576, standard errors clustered at the school level

Additionally, as initially theorized, children’s self-reported perception of interpersonal threats was positively associated with their empathy and conflict-resolution scores. We found one relationship that we had not initially theorized: children’s self-reported perception of interpersonal threats was positively associated with their stress-management scores. A one-unit change in the number of threats a child identified was positively associated with a 2- to 5-percentage-point change in their SEL skills (see Figure 8).

Figure 8: Predicted Social and Emotional Skill Scores by Number of Identified Interpersonal Threats, Controlling for Gender, Grade, Socioeconomic Status, and Exposure to Risk Factors



Note: n=576, standard errors clustered at the school level

Finally, we predicted that there would be a negative relationship between children’s SEL skills and exposure to risk factors. However, we found no relationships between these variables when controlling for the effect of important equity factors, such as grade, gender, and household wealth.

INTERNAL CONSISTENCY RELIABILITY

Across the five SEL skills, we found good internal consistency reliability of 0.70 or higher (see Table 5). This suggests that the individual items within each SEL construct are strongly correlated and measure a similar skill.

Table 5: Internal Consistency Reliability (KR-20) for SEL Constructs

Social and Emotional Competency	KR-20
Self-concept	0.91
Stress management	0.70
Empathy	0.77
Conflict resolution: SPS	0.82
Conflict resolution: Help	0.70

INTERRATER RELIABILITY

The Krippendorf’s alpha statistic for each observed variable included in the five SEL constructs was strong, above our predetermined acceptance threshold of 0.80 (see Table 6). This suggests that different assessors were administering the ISELA consistently and reliably during the data collection.

Table 6: Krippendorf’s Alpha Interrater Reliability Statistics for Observed Variables Included in Each of the SEL Constructs

Social and Emotional Competency	Krippendorf’s Alpha Range for Observed Variables
Self-concept	0.86-0.98
Stress management	0.96-1
Empathy	0.89-1
Conflict resolution: SPS	0.94-1
Conflict resolution: Help	0.86-0.96

DISCUSSION

The proliferation of SEL programs in CAFS necessitates a fuller understanding of the impact these programs have on children. However, attempts to build this knowledge have been stymied by the dearth of valid, reliable, and feasible measures that are open source and subscription free. In this article, we provide strong, positive evidence for the construct validity, internal consistency reliability, and interrater reliability of the ISELA when used with Syrian refugee children living in Iraq. We also have promising preliminary evidence for the criterion validity of the SEL skills we measured. However, designing a measure for use in CAFS comes with several tradeoffs related to assessment feasibility versus psychometric rigor, each tradeoff having a different degree of impact on the usability or rigor of the tool. Because of the resource and logistical limitations of collecting data in nonpermissive, politically volatile, or geographically challenging locations in

CAFS (Anderson, Read, and Losada 2020; Halman 2019), we need assessments that are brief, easy to use, and require limited materials or stimuli. Below we discuss the findings from this study in light of the feasibility-psychometric rigor tradeoffs we made and focus on implications for future SEL measurement in CAFS.

A challenge of measuring SEL skills is to ensure that we capture the dimensions of individual skills (Halle and Darling-Churchill 2016). For example, children's judgment about interpersonal conflict resolution is affected by several factors, including the intent they attribute to the other child, their interpersonal resolution strategies, and their help-seeking behavior (Lemerise and Arsenio 2000; Selman 2003). We attempted to capture different skill dimensions in the ISELA by ensuring that items loading on each skill construct came not only from the skill subtasks but also from the relationship subtask. This decision came with an inherent tradeoff. We increased the assessment time (about 25 minutes per child), a decision that meant we had to make cuts to other parts of the assessment. Nonetheless, we were able to establish content validity with partners who were using the tool, as well as the factor structure of self-concept, stress management, empathy, and conflict resolution.

However, we were unable to validate the factor structure for perseverance. One issue with the perseverance subtask was that 79 percent of participants completed all four drawings. This meant that there was little variation in participants' performance. This was a symptom of the tradeoff we made: by adding the relationship subtask, we shortened other ISELA subtasks; we wanted a short, performance-based measure of perseverance that was easy to use with few stimuli and materials. To measure perseverance most precisely, we need to include more complex and time-intensive items that capture the diversity of children's capacity for perseverance. In future iterations of the ISELA, we intend to work on further validating the factor structure of the perseverance latent construct by testing adaptations, including harder geometric figures and increased task complexity (e.g., not being able to lift the pencil off the paper).

The assessment-brevity tradeoff affected all sections of the ISELA, which resulted in our being judicious about the number of subtask items and the complexity of child response options. This raised a statistical challenge when trying to ensure that different items measuring each skill demonstrated strong internal consistency. All of the variables in the ISELA have a dichotomous construction and the number of items measuring each skill is small, between 4 and 12. This was done to make the assessment brief and easy to use in CAFS by limiting the need to explain complex response options to the children. Variables with larger scoring ranges and constructs with more items from the same subset of the survey provide stronger

internal consistency reliabilities (Bentler 2009; Streiner 2003). However, our analysis demonstrated that, even with these limitations, the items measuring each of the SEL skills in the ISELA demonstrated strong internal consistency reliability. Our findings suggest that measuring children's SEL skills in CAFS does not require sacrificing a measure's internal consistency. An iterative development process, like that we followed with the ISELA, can result in measures that are brief and reliable.

Another tradeoff was to remove the off-site scoring of children's responses to vignettes and scenarios; instead, we created an in situ protocol for assessors to score children's responses as socially and contextually appropriate or inappropriate. These interpretive response categories add to the ISELA's adaptability. This measure adaptability is important in CAFS because severe adversity can affect children's normative pathways of development (Betancourt et al. 2013, 2017), which requires using measures that can be changed quickly and easily to meet contextual needs. However, the adaptability of in situ scoring adds structured bias on the part of the assessor. Nonetheless, we found strong interrater reliability for all the ISELA items, which suggests that assessors had strong levels of agreement when scoring the same child. Generating in-depth response options and rigorous training can prepare assessors to score children reliably. In Iraq, we trained the assessors for three full days and conducted a field test on day four. We would recommend a similar investment in assessor training for future uses of the ISELA.

The final tradeoff we made was not with the ISELA but with the process of testing it in Iraq. Because of time and resource constraints, we were unable to include a validated measure of the children's academic or psychosocial development to establish criterion validity. This limited the scope and depth of the validity testing we could accomplish. We instead used grade (as a proxy for age because of missing data) and exposure to interpersonal threats to establish convergent validity. While we were able to provide promising evidence for the convergent validity of self-concept, stress management, empathy, and conflict resolution, these results are preliminary and need further investigation. Additionally, we were unable to establish the discriminant validity of the SEL skills with an index of distal risk factors. Unlike the extant literature that has included proximal risk factors (e.g., witnessing torture) in studies of forcibly displaced children's exposure to trauma (Morgos et al. 2007; Smith et al. 2002), we decided not to include risk factors directly related to the Syrian refugee children's experiences of trauma. We wanted to have a more robust referral system in place before asking young children sensitive questions about their displacement. Also, the floor effects observed in the number of risk factors the children experienced (the average child reported 1.2 of 7 risk factors) suggests social desirability bias or that the distal risk factors were

not pertinent. Future ISELA validation efforts should include more pertinent and robust measures of academic and psychosocial development to establish criterion validity. For example, the Inter-Agency Network for Education in Emergencies' Measurement Library (on which the ISELA is available) is compiling valid, reliable learning and development measures that can be used in CAFS. Using a measure from this initiative can further our understanding of the validity of the ISELA.

CONCLUSION

Between 2016 and 2019, we embarked on an iterative design and testing process to develop a psychometrically rigorous yet feasible open-source measure of children's SEL skills. The resulting tool—the ISELA—was specifically designed with the resource and logistic constraints of CAFS in mind. The current study provides strong evidence that the ISELA measures the self-concept, stress-management, empathy, and conflict-resolution skills of Syrian refugee children living in Iraq validly and reliably. To expand the evidence based on the impact SEL programs have on children, the tool can be used reliably by groups of assessors in the context of skills-monitoring or impact evaluations in CAFS.

REFERENCES

- Aber, J. Lawrence, Catalina Torrente, Leighann Starkey, Brian Johnston, Edward Seidman, Peter Halpin, Anjuli Shivshanker, Nina Weisenhorn, Jeannie Annan, and Sharon Wolf. 2017. "Impacts after One Year of 'Healing Classroom' on Children's Reading and Math Skills in DRC: Results from a Cluster Randomized Trial." *Journal of Research on Educational Effectiveness* 10 (3): 507-29. <https://doi.org/10.1080/19345747.2016.1236160>.
- Anda, Robert F., Vincent J. Felitti, J. Douglas Bremner, John D. Walker, Charles Whitfield, Bruce D. Perry, Shanta R. Dube, and Wayne H. Giles. 2011. "The Enduring Effects of Abuse and Related Adverse Experiences in Childhood: A Convergence of Evidence from Neurobiology and Epidemiology." *European Archives of Psychiatry and Clinical Neuroscience* 256 (3): 174-86. <https://doi.org/10.1007/s00406-005-0624-4>.
- Anderson, Kate, Lindsay Read, and Elena Losada. 2020. *Academic Learning Measurement and Assessment Tools in Education in Emergencies: Identifying, Analyzing, and Mapping Tools to Global Guidance Documents*. New York: INEE. <https://inee.org/resources/academic-learning-measurement-and-assessment-tools-education-emergencies-identifying>.

- Bandura, Albert. 2006. "Guide for Constructing Self-Efficacy Scales." In *Self-Efficacy Beliefs of Adolescents*, edited by Frank Pajares and Tim Urdan, 307-37. Greenwich, CT: Information Age Publishing.
- Bentler, Peter M. 2009. "Alpha, Dimension-Free, and Model-Based Internal Consistency Reliability." *Psychometrika* 74 (1): 137-43. <https://doi.org/10.1007/s11336-008-9100-1>.
- Betancourt, Theresa S., Ryan Mcbain, Elizabeth A. Newnham, and Robert T. Brennan. 2013. "Trajectories of Internalizing Problems in War-Affected Sierra Leonean Youth: Examining Conflict and Postconflict Factors." *Child Development* 84 (2): 455-70. <https://doi.org/10.1111/j.1467-8624.2012.01861.x>.
- Betancourt, Theresa S., Elizabeth A. Newnham, Dina Birman, Robert Lee, B. Heidi Ellis, and Christopher M. Layne. 2017. "Comparing Trauma Exposure, Mental Health Needs, and Service Utilization across Clinical Samples of Refugee, Immigrant, and U.S.-Origin Children." *Journal of Traumatic Stress* 30 (3): 209-18. <https://doi.org/10.1002/jts.22186>.
- Bong, Mimi, and Einar M. Skaalvik. 2003. "Academic Self-Concept and Self-Efficacy: How Different Are They Really?" *Educational Psychology Review* 15 (1): 1-40. <https://doi.org/10.1023/A:1021302408382>.
- Brophy, Marcia. 2017. *An Unbearable Reality: The Impact of War and Displacement on Children's Mental Health in Iraq*. London: Save the Children. https://resourcecentre.savethechildren.net/node/12182/pdf/iraq_an_unbearable_reality_june_2017_1.pdf.
- Burde, Dana, Amy Kapit, Rachel L. Wahl, Ozen Guven, and Margot Iglan Skarpeteig. 2017. "Education in Emergencies: A Review of Theory and Research." *Review of Educational Research* 87 (3): 619-58. <https://doi.org/10.3102/0034654316671594>.
- CASEL. 2015. *What Is Social and Emotional Learning?* Chicago: CASEL.
- Chirwa, Grames, and Devika Naidoo. 2014. "Structural and Social Constraints in the Teaching of Life Skills for HIV/AIDS Prevention in Malawi Primary Schools." *South African Journal of Childhood Education* 4 (1): 61-79. <https://doi.org/10.4102/sajce.v4i1.67>.

- Darling-Churchill, Kristen E., and Laura Lippman. 2016. "Early Childhood Social and Emotional Development: Advancing the Field of Measurement." *Journal of Applied Developmental Psychology* 45:1-7. <https://doi.org/10.1016/j.appdev.2016.02.002>.
- Denham, Susanne A., Peter Ji, and Bridget Hamre. 2010. *Compendium of Preschool through Elementary School Social-Emotional Learning and Associated Assessment Measures*. Chicago: CASEL. <https://casel.org/wp-content/uploads/2016/06/compendium-of-preschool-through-elementary-school-social-emotional-learning-and-associated-assessment-measures.pdf>.
- Denham, Susanne A., Erin Way, Sara C. Kalb, Heather K. Warren-Khot, and Hideko H. Bassett. 2013. "Preschoolers' Social Information Processing and Early School Success: The Challenging Situations Task." *British Journal of Developmental Psychology* 31 (2): 180-97. <https://doi.org/10.1111/j.2044-835X.2012.02085.x>.
- Dodge, Kenneth A., Patrick S. Malone, Jennifer E. Lansford, Emma Sorbring, Ann T. Skinner, Sombat Tapanya, Liliana Maria Uribe Tirado et al. 2015. "Hostile Attributional Bias and Aggressive Behavior in Global Context." *Proceedings of the National Academy of Sciences* 112 (30): 9310-15. <https://doi.org/10.1073/pnas.1418572112>.
- D'Sa, Nikhit. 2019a. *International Social & Emotional Learning Assessment: Adaptation and Administration Guidance*. Washington, DC: Save the Children. https://inee.org/resources/international-social-and-emotional-learning-assessment-isela?webform_id=toolkit_resources.
- D'Sa, Nikhit. 2019b. "Adopt, Contextualize, or Adapt? Understanding the Complexities of Modifying or Developing a Measure of Children's Social and Emotional Competencies." In *NISSEM Global Briefs: Education for the Social, the Sustainable and the Emotional*, edited by Aaron Benavot, Jean Bernard, Colette Chabbott, S. Garnett Russell, Margaret Sinclair, Andy Smart, and James H. Williams, 60-71. Washington, DC: NISSEM.
- Durable Solutions Platform. 2019. *Far from Home: Future Prospects for Syrian Refugees in Iraq*. Durable Solutions Platform. <https://reliefweb.int/report/iraq/far-home-future-prospects-syrian-refugees-iraq>.

- Durlak, Joseph A., Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, and Kriston B. Schellinger. 2011. "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions." *Child Development* 82 (1): 405-32. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>.
- Ferrás, Silvia Diazgranados, Jamie Weiss-Yagoda, and Carly Tubbs Dolan. 2019. "Measuring Social-Emotional Learning Outcomes for Children in Crisis: An Education in Emergencies Open-Source Measurement Library." In *NISSEM Global Briefs: Education for the Social, the Sustainable and the Emotional*, edited by Aaron Benavot, Jean Bernard, Colette Chabbott, S. Garnett Russell, Margaret Sinclair, Andy Smart, and James H. Williams, 752-59. Washington, DC: NISSEM.
- Goodman, Anna, and Robert Goodman. 2009. "Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health." *Journal of the American Academy of Child and Adolescent Psychiatry* 48 (4): 400-03. <https://doi.org/10.1097/CHI.0b013e3181985068>.
- Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: A Research Note." *Journal of Child Psychology and Psychiatry* 38 (5): 581-86. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>.
- Greenberg, Mark T., Elene E. Domitrovich, Roger P. Weissberg, and Joseph A. Durlak. 2017. "Social and Emotional Learning as a Public Health Approach to Education." *The Future of Children* 27 (1): 13-32. <https://doi.org/10.1353/foc.2017.0001>.
- Haggerty, Kevin, Jenna Elgin, Andrew Woolley. 2010. *Social-Emotional Learning Assessment Measures for Middle School Youth*. Seattle: Raikes Foundation. <https://www.casel.org/wp-content/uploads/2016/01/DAP-Raikes-Foundation-Review-1.pdf>.
- Halle, Tamara G., and Kristen E. Darling-Churchill. 2016. "Review of Measures of Social and Emotional Development." *Journal of Applied Developmental Psychology* 45:8-18. <https://doi.org/10.1016/j.appdev.2016.02.003>.
- Halman, Paul G. 2019. "Challenges in Gathering Data after Natural Disasters and the Potential for Local Knowledge: The Case of the Nepal Earthquake 2015." *NORRAG Special Issue 02*, 28-30.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33 (2-3): 61-83. <https://doi.org/10.1017/S0140525X0999152X>.

INEE (Inter-agency Network for Education in Emergencies) and EASEL (Ecological Approaches to Social Emotional Learning) Laboratory. 2020. *SEL and PSS Measurement and Assessment Tools in Education in Emergencies: Identifying, Analyzing, and Mapping Tools to Global Guidance Documents*. New York: INEE. <https://inee.org/resources/sel-and-pss-measurement-and-assessment-tools-education-emergencies-identifying-analyzing>.

Jones, Damon E., Mark Greenberg, and Max Crowley. 2015. "Early Social-Emotional Functioning and Public Health: The Relationship Between Kindergarten Social Competence and Future Wellness." *American Journal of Public Health* (November): e1-8. <https://doi.org/10.2105/AJPH.2015.302630>.

Lange-Nielsen, Ida Ingridsdatter, Silje Kolltveit, Abdel Aziz Mousa Thabet, Atle Dyregrov, Ståle Pallesen, Tom Backer Johnsen, and Jon Christian Laberg. 2012. "Short-Term Effects of a Writing Intervention among Adolescents in Gaza." *Journal of Loss and Trauma* 17 (5): 403-22. <https://doi.org/10.1080/15325024.2011.650128>.

Lemerise, Elizabeth A., and William F. Arsenio. 2000. "An Integrated Model of Emotion Processes and Cognition in Social Information Processing." *Child Development* 71 (1): 107-18. <https://doi.org/10.1111/1467-8624.00124>.

MacCallum, Robert C., Michael W. Browne, and Hazuki M. Sugawara. 1996. "Power Analysis and Determination of Sample Size for Covariance Structure Modeling." *Psychological Methods* 1 (2): 130-49. <https://doi.org/10.1037/1082-989X.1.2.130>.

Markus, Hazel, and Paula Nurius. 1986. "Possible Selves." *American Psychologist* 41 (9): 954-69. <https://doi.org/10.1037/0003-066X.41.9.954>.

McDonald, Alun, Misty Buswell, Sonia Khush, and Marcia Brophy. 2017. *Invisible Wounds: The Impact of Six Years of War on the Mental Health of Syria's Children*. London: Save the Children. https://resourcecentre.savethechildren.net/node/10384/pdf/invisible_wounds.pdf.

McHugh, Mary L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22 (3): 276-82. <https://doi.org/10.11613/BM.2012.031>.

- McNatt, Zahirah, Neil Boothby, Mike Wessells, and Rita Lo. 2018. *Guidance on PSS Facilitating Psychosocial Wellbeing and Social and Emotional Learning*. New York: INEE. <https://inee.org/resources/inee-guidance-note-psychosocial-support>.
- Mellor, David, and Kathleen A. Moore. 2014. "The Use of Likert Scales with Children." *Journal of Pediatric Psychology* 39 (3): 369-79. <https://doi.org/10.1093/jpepsy/jst079>.
- Miller, Kenneth E., and Andrew Rasmussen. 2010. "War Exposure, Daily Stressors, and Mental Health in Conflict and Post-Conflict Settings: Bridging the Divide between Trauma-Focused and Psychosocial Frameworks." *Social Science and Medicine* 70 (1): 7-16. <https://doi.org/10.1016/j.socscimed.2009.09.029>.
- Ministry of Education and Sports. 2010. *Life Skills Curriculum for Primary School Teachers in Uganda*. Kampala: Government of Uganda. <https://healtheducationresources.unesco.org/library/documents/life-skills-curriculum-primary-school-teachers-uganda>.
- Morgos, Dorothy J., William Worden, and Leila Gupta. 2007. "Psychosocial Effects of War Experiences among Displaced Children in Southern Darfur." *Omega: Journal of Death and Dying* 56 (3): 229-53. <https://doi.org/10.2190/OM.56.3.b>.
- Oyserman, Daphna, and Stephanie Fryberg. 2006. "The Possible Selves of Diverse Adolescents: Content and Function across Gender, Race and National Origin." In *Possible Selves: Theory, Research, and Applications*, edited by Curtis Dunkel and Jennifer Kerpelman, 17-39. Hauppauge, NY: Nova Science Publishers.
- Payton, John, Roger P. Weissberg, Joseph A. Durlak, Allison B. Dymnicki, Rebecca D. Taylor, Kriston B. Schellinger, and Molly Pachan. 2008. *The Positive Impact of Social and Emotional Learning for Kindergarten to Eighth-Grade Students: Findings from Three Scientific Reviews*. Chicago: CASEL. <https://casel.org/wp-content/uploads/2016/06/the-positive-impact-of-social-and-emotional-learning-for-kindergarten-to-eighth-grade-students-technical-report.pdf>.
- Pisani, Lauren, Ivelina Borisova, and Amy Jo Dowd. 2018. "Developing and Validating the International Development and Early Learning Assessment (IDELA)." *International Journal of Educational Research* 91 (January): 1-15. <https://doi.org/10.1016/j.ijer.2018.06.007>.
- Preacher, Kristopher J., and Donna L. Coffman. 2006. "Computing Power and Minimum Sample Size for RMSEA." <http://quantpsy.org/rmsear/rmsear.htm>.

- Puerta, Maria Laura Sánchez, Alexandria Valerio, and Marcela Gutiérrez Bernal. 2016. *Taking Stock of Programs to Develop Socioemotional Skills: A Systematic Review of Program Evidence*. Washington, DC: World Bank. <https://doi.org/10.1596/978-1-4648-0872-2>.
- Qouta, Samir R., Esa Palosaari, Marwan Diab, and Raija-Leena Punamaki. 2012. "Intervention Effectiveness among War-Affected Children: A Cluster Randomized Controlled Trial on Improving Mental Health." *Journal of Traumatic Stress* 25:188-298. <https://doi.org/10.1002/jts.21707>.
- Ring, Hannah Reeves, and Amy R. West. 2015. "Teacher Retention in Refugee and Emergency Settings: The State of the Literature." *International Education Journal: Comparative Perspectives* 14 (3): 106-21.
- Royeen, Charlotte Basic. 1985. "Adaptation of Likert Scaling for Use with Children." *Occupational Therapy Journal of Research* 5 (1): 59-69. <https://doi.org/10.1177/153944928500500104>.
- Rutter, Michael. 1981. "Stress, Coping and Development: Some Issues and Some Questions." *Journal of Child Psychology and Psychiatry* 22 (4): 323-56. <https://doi.org/10.1111/j.1469-7610.1981.tb00560.x>.
- Schreiber, James B., Frances K. Stage, Jamie King, Amaury Nora, and Elizabeth A. Barlow. 2006. "Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review." *The Journal of Educational Research* 99 (6): 323-37. <https://doi.org/10.3200/JOER.99.6.323-338>.
- Schultz, David, Carroll E. Izard, and George Bear. 2004. "Children's Emotion Processing: Relations to Emotionality and Aggression." *Development and Psychopathology* 16:371-87. <https://doi.org/10.1017/S0954579404044566>.
- Selman, Robert L. 2003. *The Promotion of Social Awareness: Powerful Lessons from the Partnership of Developmental Theory and Classroom Practice*. New York: Russell Sage Foundation.
- Sklad, Marcin, Rene Diekstra, Monique De Ritter, and Jehonathan Ben. 2012. "Effectiveness of School-Based Universal Social, Emotional, and Behavioral Programs: Do They Enhance Students' Development in the Area of Skill, Behavior, and Adjustment?" *Psychology in the Schools* 49 (9): 892-909. <https://doi.org/10.1002/pits.21641>.

- Smith, Patrick, Sean Perrin, William Yule, Berima Hacam, and Rune Stuvland. 2002. "War Exposure among Children from Bosnia-Herzegovina: Psychological Adjustment in a Community Sample." *Journal of Traumatic Stress* 15 (2): 147-56. <https://doi.org/10.1023/A:1014812209051>.
- Stewart, Jay. 2013. "Tobit or Not Tobit?" *Journal of Economic and Social Measurement* 38: 263-90. <https://doi.org/10.3233/JEM-130376>.
- Streiner, David L. 2003. "Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency." *Journal of Personality Assessment* 80 (1): 99-103. https://doi.org/10.1207/S15327752JPA8001_18.
- Super, Charles M., and Sara Harkness. 1986. "The Developmental Niche: A Conceptualization at the Interface of Child and Culture." *International Journal of Behavioral Development* 9 (4): 545-69. <https://doi.org/10.1177/016502548600900409>.
- Taylor, Rebecca D., Eva Oberle, Joseph A. Durlak, and Roger P. Weissberg. 2017. "Promoting Positive Youth Development through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-up Effects." *Child Development* 88 (4): 1-16. <https://doi.org/10.1111/cdev.12864>.
- Torrente, Catalina, John Lawrence Aber, Leighann Starkey, Brian Johnston, Anjali Shivshanker, Nina Weisenhorn, Jeannie Annan, Edward Seidman, Sharon Wolf, and Carly Tubbs Dolan. 2019. "Improving Primary Education in the Democratic Republic of the Congo: End-Line Results of a Cluster-Randomized Wait-List Controlled Trial of Learning in a Healing Classroom." *Journal of Research on Educational Effectiveness* 12 (3): 413-47. <https://doi.org/10.1080/19345747.2018.1561963>.
- Ungar, Michael, Gerry Connelly, Linda Liebenberg, and Linda Theron. 2017. "How Schools Enhance the Development of Young People's Resilience." *Social Indicators Research* 145: 615-27. <https://doi.org/10.1007/s11205-017-1728-8>.
- Upadhyay, Ashish Kumar, Swati Srivastava, and Chhavi Paul. 2017. "Consequences of Forced Migration during Early Childhood on Cognitive Well-Being in Later Childhood in Andhra Pradesh, India." *International Journal of Population Studies* 3 (2): 16. <https://doi.org/10.18063/ijps.v3i2.348>.

- Von Culin, Katherine R., Eli Tsukayama, and Angela L. Duckworth. 2014. "Unpacking Grit: Motivational Correlates of Perseverance and Passion for Long-Term Goals." *The Journal of Positive Psychology* 9 (4): 306-12. <https://doi.org/10.1080/17439760.2014.898320>.
- Wang, Huan, James Chu, Prashant Loyalka, Tao Xin, Yaojiang Shi, Qinghe Qu, and Chu Yang. 2016. "Can Social-Emotional Learning Reduce School Dropout in Developing Countries?" *Journal of Policy Analysis and Management* 35 (4): 818-47. <https://doi.org/10.1002/pam.21915>.
- Winthrop, Rebecca, and Jackie Kirk. 2008. "Learning for a Bright Future: Schooling, Armed Conflict, and Children's Well-Being." *Comparative Education Review* 52 (4): 639-61. <https://doi.org/10.1086/591301>.
- Woerner, Wolfgang, Bacy Fleitlich-Bilyk, Rhonda Martinussen, Janet Fletcher, Giulietta Cucchiario, Paulo Dalgalarondo, Mariko Lui, and Rosemary Tannock. 2004. "The Strengths and Difficulties Questionnaire Overseas: Evaluations and Applications of the SDQ beyond Europe." *European Child and Adolescent Psychiatry*, Supplement 13 (2): 47-54. <https://doi.org/10.1007/s00787-004-2008-0>.
- Zinsser, Katherine M., Roger P. Weissberg, and Linda Dusenbury. 2013. *Aligning Preschool through High School Social and Emotional Learning Standards: A Critical and Doable Next Step*. Chicago: CASEL. <https://files.eric.ed.gov/fulltext/ED574866.pdf>.

TEACHERS' OBSERVATIONS OF LEARNERS' SOCIAL AND EMOTIONAL LEARNING: PSYCHOMETRIC EVIDENCE FOR PROGRAM EVALUATION IN EDUCATION IN EMERGENCIES

HA YEON KIM, KALINA GJICALI, ZEZHEN WU, AND CARLY TUBBS DOLAN

ABSTRACT

Rigorous evaluation of social and emotional learning programs requires the use of measures that provide reliable and valid information on the meaningful differences in children's social emotional skills across treatment and control groups, as well as changes over time. In contexts affected by conflict and crisis, few measures can provide the evidence required to support their use in program evaluations, which limits stakeholders' ability to determine whether a program is working, how well it is working, and for whom. The Teachers' Observation of Learners' Social Emotional Learning, known as the TOOLSEL, holds promise for addressing this gap. The TOOLSEL is a teacher-report questionnaire about children's behavior as observed in natural classroom settings. It is used to assess a set of social, emotional, behavioral, and cognitive competencies among primary school-age children in fragile, conflict-affected settings. In this article, using the data from a sample of 3,661 Syrian refugee children who were enrolled in formal Lebanese public schools and had access to a nonformal remedial support program, we report evidence on the psychometric soundness of the TOOLSEL. We provide empirical evidence of the TOOLSEL's reliability and validity, and that the TOOLSEL captured these Syrian refugee children's social and emotional learning skills in ways that were unbiased and comparable across treatment groups, gender, age, and time. We also provide recommendations for using the TOOLSEL, including ways to improve its feasibility, reliability, and validity.

Received March 14, 2020; revised November 30, 2020, and April 21, 2021; accepted August 2, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

INTRODUCTION

Diverse stakeholders are increasingly investing in the implementation of social and emotional learning (SEL) programs in humanitarian contexts (UNESCO 2018). SEL programs provide safe, predictable learning environments for conflict-affected children that can promote the social and emotional skills that are critical in bolstering their resilience, addressing risks proactively, and building competencies at scale (Betancourt et al. 2013; Burde et al. 2017; Jordans, Pigott, and Tol 2016). These skills are important developmental indices, and they promote better academic outcomes (Durlak et al. 2011), as well as labor market attainment and wellbeing over the longer term (Heckman, Stixrud, and Urzua 2006; Jones, Greenberg, and Crowley 2015). However, little rigorous research has been conducted on the impact SEL programming has on refugee children living in humanitarian contexts, which leaves a critical knowledge gap when making programmatic decisions about how to support conflict-affected children most effectively (UNESCO 2018; Bakrania et al. 2021).

Building the evidence base on SEL in humanitarian contexts requires having field-feasible measures of children's social and emotional skills that are psychometrically sound, fit for program-evaluation purposes, and appropriate for the context in which they are being implemented (Tubbs Dolan and Caires 2020). Historically, many measures of social and emotional skills have been adopted from existing tools and used "off-the-shelf" in crisis contexts, with little consideration of their intended purpose (e.g., screening test, formative assessment, program evaluation) or whether they can provide reliable, valid information about the target population and context (Tubbs Dolan 2017). However, merely translating a tool designed for a different culture and context into a new language does not guarantee that it will provide a valid measurement of SEL in a new context. At a minimum, stakeholders must assess the psychometric properties of existing measures when they are used in a new context or with a new population (AERA, APA, and NCME 2014).

This study is one attempt to generate evidence on the reliability and validity of a measure assembled from existing measures used in humanitarian contexts. The Teachers' Observation of Learners' Social and Emotional Learning, known as TOOLSEL, is designed to capture teachers' perceptions of primary school-age children's social, emotional, behavioral, and cognitive skills. It was specifically developed to evaluate an SEL program that targets these skills in nonformal education settings for Syrian refugee children living in Lebanon. In this article, we present the data we used from a large randomized controlled trial to provide psychometric evidence of TOOLSEL's effectiveness with these children.

BUILDING SOCIAL AND EMOTIONAL COMPETENCIES IN EDUCATION IN EMERGENCIES

Education programming in emergency contexts can provide children with a safe space and a structured routine that creates a sense of normalcy, as well as opportunities to develop supportive relationships and attain meaningful learning outcomes (UNESCO 2018; Davies and Talbot 2008). However, children in education in emergency (EiE) settings may enter their classrooms with psychosocial challenges stemming from their experiences of violence, forced migration, and exploitation, as well as myriad daily stressors (Betancourt et al. 2013; Burde et al. 2015), all of which can interfere with their ability to learn and to connect with their teachers and classmates (Burde et al. 2017; Kim et al. 2020). Given research suggesting that children in crisis settings may be at particular risk for difficulties with social and self-regulation skills, practitioners working in emergency contexts have targeted these skills as key components of SEL programs, such as the Better Learning Program (Shah 2017) and Five-Component SEL (Kim et al. 2021).

TOOLSEL was designed to address the need for measures that can be used in EiE classrooms to assess the status and improvement of such SEL skills reliably and validly. It captures a range of cognitive, emotional, and behavioral competencies that are hypothesized to be important for children's successful social and academic adaptation in classrooms in EiE settings, which teachers can observe through daily classroom interactions. TOOLSEL focuses specifically on capturing several important social competencies and challenges that children display in classroom environments, as well as the self-regulatory functions necessary for learning, such as executive function, and emotional and behavioral regulation. We briefly discuss these competencies below.

Classrooms—both physical classrooms in formal schools and other nonformal peer-group learning spaces—are a primary setting where many school-age children who have access to education are able to develop and maintain relationships. Research in non-EiE contexts has found that successful social adjustment—as indicated by positive social interactions such as prosocial behavior and peer acceptance—is related to concurrent and future academic outcomes (Furrer and Skinner 2003), and to social competence, emotional health, and positive school behaviors (Hartup 1996). On the other hand, social difficulties indicated by aggression, peer rejection, and victimization put children at increased risk of maladaptive social-emotional functioning, both in the present and over time (Gest, Welsh, and Domitrovich 2005; LaFontana and Cillessen 2002).

Self-regulation is another of the SEL skills relevant to and observable in classroom settings. Self-regulation involves a complex system of cognitive, emotional, and behavioral processes that inhibit or modulate children's predominant responses to stimuli, and that enable them to display more adaptive emotions and behaviors (Eisenberg, Smith, and Spinrad 2011; Rothbart and Rueda 2005). Indeed, US studies suggest that self-regulation is critical for children's ability to develop successful social relationships (Kochanska, Murray, and Harlan 2000) and academic competence (Raver et al. 2011). A recent study conducted with Syrian refugee children living in Lebanon (Kim et al. 2020), which used measures that were tested for reliability and validity with the sample, also confirmed that children's cognitive and behavioral regulation skills are predictive of their academic performance.

The cognitive aspects of self-regulation skills are often represented as executive-function skills, which refers to a broad set of cognitive capacities, including working memory (i.e., the ability to keep in mind goal-relevant information) and inhibitory control (i.e., the ability to stop oneself from performing a prepotent response; Blair and Razza 2007). Extensive research suggests that executive function is a key mechanism for children's self-regulation in school, which is foundational to their learning and school success (Hughes and Ensor 2011; Jacob and Parkinson 2015). Regulation of emotions is another aspect of self-regulation that represents the capacity to regulate one's emotions and behavior in order to produce adaptive responses to the demands of a situation (Rothbart and Rueda 2005). Evidence from non-EiE contexts suggests that regulation of emotions is related to children's academic success (Boekaerts and Pekrun 2015), and to their social competence and peer acceptance (Valiente et al. 2011). Lastly, behavioral regulation—that is, the capacity to modulate behavior to achieve a specific goal—is a third foundational skill that enables children to adjust and learn successfully in classroom settings (Duncan, McClelland, and Acock 2017).

MEASURING THE IMPACT OF SEL PROGRAMMING ON SOCIAL AND EMOTIONAL SKILLS IN EIE SETTINGS

Evaluating the impact of SEL programs on children's social and emotional skills in EiE settings requires measures that are field feasible and have strong evidence of psychometric soundness.

FIELD FEASIBILITY

Using teacher rating measures, such as TOOLSEL, in an EiE context has several advantages in terms of feasibility, including that teachers' reports (1) are based on

accumulated knowledge of a particular child in various social and academic settings over a period of time, as compared to observation-based assessments that rely on a small set of short observation sessions; (2) are less likely to be subject to social-desirability bias or be dependent on children's self-awareness skills, as compared to self-report measures that require children to reflect and respond objectively about their own thoughts, feelings, and behaviors (Van de Mortel 2008); and (3) are low cost and easy to incorporate into the platforms commonly used for monitoring and evaluation, unlike interview protocols and performance-based measures that are expensive to develop and adapt, and that require lengthy data collection on individual children. While performance- or observation-based measures hold promise for measuring task- and context-specific skills and performance (Taylor et al. 2018), the cost to develop measures and collect data that are appropriate to a particular context and population may be prohibitive.

PSYCHOMETRIC CRITERIA

For a measure to be suitable for evaluation purposes, it must meet several psychometric criteria (Tubbs Dolan and Cairns 2020). First, measures used for program-evaluation purposes must have strong evidence of coherence by consistently providing information on the unique and meaningful constructs the measures are intended to capture. Second, data from program-evaluation measures must be highly reliable, as an error in the data can attenuate the ability to determine the impact of a program (Raudenbush and Sadoff 2008). Third, data from program-evaluation measures should provide evidence that the measures function and that they capture the same SEL skills of children from different subgroups (e.g., of different gender and age groups) and over time, in order to assess differences by group and changes in the same set of skills. This criterion is known as measurement invariance. Fourth, measures developed to evaluate impact should be sensitive to program-induced change that may occur during the program. Lastly, the measure should capture the key behaviors of social, emotional, and cognitive skills by providing evidence of expected relations in terms of direction—that is, whether they are positively or negatively related—and of magnitude, relative to other theoretically related variables.

POTENTIAL CORRELATES OF TEACHER-RATED SEL SKILLS

A variety of factors beyond the skills themselves are likely to be related to teachers' ratings of children's SEL skills. These include characteristics such as age and gender, similar or related social and emotional skills, and experiences reported by other sources.

First, as children mature, they build the capacity to regulate their emotions and behavior (Cole, Michel, and Teti 1994), become aware of others' perspectives in a social situation and display more prosocial behaviors (Fabes and Eisenberg 1998), and become able to sustain their attention for longer periods of time (Lumley et al. 2002). Research suggests that children become better with age at planning their actions and controlling their impulses (Zelazo, Carlson, and Kesek 2008). As they develop (Zimmermann and Iwanski 2014), children also gradually develop adaptive emotional and behavioral regulation strategies.

Second, gender differences in social-emotional skills and behaviors are prominent across domains. A meta-analysis of gender differences in children's prosocial behavior confirms that girls generally exhibit more prosocial behavior than boys (Fabes and Eisenberg 1998). Evidence from studies with war-affected children is consistent with findings from those in non-EiE contexts, with teachers rating girls lower than boys in aggression and higher in prosocial behaviors (Elzein and Ammar 2010; Keresteš 2006). Research has found that boys tend to exhibit more problems paying attention and more disruptive behavior disorders than girls (Lumley et al. 2002). However, such differences could be blurred in cultural contexts where culture-specific beliefs, values, and gender stereotypes appear to be different (Brody 2000) and different measurement methods are considered (McRae et al. 2008).

Lastly, teachers' rating of students' SEL in classrooms is likely to be modestly correlated with similar concepts where different measures were used by different reporters. For example, social competence and prosocial behavior are expected to be negatively related to self-reports of bullying and victimization experienced in school, whereas social problems are likely to be positively correlated with victimization (Ellis et al. 2016). In addition, executive function measured using performance-based assessments would likely be related to teachers reports of children's working memory and classroom behaviors related to inhibitory control. Observer reports of behavioral regulation are also likely to be related to teachers' ratings of behavioral regulation.

While typically not highly correlated with performance- and observation-based or child self-report measures (Buckley and Krachman 2016), teachers' reports provide meaningful information, as their perception and interpretation of children's behavior can affect their interaction with the children and the children's outcomes (McKown and Weinstein 2008). Ultimately, examining the divergence and convergence of different measurement methods provides multifaceted information that is valuable in understanding children's social and emotional development in emergency contexts (De Los Reyes et al. 2015).

CURRENT STUDY

This study utilizes data collected from Syrian refugee children in nonformal education classrooms in Lebanon—a typical education setting in EiE contexts—and examines the psychometric properties of TOOLSEL, a teacher-report measure of children’s SEL, in order to provide evidence of the tool’s validity and reliability. We first provide evidence of the measure’s internal coherence by identifying unique SEL constructs captured through the nonformal education teachers’ perspectives on the TOOLSEL and report the internal consistency of the items for each construct. Then we test whether these SEL constructs are consistently measured across treatment groups, different age groups and genders, and across time (fall to spring). We next examine whether the SEL constructs differ by changes occurring during the programming period, by age, and by gender. Finally, we test the hypothesized association between the SEL constructs captured with TOOLSEL and the children’s experience of victimization at school, behavioral regulation, and executive function, which are measured using different tools.

METHODS

PARTICIPANTS

We utilize data from a sample of Syrian refugee children living in Lebanon who were enrolled in nonformal remedial support programs; the data were collected as a part of a large, randomized controlled trial. During the 2017-2018 school year, the International Rescue Committee delivered nonformal remedial tutoring programming that was infused with SEL principles to Syrian refugee children in Lebanon’s Bekaa and Akkar regions. The program was offered in community sites located close to the area where a large number of the Syrian refugees reside, either in spaces rented in buildings in urban/residential areas or in tent schools and classrooms built for the program in the informal settlement communities located in more rural areas. The parents or guardians of all participants provided written consent for their children to participate in the research. The participants included 3,661 students ages 5 to 16 ($M=9.38$, $SD=2.27$; 50% female) who were enrolled in grades 1 to 7 in Lebanese public schools; they came from 169 classrooms in the 57 community sites. At the time of the study, the children had been living in Lebanon an average of four years ($M=4.13$, $SD=1.50$), and the majority of them (86%) had not reported any interruption in their schooling. Students in 29 sites were

randomly assigned to a treatment condition, where an additional SEL intervention was implemented as a part of the tutoring programming. All programming was offered in Arabic. Data were collected in the fall at the beginning of the program (November: $n=3,254$) and at the end in the spring (May: $n=2,952$).

MEASURES

All items of each measure used in this study were translated from English into Arabic. They were adapted through rounds of iterative feedback from the International Rescue Committee's local practitioners, who were working closely with teachers and students in Lebanon to ensure an adequate linguistic, cultural, and contextual fit.

TOOLSEL

Given the scarcity of SEL measures developed locally with the Syrian refugee population, TOOLSEL is assembled from various teacher-report surveys of children's classroom behaviors that were developed and tested in the US. The TOOLSEL items are drawn from three measures: the Teacher Observation of Child Adaptation-Checklist (TOCA-C; Koth, Bradshaw, and Leaf 2009); the Social Competence Scale (SCS; Conduct Problems Prevention Research Group 1990); and the Classroom Executive Function Survey (CEFS; Jones, Bailey, and Barnes 2015). Each item was rated on a five-point Likert scale ranging from 1="Never" to 5="Almost always." See Table A1 in the Appendix for the full list of items.

TOCA-C (Koth et al. 2009) is a teacher-report checklist, originally developed in the US to assess the social adaptive classroom behaviors of first-grade students as viewed and defined by their teachers. Selected items from the Prosocial Behavior, Concentration Problems, and Disruptive Behavior subscale were included in TOOLSEL. In studies in the US (Koth et al. 2009) and Greece (Kourkounasiou and Skordilis 2014), internal consistency was high for each of the subscales, with Cronbach's alpha ranging from 0.87 to 0.97.

TOOLSEL also includes items from the Emotion Regulation subscale of the SCS, which was originally created for the Fast Track Project (Conduct Problems Prevention Research Group 1990). Lastly, eight items from the CEFS (Jones et al. 2015) were included to capture teachers' perceptions of students' executive function skills. CEFS was specifically designed to measure children's demonstrated working memory, inhibitory control, and attention skills; it has been used previously in the EiE context, including in Lebanon.

VICTIMIZATION EXPERIENCE IN PUBLIC SCHOOLS

The school victimization experience was captured via a six-item questionnaire that asked children to reflect on their experience in public schools in the previous two weeks. The questions included the four items of the Victim subscale in the Illinois Bully Scale (Espelage and Holt 2001; e.g., “Other students pick on me.” “I got hit and pushed by other students.”), and an additional two items to reflect receiving harsh treatment from adults in school; this was common among the Syrian refugee children attending the public schools, according to anecdotal reports from the partner organization field practitioners (“Teachers, school directors, or other adults in public school pinched, pulled hair, or pulled ears.” “Teachers, school directors, or other adults in public school hit me with an object such as a ruler, stick, or *tuyau* [PVC pipe].” Responses were measured on a scale of 0=“Not at all.” to 4=“Absolutely yes.” Internal consistency reliability was $\alpha=0.75$ in the fall.

PRESCHOOL SELF-REGULATION ASSESSMENT: ASSESSOR REPORT

Children’s behavior regulation was rated by assessors using a 13-item version of the Preschool Self-Regulation Assessment: Assessor Report (PSRA-AR; Smith-Donald et al. 2007) adapted for a study in Zambia (McCoy et al. 2017). The PSRA-AR was originally designed to include assessors’ ratings of each child’s behavior as displayed during the performance-based PSRA assessment (e.g., “Pays attention to instructions and demonstration.” “Remains in seat appropriately during test.”). Each item was scored on a four-point Likert scale, with higher scores indicating better behavioral regulation.

RAPID ASSESSMENT OF COGNITIVE AND EMOTIONAL REGULATION

The Rapid Assessment of Cognitive and Emotional Regulation (RACER; Ford et al. 2019) was used to assess two aspects of executive function, working memory and inhibitory control, on a random half of the current sample. RACER demonstrated good accuracy and reliability in testing in Peru (Hamoudi and Sheridan 2015), Lebanon, and Niger (Ford et al. 2018), and also was used in Ghana, Bangladesh, and Ethiopia. Working memory was measured using a Spatial Delayed Match to Sample Task (Goldman-Rakic 1996). Inhibitory control was measured using a Simon Task (Simon and Rudell 1967).

ANALYTIC APPROACH

When using a measure in a new context and with a new population, conducting an empirical assessment of the psychometric properties is a necessary first step toward developing a locally developed and/or contextualized measure (AERA, APA, and NCME 2014). To do this, we conducted the analyses described below.

All descriptive analyses for this study were conducted using Stata SE15.1, and all factor analyses were conducted using Mplus 8.3 (Muthén and Muthén 2014).¹ First, to identify the unique SEL constructs underlying the TOOLSEL items, we identified and confirmed the TOOLSEL factor structure by conducting exploratory factor analyses (EFA) and confirmatory factor analyses (CFAs) at each time point (fall and spring). All items in the measurement models were estimated using weighted least squares mean and a variance adjusted estimator with a probit-link function (Lei 2009). The following criteria were used to assess the models' goodness of fit (Hu and Bentler 1999): RMSEA<0.08; CFI/ TLI>0.9; and SRMR<0.08.

Second, to assess internal consistency, we report Cronbach's α and McDonald's ω (Hayes and Coutts 2020; McDonald 1999) of each latent factor; ω does not assume equal factor loadings (i.e., all items contribute equally to measuring the construct of interest) and therefore is a better estimate of internal consistency than the conventional α (Revelle and Zinbarg 2009). While there are no definitive and universal guidelines, α >0.70 and ω >0.80 are generally considered acceptable/highly reliable (Catalán 2019).

Third, we tested measurement invariance across the treatment and control groups, different age groups, and gender groups for each time point, and longitudinal invariance across time. Measurement invariance refers to the extent to which a set of items measures an underlying construct of interest in the same way across groups or time (Reise, Widaman, and Pugh 1993). This is done by testing the equivalence of (a) the factor structure in treatment, gender, and age groups, and across timepoints (configural invariance) to evaluate whether and to what extent the same latent constructs could be identified by the same manifest observations across groups and time points; (b) the factor loadings of the items across groups/timepoints (metric invariance) in order to test whether the psychological meanings

¹ To account for nested data structure where teachers reported on all individual children's SEL, all analyses were conducted using robust standard errors, adjusted for clustering at the classroom level. In all factor-analysis models, missing data at the item level were pairwise deleted (i.e., all available information was used from all cases) to preserve the full sample (Asparouhov and Muthén 2010).

of the measured latent constructs are equivalent across groups and time points; and (c) the item intercepts or thresholds across groups/timepoints (scalar invariance) to evaluate whether the means of different groups or observations at different time points can be compared on the same scale (Vandenberg and Lance 2000).²

Fourth, we tested hypothesized differences of the TOOLSEL constructs across treatment groups, age groups, gender groups, and assessment time (fall to spring) by comparing the intercept of the latent factors in the measurement invariance models. For example, to compare male and female students, we report intercepts of the latent factors for females in the scalar invariance model of the gender invariance analysis, where male students' mean is fixed at zero. And, lastly, we examined the extent of the measurement validity of TOOLSEL by investigating (a) the bivariate association of the TOOLSEL constructs across time; (b) the bivariate associations with other related constructs; and (c) partial correlations controlling for child demographic characteristics (age, grade, gender) using the ordinary least squares regression approach.

RESULTS

IDENTIFYING TOOLSEL CONSTRUCTS

EXPLORATORY FACTOR ANALYSIS

Given the poor model fit of the five-factor confirmatory factor analysis models that reflect the original subscales the items came from (Table B1 in the Appendix), a series of exploratory factor analyses was used to conduct an empirically based exploration of the factor structure. All 28 items were included in the initial EFA models (see Table A1 for a full list of items and the items that were removed; see descriptive statistics of all items in Table A2). A four-factor solution consisting of 23 items was chosen due to the acceptable model fit and consistent patterns in the factor structure across the fall and the spring (Table B2). A list of items for the four subscales identified from the EFA are presented in Table 1.

² The relative fit of each of these models was assessed against the configural model using criteria suggested by Chen (2007); metric invariance: $\Delta CFI < 0.01$, $\Delta RMSEA < 0.015$, $\Delta SRMR < 0.030$; scalar invariance: $\Delta CFI < 0.01$, $\Delta RMSEA < 0.015$, $\Delta SRMR < 0.010$.

Table 1: TOOLSEL Items by Subscales

Number	Construct	Item Code and Description
1	Prosocial Behavior and Academic Engagement	TOC1: In the last two weeks [your child]: Concentrates
2		TOC2: In the last two weeks [your child]: Is friendly
3		TOC3: In the last two weeks [your child]: Pays attention
4		TOC7: In the last two weeks [your child]: Works hard
5		TOC5: In the last two weeks [your child]: Is liked by classmates
6		TOC9: In the last two weeks [your child]: Shows empathy & compassion for other's feelings
7	Social Problems	TOC10: In the last two weeks [your child]: Gets angry when provoked by other children
8		TOC15: In the last two weeks [your child]: Fights
9		TOC12: In the last two weeks [your child]: Yells at others
10		TOC14: In the last two weeks [your child]: Is rejected by classmates
11		TOC20: In the last two weeks [your child]: Teases classmates
12	Working Memory Functioning	TOC21: In the last two weeks [your child]: Learns up to ability
13		CEFS1: In the last two weeks [your child]: Remembers lists or items in the correct order
14		CEFS2: In the last two weeks [your child]: Follows multiple-step instructions
15		CEFS3: In the last two weeks [your child]: Uses multiple rules to complete a task

EVIDENCE FOR TOOLSEL IN EIE PROGRAM EVALUATION

Number	Construct	Item Code and Description
16	Emotional and Behavioral Regulation	CEFS4: In the last two weeks [your child]: Waits to be called on before responding
17		SCS11: In the last two weeks [your child]: Can calm down when excited or all wound up
18		CEFS6: In the last two weeks [your child]: Transitions easily to new activities, tasks, or major parts of the day (e.g., from recess)
19		CEFS8: In the last two weeks [your child]: Uses self-control techniques
20		SCS12: In the last two weeks [your child]: Can wait in line patiently when necessary
21		CEFS9: In the last two weeks [your child]: Waits patiently for her/his turn
22		CEFS10: In the last two weeks [your child]: Uses listening skills
23		SCS18: In the last two weeks [your child]: Controls temper when there is a disagreement

Note: Full set of items included in the initial analysis is available in Appendix A. Items labeled starting with TOC are taken from TOCA-C, with original item numbers used in TOCA-C. Similarly, items labeled starting with SCS were taken from SCS Emotional Regulation Scale, and items labeled starting with CEFS were taken from CEFS.

CONFIRMATORY FACTOR ANALYSIS

CFA with the four factors extracted from the EFA was run with the fall data and then modified to include two additional residual covariances (Table 2, Figure 1). This same final model obtained from the fall was tested with the endline (spring) data and yielded a result with an acceptable model fit (Table B3). All items loaded onto their respective factors with high factor loadings at $\lambda > 0.50$. The final factor structure revealed that the TOOLSEL constructs represented a considerable departure from the original subscales. Specifically, Factor 1: Prosocial Behavior and Academic Engagement, was a combination of the positively worded items from the Prosocial Behavior and Concentration Problems subscales of TOCA-C. Factor 2: Social Problems consisted of items from the Disruptive Behavior and negatively worded items from the Prosocial Behavior subscales of the TOCA-C. Factor 3: Working Memory Functioning was composed of one item from the Concentration Problem subscale from the TOCA-C, "Learn up to ability," and three items from the CEFS that described the children's working memory capacity. Lastly, Factor 4: Emotional and Behavioral Regulation, consisted of three items from the SCS Emotion Regulation subscale and five items from CEFS that describe children's ability to inhibit impulsive behaviors and to participate in classroom activities. The final model allowed two sets of item covariance for Factor 4 for a better model fit, based on conceptual similarity: (a) items CEFS4, "Waits to be called on," and SCS11, "Can calm down when excited," and (b) items SCS12, "Can wait in line patiently," and CEFS9, "Waits patiently for turn." See Table 3 for the factor loadings of each item in both the fall and the spring. These four latent factors of teacher-reported SEL skills were highly correlated to each other, ranging from $r = -0.453$ to 0.877 in the fall, and from $r = -0.351$ to 0.889 in the spring (Figure 1).

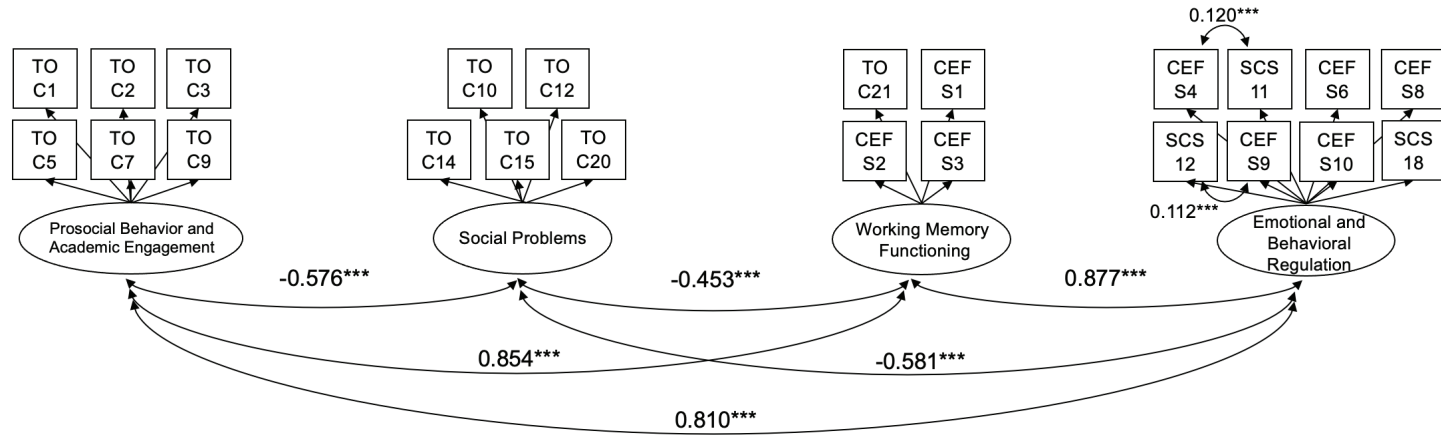
Table 2: Factor Loadings of the TOOLSEL at Fall and Spring from the Confirmatory Factor Analysis Final Model

		Fall			Spring		
		<i>b</i>	<i>SE</i>	<i>p</i>	<i>b</i>	<i>SE</i>	<i>p</i>
<i>Prosocial Behavior and Academic Engagement</i> (Fall $\alpha=0.921, =0.945$; Spring $\alpha=0.932, =0.950$)							
1	TOC1: Concentrates	0.913	0.007	0	0.932	0.006	0
2	TOC2: Is friendly	0.883	0.009	0	0.905	0.008	0
3	TOC3: Pays attention	0.903	0.008	0	0.905	0.008	0
4	TOC7: Works hard	0.778	0.014	0	0.805	0.013	0
5	TOC5: Is liked by classmates	0.882	0.009	0	0.896	0.008	0
6	TOC9: Shows empathy & compassion	0.781	0.015	0	0.805	0.014	0
<i>Social Problems</i> (Fall $\alpha=0.847, =0.900$; Spring $\alpha=0.847, =0.886$)							
1	TOC10: Gets angry when provoked	0.647	0.024	0	0.560	0.032	0
2	TOC15: Fights	0.875	0.014	0	0.879	0.014	0
3	TOC12: Yells at others	0.847	0.021	0	0.864	0.021	0
4	TOC14: Is rejected by classmates	0.892	0.014	0	0.892	0.014	0
5	TOC20: Teases classmates	0.714	0.022	0	0.737	0.020	0
<i>Working Memory Functioning</i> (Fall $\alpha=0.877, =0.909$; Spring $\alpha=0.910, =0.928$)							
1	TOC21: Learns up to ability	0.709	0.018	0	0.804	0.016	0
2	CEFS1: Remembers lists or items	0.851	0.009	0	0.899	0.009	0
3	CEFS2: Follows multistep instructions	0.901	0.009	0	0.927	0.007	0
4	CEFS3: Uses multiple rules	0.883	0.009	0	0.905	0.009	0

		Fall			Spring		
		<i>b</i>	<i>SE</i>	<i>p</i>	<i>b</i>	<i>SE</i>	<i>p</i>
<i>Emotional and Behavioral Regulation</i>							
<i>(Fall $\alpha=0.960$, $=0.972$; Spring $\alpha=0.964$, $=0.973$)</i>							
1	CEFS4: Waits to be called on	0.881	0.009	0	0.911	0.007	0
2	SCS11: Can calm down when excited	0.872	0.009	0	0.897	0.009	0
3	CEFS6: Transitions easily to new activities	0.901	0.007	0	0.925	0.006	0
4	CEFS8: Uses self-control techniques	0.915	0.006	0	0.927	0.007	0
5	SCS12: Can wait in line patiently	0.909	0.007	0	0.924	0.007	0
6	CEFS9: Waits patiently for turn	0.905	0.008	0	0.919	0.006	0
7	CEFS10: Uses listening skills	0.919	0.007	0	0.928	0.007	0
8	SCS18: Controls temper	0.864	0.01	0	0.840	0.013	0

Note: Items labeled starting with TOC are taken from TOCA-C, with original item numbers used in TOCA-C. Similarly, items labeled starting with SCS is taken from SCS Emotional Regulation Scale, and items labeled starting with CEFS were taken from CEFS.

Figure 1: Factor-Structure Diagrams Displaying Model Parameters at Fall (top) and Spring (bottom)



INTERNAL CONSISTENCY OF SUBSCALES

Table 3 also presents Cronbach's alpha estimates for scores from the empirically derived TOOLSEL subscales. All subscales have high internal reliability, ranging from $\alpha=0.85-0.96$ to $\omega=0.87-0.97$.

MEASUREMENT INVARIANCE

Using the final, empirically derived four-factor structure, we tested measurement invariance across subgroups within the sample by treatment condition, gender, and age, and across timepoints.

TREATMENT INVARIANCE

We found evidence of scalar invariance in both the fall and the spring between the treatment and control groups (see Table B4 for model fits). This means that the latent factors across two different treatment groups measure the same constructs on an equivalent scale, and therefore we can directly compare treatment and control group students on the same TOOLSEL constructs and on the same scale without bias.

GENDER AND AGE MEASUREMENT INVARIANCE

We found that TOOLSEL is scalar invariant at both waves across gender and age groups (Tables B5 and B6), which suggests that we can compare the differences by gender and age on the TOOLSEL constructs without measurement bias based on a child's gender or age.

INVARIANCE ACROSS TIME

A series of longitudinal invariance models was tested to confirm that the change from the fall to the spring for the same constructs can be estimated (Table B7). Model fit difference between configural, metric, and scalar models suggested that the factor structure, loadings, and thresholds of the items were invariant from the fall to the spring. In other words, we found no significant difference in the item and measure functioning across timepoints, thus we can compare the fall and the spring scores on these constructs.

DIFFERENCE OF SEL ACROSS GENDER, AGE, AND TIME

Table 3 and Figures 2, 3, and 4 provided differences in TOOLSEL constructs by gender, age, and time. We found significant gender differences. Girls were rated

higher than boys on all the favorable TOOLSEL constructs—Prosocial Behavior and Academic Engagement, Working Memory Functioning, Emotional and Behavioral Regulation—and lower on social problems. Interestingly, we found no statistical difference by age in the TOOLSEL constructs, despite the pattern of increase in means with age. On average, teachers reported decreased Prosocial Behavior and Academic Engagement (standardized difference=-0.106, $p<.05$) and increased Social Problems (standardized difference=0.165, $p=.001$) in the spring as compared to the fall, while they did not report a significant difference in Working Memory Functioning and Emotional and Behavioral Regulation.

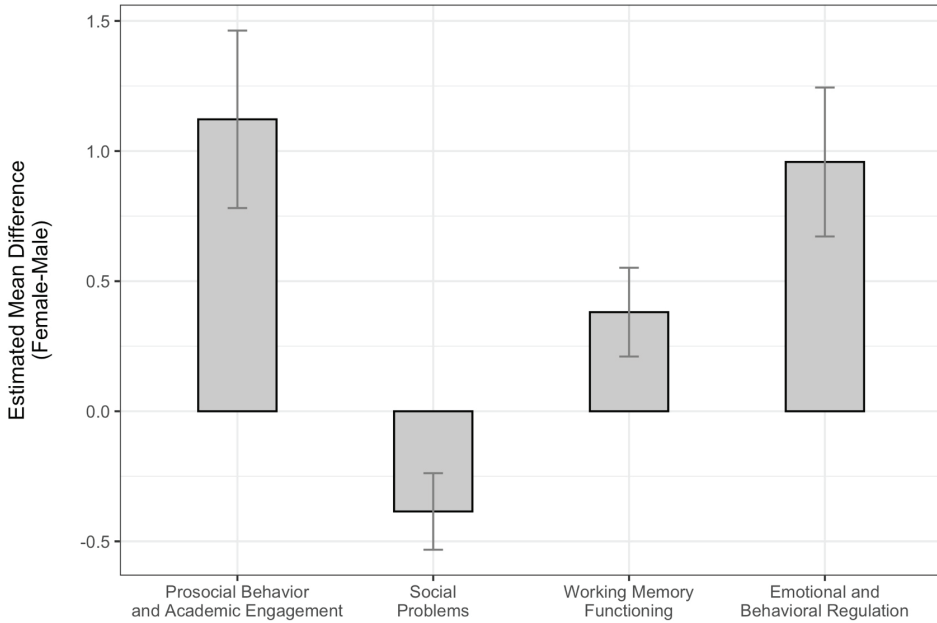
Table 3: Model-Based Estimates of TOOLSEL Subconstructs by Data Collection Wave, Age, and Gender

Estimated Latent Factor Mean (SE)				
	Prosocial Behavior and Academic Engagement	Social Problems	Working Memory Functioning	Emotional and Behavioral Regulation
Data Collection				
Fall	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)
Spring	-0.106* (0.049)	0.165** (0.055)	-0.003 (0.049)	-0.037 (0.053)
Age (years old)				
7 years or younger	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)
8-9 years	0.027 (0.234)	0.108 (0.098)	0.035 (0.121)	-0.004 (0.199)
10-11 years	0.155 (0.233)	0.082 (0.106)	0.11 (0.118)	0.052 (0.185)
≥ 12 years	0.393 (0.237)	0.074 (0.11)	0.188 (0.117)	0.123 (0.197)
Gender				
Male	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)
Female	1.122*** (0.174)	-0.385*** (0.075)	0.381*** (0.087)	0.958*** (0.146)

Note: In the fall, children age seven or younger and male were referenced for estimating means of other timepoints and subgroups in the models, and therefore fixed at a mean of 0 and variance of 1. * $p<.05$, ** $p<.01$, *** $p<.001$

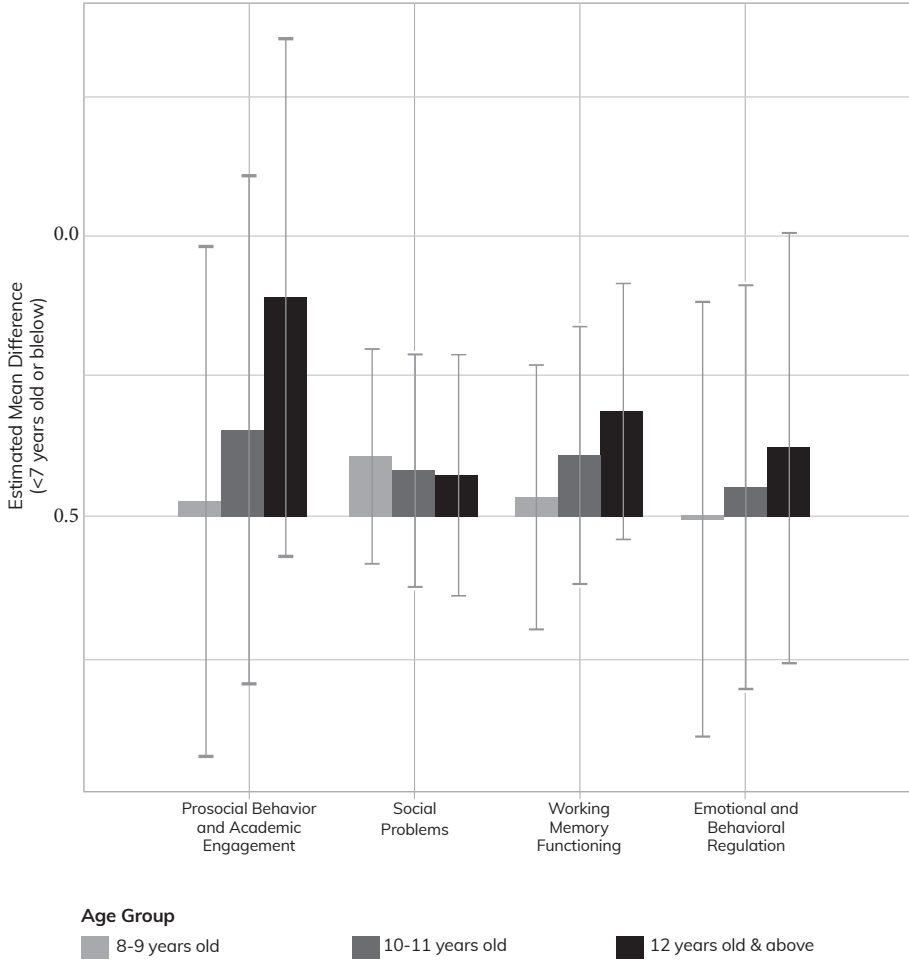
Correction: The original publication of this table in December 2021 incorrectly reported the signs of the coefficients in the row for the Spring data collection. The signs were inverted (positive to negative, and vice versa) and have been corrected in this version (August 2022). The description in the manuscript text is accurate and gives the correct direction sign of the estimates; the interpretation of the study’s findings are unaffected.

Figure 2: Gender Differences in TOOLSEL Constructs: (1) Prosocial Behavior and Academic Engagement, (2) Social Problems, (3) Working Memory Functioning, and (4) Emotional and Behavioral Regulation



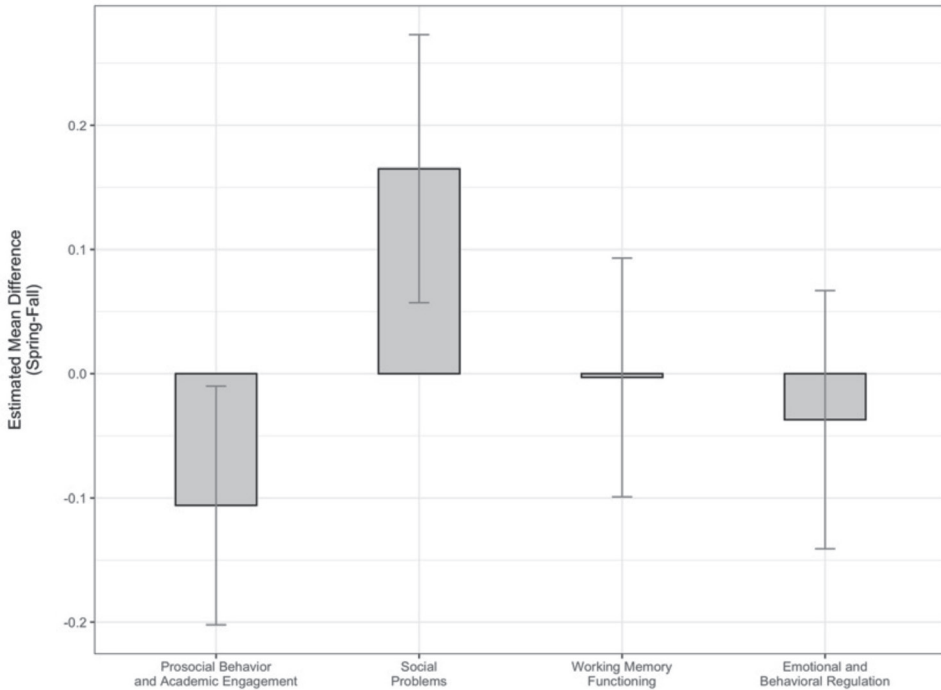
Note: Male is the reference group in estimating the mean difference.

Figure 3: Age Differences in TOOLSEL Constructs: (1) Prosocial Behavior and Academic Engagement, (2) Social Problems, (3) Working Memory Functioning, and (4) Emotional and Behavioral Regulation



Note: Children aged seven or younger is the reference group in estimating mean difference.

Figure 4: Spring-Fall Differences in TOOLSEL Constructs: (1) Prosocial Behavior and Academic Engagement, (2) Social Problems, (3) Working Memory Functioning, and (4) Emotional and Behavioral Regulation



Note: Fall is the reference group in estimating mean difference.

CORRELATIONAL EVIDENCE OF VALIDITY

BIVARIATE ASSOCIATIONS ACROSS TIME: FALL TO SPRING

We expect teachers' perceptions of their children in a specific dimension to change somewhat, but generally to remain stable over the course of a school year. Bivariate correlations of the factor scores of all four of the TOOLSEL constructs were positively correlated across time points, $r=0.585$ for Prosocial Behavior and Academic Engagement, $r=0.603$ for Social Problems, $r=0.569$ for Working Memory Functioning, $r=0.510$ for Emotional and Behavioral Regulation. This indicates that teachers' perceptions of children's behavior were fairly consistent, displaying some continuity and some change across the six-month period (Table 4).

Table 4: Bivariate Correlations among TOOLSEL Factor Scores at Fall and Spring

	1	2	3	4	5	6	7
1. Prosocial Behavior and Academic Engagement T1	--						
2. Social Problems T1	-0.637	--					
3. Working Memory Functioning T1	0.905	-0.534	--				
4. Emotional and Behavioral Regulation T1	0.862	-0.640	0.911	--			
5. Prosocial Behavior and Academic Engagement T2	0.585	-0.456	0.527	0.487	--		
6. Social Problems T2	-0.364	0.603	-0.284	-0.376	-0.570	--	
7. Working Memory Functioning T2	0.572	-0.368	0.569	0.489	0.931	-0.427	--
8. Emotional and Behavioral Regulation T2	0.534	-0.449	0.501	0.510	0.882	-0.585	0.920

Note: All correlation coefficients were statistically significant at $p < .001$.

BIVARIATE ASSOCIATIONS WITH OTHER MEASURES

Bivariate correlations between the TOOLSEL constructs and other external measures (Table 6) showed additional support for validity. That is, the TOOLSEL constructs were correlated in the expected directions with external measures of similar constructs. The Prosocial Behavior and Academic Engagement factor was positively correlated with both the assessor report of behavioral regulation and the performance-based assessment of working memory ($r=0.147, p<.001$, and $r=0.152, p<.001$, respectively). In addition, it was negatively correlated with child self-reports of public school victimization ($r=-0.117, p<.001$), but not correlated with RACER inhibitory control ($r=-0.008, p>.05$). Social problems were positively correlated with child self-report of public school victimization ($r=0.144, p<.001$), as expected. However, it had a statistically significant but very small correlation with the assessor report of behavioral regulation ($r=-0.061, p<.001$), RACER working memory ($r=-0.050, p<.05$), and RACER inhibitory control ($r=0.053, p<.05$). TOOLSEL's Working Memory Functioning was positively correlated with the assessor report of behavioral regulation ($r=0.152, p<.001$) and RACER working memory ($r=0.167, p<.001$). In addition, Working Memory

Functioning was negatively correlated to a small degree ($r=-0.091$, $p<.001$) with child self-reported public school victimization and not correlated with RACER inhibitory control ($r=0.025$, $p>.05$). Emotional and Behavioral Regulation was positively correlated with assessor-report behavioral regulation ($r=0.112$, $p<.001$) and RACER working memory ($r=0.114$, $p<.001$), and negatively correlated with child self-report of school victimization ($r=-0.128$, $p<.001$). Interestingly, Emotional and Behavioral Regulation were not associated with the RACER inhibitory control.

Table 5: Bivariate Correlations between TOOLSEL Factor Scores and PSRA, RACER, and Victimization Scale in the Fall

	1	2	3	4	5	6	7
1. Prosocial Behavior and Academic Engagement T1	--						
2. Social Problems T1	-0.637***	--					
3. Working Memory Functioning T1	0.905***	-0.534***	--				
4. Emotional and Behavioral Regulation T1	0.862***	-0.640***	0.911***	--			
5. Public School Victimization	-0.117***	0.144***	-0.091***	-0.128***	--		
6. Behavioral Regulation	0.147***	-0.061***	0.152***	0.112***	-0.065***	--	
7. RACER Working Memory	0.152***	-0.050*	0.167***	0.114***	0.025	0.256***	--
8. RACER Inhibitory Control	-0.008	0.053*	0.025	-0.022	-0.050*	0.048	0.130***

Note: *** $p<0.001$, ** $p<0.01$, * $p<0.05$

PARTIAL CORRELATION

Table 6 presents the ordinary least squares regression models testing partial correlations between TOOLSEL constructs and other related constructs, controlling for child demographic characteristics (age, grade, gender). In addition to child demographic characteristics, measures of school victimization, working memory, inhibitory control, and behavioral regulation explained 9 percent to 12 percent of the variance in TOOLSEL constructs. Controlling for child characteristics and other measures, public school victimization was significantly associated with all TOOLSEL constructs. Specifically, a higher degree of victimization was related to lower Prosocial Behavior and Academic Engagement ($b=-0.156, p<.001$), lower Working Memory Functioning ($b=-0.124, p<.001$), lower Emotional and Behavioral Regulation ($b=-0.171, p<.001$), and more Social Problems ($b=0.180, p<.001$). Assessor-report behavioral regulation was positively related to Prosocial Behavior and Academic Engagement ($b=0.104, p<.01$), Working Memory Functioning ($b=0.112, p<.001$), and Emotional and Behavioral Regulation ($b=0.090, p<.01$). RACER working memory was positively associated with teacher-reported Prosocial Behavior and Academic Engagement ($b=0.222, p<.001$), Working Memory Functioning ($b=0.234, p<.001$), Emotional and Behavioral Regulation ($b=0.185, p<.001$), and negatively associated with Social Problems ($b=-0.093, p<.01$). The RACER cognitive inhibitory control measure was not related to any of the TOOLSEL constructs.

Table 6: Ordinary Least Squares Regression Models Predicting TOOLSEL Constructs

	Prosocial Behavior and Academic Engagement			Social Problems			Working Memory Functioning			Emotional and Behavioral Regulation		
	<i>Beta</i>	<i>b</i>	<i>SE</i>	<i>Beta</i>	<i>b</i>	<i>SE</i>	<i>Beta</i>	<i>b</i>	<i>SE</i>	<i>Beta</i>	<i>b</i>	<i>SE</i>
(Intercept)	0.000	-0.199	0.167	0.000	0.073	0.169	0.000	-0.174	0.173	0.000	-0.185	0.176
Public School Victimization	-0.123	-0.156***	0.031	0.149	0.180***	0.030	-0.097	-0.124***	0.032	-0.130	-0.171***	0.032
Behavioral Regulation	0.086	0.104**	0.035	-0.017	-0.019	0.033	0.092	0.112***	0.034	0.071	0.090**	0.033
RACER Working Memory	0.137	0.222***	0.045	-0.060	-0.093*	0.041	0.143	0.234***	0.047	0.110	0.185***	0.046
RACER Inhibitory Control	-0.018	-0.017	0.022	0.046	0.041	0.023	0.010	0.009	0.023	-0.021	-0.020	0.025
Age	-0.001	0.000	0.019	0.032	0.013	0.021	-0.007	-0.003	0.020	0.006	0.003	0.020
Grade	-0.002	-0.001	0.029	-0.003	-0.002	0.032	0.009	0.006	0.033	-0.036	-0.022	0.032
Female (reference=Male)	0.257	0.498***	0.053	-0.240	-0.442***	0.056	0.223	0.437***	0.052	0.257	0.517***	0.056
R ²	0.122			0.094			0.104			0.110		

DISCUSSION

TOOLSEL was assembled from parts of existing measures to assess teachers' perceptions of students' classroom behaviors that reflect a set of social, emotional, and cognitive skills. TOOLSEL was intended to be used to evaluate a classroom-based SEL intervention for Syrian refugee children in nonformal education settings in Lebanon. Measures used to evaluate programs must meet a high standard of evidence for validity and reliability, given that the results often are used for accountability purposes and for program and policy decisionmaking that can have widespread consequences. Evidence indicates that TOOLSEL holds promise for use as a program-evaluation measure; however, we make several recommendations that would strengthen the data resulting from the use of this tool.

First, we found evidence of TOOLSEL's internal coherence, with a consistent factor structure that is meaningful and unique to the population and context. While the empirical data did not support the originally hypothesized factors for the five discrete subscales assembled across different tools, a series of exploratory and confirmatory factor analyses provided consistent support for a four-factor structure measuring teachers' perceptions of student behaviors in a classroom context: (1) Prosocial Behavior and Academic Engagement, (2) Social Problems, (3) Working Memory Functioning, and (4) Emotional and Behavioral Regulation. It is important to note that some of these final TOOLSEL subconstructs consist of items from across multiple, theoretically distinct subdomains of social and emotional skills. These results suggest that teachers are identifying the behaviors of "good" or "well-functioning" students, but not distinguishing between specific behavior subdomains; for example, prosocial versus classroom engagement behaviors (e.g., "Showing empathy" vs. "Working hard"); and emotional versus behavioral regulation skills (e.g., "Can calm down when excited or all wound up" vs. "Waits to be called on before responding"). In addition, the Prosocial Behavior and Academic Engagement subscale was highly correlated with the Working Memory Functioning and Emotional and Behavioral Regulation subscales. These findings may indicate cultural and contextual specificity in teachers' perceptions of children's social and emotional competence, and the subscales generated from this study may capture the children's skills that are better aligned with the cultural and contextual understanding of child development. On the other hand, it also may point to a limitation of teachers' reporting SEL skills. The patterns of high correlation among teacher-reported measures of related constructs are also observed in the non-EiE settings, such as the previous studies conducted in the US and Greece (Koth et al. 2009; Kourkounasiou and Skordilis 2014). Teachers are not typically trained in observing specific, distinct, social and emotional skills,

and they may rely on their global perceptions of individual children as good or bad, or as well-behaved or disruptive. This lack of specificity in teacher ratings may be important to consider when using teacher-reported measures for purposes that require an assessment of specific social, emotional, and cognitive processes.

Second, all of the empirically derived subscales for these four factors were consistent internally and over time with this sample of Syrian refugee children who were attending Lebanese public schools and taught by Lebanese teachers, which provides strong evidence of reliability. Such evidence of reliability is an important criterion for measures used for program-evaluation purposes, given that measurement error can attenuate the detection of treatment effects (Raudenbush and Sadoff 2008). Specifically, the subscales showed high internal consistency, which indicates that teachers generally provided consistent ratings on items within a subscale.

Third, we found evidence of measurement invariance with TOOLSEL by treatment, age, and gender groups, and across time (fall and spring). This means that the measure functions in the same way and is not biased against any subgroup by treatment condition, gender, or age when comparing the differences in TOOLSEL constructs. TOOLSEL also can be used without bias for program-evaluation purposes with pre- and posttest design, due to the differential functioning of the measure before and after the program implementation. In this case, some of the TOOLSEL constructs showed increases (Social Problems) or decreases (Prosocial Behavior and Academic Engagement) over the duration of the program period (six months, from fall to spring). While we do not have enough information on the normative developmental patterns and change in teachers' perceptions over time for Syrian refugee children in Lebanon to determine whether these changes are in the expected direction or at the expected magnitude, these results provide some support for their use in program evaluation to detect change over the program implementation period.

Fourth, the correlational evidence provides initial support for the validity of TOOLSEL. Specifically, the four constructs showed moderate autocorrelations over the course of six months and suggested that the teachers' perceptions of children's SEL skills display some degree of continuity and some degree of change (i.e., they are relatively stable over time). While these correlations are not very high, they are aligned with US research suggesting that SEL constructs tend to be more influenced by contextual factors and are likely to vary over time, as compared to academic skills, which tend to be highly stable over time (Soland et al. 2019). We also found significant gender differences in the expected directions,

given the current literature (Zimmermann and Iwanski 2014), which suggests that TOOLSEL is sensitive to detecting teachers' perceptions of gender difference in children's SEL skills (Elzein and Ammar 2010; Keresteš 2006; Lumley et al. 2002). Specifically, teachers rated girls higher than boys on Prosocial Behavior and Academic Engagement, Working Memory Functioning, and Emotional and Behavioral Regulation factors, and lower on Social Problems. However, it was not sensitive to detecting age differences, and there is not yet evidence that TOOLSEL can be used to detect developmental differences in the SEL constructs it has been designed to measure.

In addition, teacher ratings for each of the TOOLSEL subconstructs were generally correlated with other similar concepts in the expected directions, albeit at a relatively small magnitude ($rs < 0.2$). This includes an assessor-report measure of behavioral regulation, a performance-based tablet assessment of cognitive function, and child self-reports of experiencing victimization at school. It is not uncommon for reports from different raters to provide discrepant information (Buckley and Krachman 2016). While such discrepancies are often treated as a nuisance, recent research has demonstrated that discrepancies across informants can contain useful information that is helpful in interpreting program impacts, and for predicting longer-term adjustment and wellbeing (De Los Reyes 2011). While teacher reports provide meaningful information about the teachers' perception and interpretation of children's classroom behaviors, the use of multiple measurement methods and informants will be valuable in understanding children's social and emotional development in emergency contexts—especially when the purpose of assessment demands understanding children's behaviors, attitudes, and skills across multiple settings.

IMPLICATIONS FOR USE

FEASIBILITY CONSIDERATIONS

Given the resource constraints common in EiE contexts, it is important to consider the field feasibility of a measure and to use caution in interpreting the evidence from teacher reports in EiE settings, for the following reasons: (1) teachers may not know students very well if the student population they serve is highly mobile or attends lessons infrequently; (2) teachers may not have time to provide thoughtful and reliable information on individual children, as they are balancing a number of competing demands—including coping with their own experiences of trauma and adversity—and also may have limited training and experience in observing

and working with children; (3) reports from teachers in refugee contexts who come from a host community with a different cultural background and context than that of the refugee children may project systematic bias against the refugees that reflects the tension between the refugee and host communities. Given these considerations, we provide several more regarding the adaptation and use of TOOLSEL.

ADAPTATIONS AND CONSIDERATIONS FOR USE

While the evidence provided in this study largely supports the use of TOOLSEL for evaluation purposes with Syrian refugee children living in Lebanon, the findings are not assumptively generalizable to different populations and contexts. Hence, we strongly recommend piloting, adapting, and reevaluating the psychometric properties of the measure before using it with different populations and in different contexts. We provide a few suggestions for researchers and practitioners considering the use of TOOLSEL.

Most importantly, researchers and practitioners should ensure that the setting and structure of the program are suitable for using TOOLSEL, and that they are using it to evaluate the program's impact. TOOLSEL is designed for use in classrooms and learning spaces by teachers or facilitators who have regular and extensive interactions with individual children. This means TOOLSEL is appropriate to use with small to medium-size classes or learning groups where the children are engaged in learning activities facilitated by adults. It only can be used after the program has been launched and the teachers have had time to get to know the children well. This may not be the case in many EiE settings, where teachers often work with large groups of children and are too overwhelmed by multiple demands to get to know the children individually; moreover, children may not attend the program regularly, due to the safety and economic concerns common in EiE settings. Finally, while it may be tempting to use a measure like TOOLSEL for multiple purposes in resource-strained EiE settings, we emphasize that TOOLSEL should not be used for purposes other than program evaluation and research. Given the limited specificity of the teacher ratings we found in this report, we strongly recommend against using TOOLSEL for screening or formative assessment purposes.

Once TOOLSEL is deemed appropriate for a particular setting and purpose, we recommend a set of strategies to ensure that teachers can differentiate meaningfully between children and report on their individual behaviors in class, and thus improve the validity of the teacher-reporting scales. First, cognitive

interviewing techniques can be used during the measure pilot to understand how teachers are interpreting and responding to items, and their perceptions of the utility, reliability, and cultural and ecological validity in crisis contexts. This information can be used to refine items and assessment directions/procedures to help teachers distinguish clearly between social skills and learning-based cognitive processes, and to improve the measure's utility and validity in reflecting teachers' perspectives.

Second, explicit assessor training for teachers in filling out the survey can improve the validity of their reports. Teachers in EiE settings may not have enough experience or training to observe carefully and report on the children's individual behaviors. They also may lack sufficient literacy to understand the questions fully, especially when the written instructional language is not their first language (Dryden-Peterson 2015).³ Therefore, establishing common understanding of the meaning of items presented in TOOLSEL for the concepts each item is intended to capture may increase the specificity of the concepts TOOLSEL can capture, and improve its reliability and validity.

Third, in planning for the assessment, we recommend implementing strategies that reduce the burden of reporting for teachers. This may include selecting a random subset of children for teachers to report on or providing coverage in the classroom to give the teacher time to fill out the survey. Fourth, we recommend using behavioral "nudge" strategies during the assessment that prime teachers to think about the many different behaviors of the focal child. Trained enumerators or tablet algorithms also could be used to quickly identify when teachers are providing a child with the same score on all items, which will result in statistics with low reliability. Fifth, we recommend that the items on the measure be adapted for each age group (i.e., early childhood, middle childhood, adolescence) so that each item is situated within an appropriate developmental trajectory. This may partially remedy the teacher reference bias and provide teachers with different forms of the measure that are based on the age of the child, rather than receiving the same measure regardless of the child's characteristics. Finally, we recommend collecting data from multiple sources to triangulate the data most effectively.

³ All teachers in our study had sufficient literacy, as their native/first language was Arabic (the language of instruction and research for this study) and they had a high school education or higher.

CONCLUSION

This study provides evidence that TOOLSEL offers coherent, reliable, consistent, and empirically valid information that is unbiased across treatment groups, gender and age groups, and the timing of the assessment. In addition, we find additional support for using TOOLSEL in program evaluation, given its ability to detect changes during a six-month implementation of the program with Syrian refugee children living in Lebanon. While testing the sensitivity to treatment is beyond the scope of this study and only can be done as a part of an impact evaluation of a program that is proven to show changes in these SEL skills, the evidence produced in the current study provides some confidence in the decision to use TOOLSEL for evaluation purposes. We acknowledge that the TOOLSEL construction relies on knowledge and tools that are based mainly on research in non-EiE contexts and thus that make a limited contribution to the decolonization of research and knowledge (Bermúdez, Muruthi, and Jordan 2016; Zavala 2013). When possible, it is more desirable to develop and adapt SEL measures that fully reflect the local context and culture and to use methodological approaches that are rooted in participant-informed coconstruction of knowledge, such as participatory research methods (Javdani, Singh, and Sichel 2017). When the tools, time, and resources needed to generate such measures are not available, TOOLSEL provides a feasible and practical alternative for assessing SEL skills that is suitable for program evaluation in EiE settings.

Indeed, research that, like this study, empirically evaluates tools or hypotheses that are developed primarily in non-EiE settings holds promise as a starting point for valuable culturally and contextually grounded research. Not all research can be built from the ground up, especially in conflict- and crisis-affected and resource-poor contexts, where the effective and prompt provision of services that support the population's urgent needs is prioritized. In such cases, this type of research can provide a practical alternative that takes the current status quo—which relies on imposing “evidence-based” knowledge from the non-EiE context—a step further toward building contextually and culturally relevant knowledge in situations and with populations that have traditionally been underrepresented, misrepresented, and marginalized.

REFERENCES

- AERA (American Educational Research Association), APA (American Psychological Association), and NCME (National Council on Measurement in Education). 2014. *Standards for Educational and Psychological Testing*, 3rd ed. Washington, DC: AERA.
- Asparouhov, Tihomir, and Bengt Muthén. 2010. “Weighted Least Squares Estimation with Missing Data.” *Mplus Technical Appendix* (2010):1-10.
- Bakrania, Shivit, Nikola Balvin, Silvio Daidone, and Jacobus de Hoop. 2021. *Impact Evaluation in Settings of Fragility and Humanitarian Emergency*. Innocenti Discussion Papers 2021-02. Florence: UNICEF Office of Research–Innocenti. <https://www.unicef-irc.org/publications/pdf/Impact-Evaluation-in-Settings-of-Fragility-and-Humanitarian-Emergency.pdf>.
- Bermúdez, J. Maria, Bertranna A. Muruthi, and Lorien S. Jordan. 2016. “Decolonizing Research Methods for Family Science: Creating Space at the Center.” *Journal of Family Theory & Review* 8 (2): 192-206. <https://doi.org/10.1111/jftr.12139>.
- Betancourt, Theresa, Sarah Meyers-Ohki, Alexandra Charrow, and Wietse Tol. 2013. “Interventions for Children Affected by War: An Ecological Perspective on Psychosocial Support and Mental Health Care.” *Harvard Review of Psychiatry* 21 (2): 70-91. <https://doi.org/10.1097/HRP.0b013e318283bf8f>.
- Blair, Clancy, and Rachel Razza. 2007. “Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten.” *Child Development* 78 (2): 647-63. <https://doi.org/10.1111/j.1467-8624.2007.01019.x>.
- Boekaerts, Monique, and Reinhard Pekrun. 2015. “Emotions and Emotion Regulation in Academic Settings.” In *Handbook of Educational Psychology*, edited by Lyn Corno and Eric M. Anderman, 90-104. New York: Routledge.
- Brody, Leslie. 2000. “The Socialization of Gender Differences in Emotional Expression: Display Rules, Infant Temperament, and Differentiation.” *Gender and Emotion: Social Psychological Perspectives* 2: 24-47. <https://doi.org/10.1017/CBO9780511628191.003>.

- Buckley, Katie, and Sara Krachman. 2016. *Initial Findings from the Boston Charter Research Collaborative*. Boston, MA: Transforming Education. <https://www.transformingeducation.org/wp-content/uploads/2017/04/TE-BCRCWorkingPaperFINAL.pdf>.
- Burde, Dana, Ozen Guven, Jo Kelcey, Heddy Lahmann, and Khaled Al-Abbadi. 2015. "What Works to Promote Children's Educational Access, Quality of Learning, and Wellbeing in Crisis-Affected Contexts." *Education Literature Review*. London: Department for International Development. https://assets.publishing.service.gov.uk/media/57a0897ee5274a31e00000e0/61127-Education-in-Emergencies-Rigorous-Review_FINAL_2015_10_26.pdf.
- Burde, Dana, Amy Kapit, Rachel Wahl, Ozen Guven, and Margot Skarpeteig. 2017. "Education in Emergencies: A Review of Theory and Research." *Review of Educational Research* 87 (3): 619-58. <https://doi.org/10.3102/0034654316671594>.
- Catalán, Héctor. 2019. "Reliability, Population Classification and Weighting in Multidimensional Poverty Measurement: A Monte Carlo Study." *Social Indicators Research* 142 (3): 887-910. <https://doi.org/10.1007/s11205-018-1950-z>.
- Chen, Fang Fang. 2007. "Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance." *Structural Equation Modeling: A Multidisciplinary Journal* 14 (3): 464-504. <https://doi.org/10.1080/10705510701301834>.
- Cole, Pamela, Margaret Michel, and Lauren O'Donnell Teti. 1994. "The Development of Emotion Regulation and Dysregulation: A Clinical Perspective." *Monographs of the Society for Research in Child Development* 59 (2-3): 73-102.
- Conduct Problems Prevention Research Group. 1990. "Social Competence Scale (Teacher Version)." *Zugriff Am* 5: 2007.
- Davies, Lynn, and Christopher Talbot. 2008. "Learning in Conflict and Postconflict Contexts." *Comparative Education Review* 52 (4): 509-18. <https://doi.org/10.1086/591295>.

- De Los Reyes, Andres. 2011. "Introduction to the Special Section: More Than Measurement Error. Discovering Meaning behind Informant Discrepancies in Clinical Assessments of Children and Adolescents." *Journal of Clinical Child & Adolescent Psychology* 40 (1): 1-9. <https://doi.org/10.1080/15374416.2011.533405>.
- De Los Reyes, Andres, Tara Augenstein, Mo Wang, Sarah Thomas, Deborah Drabick, Darcy Burgers, and Jill Rabinowitz. 2015. "The Validity of the Multi-Informant Approach to Assessing Child and Adolescent Mental Health." *Psychological Bulletin* 141 (4): 858. <https://doi.org/10.1037/a0038498>.
- Duncan, Robert, Megan McClelland, and Alan Acock. 2017. "Relations between Executive Function, Behavioral Regulation, and Achievement: Moderation by Family Income." *Journal of Applied Developmental Psychology* 49: 21-30. <https://doi.org/10.1016/j.appdev.2017.01.004>.
- Durlak, Joseph, Roger Weissberg, Allison Dymnicki, Rebecca Taylor, and Kriston Schellinger. 2011. "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-based Universal Interventions." *Child Development* 82 (1): 405-32. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>.
- Eisenberg, Nancy, Cynthia L. Smith, and Tracy L. Spinrad. 2011. "Effortful Control: Relations with Emotion Regulation, Adjustment, and Socialization in Childhood." In *Handbook of Self-Regulation: Research, Theory, and Applications*, 2nd ed., edited by Kathleen D. Vohs and Roy F. Baumeister, 263-83. New York: Guilford Press.
- Ellis, Bruce, Anthony Volk, Jose-Michael Gonzalez, and Dennis Embry. 2016. "The Meaningful Roles Intervention: An Evolutionary Approach to Reducing Bullying and Increasing Prosocial Behavior." *Journal of Research on Adolescence* 26 (4): 622-37. <https://doi.org/10.1111/jora.12243>.
- Elzein, Heyam, and DIALA Ammar. 2010. "Parent and Teacher Perceptions of Assessing Lebanese Children's Reaction to War-Related Stress: A Survey of Psychological and Behavioral Functioning." *Journal of Child & Adolescent Trauma* 3 (4): 255-78. <https://doi.org/10.1080/19361521.2010.523060>.
- Espelage, Dorothy, and Melissa Holt. 2001. "Bullying and Victimization during Early Adolescence: Peer Influences and Psychosocial Correlates." *Journal of Emotional Abuse* 2 (2-3): 123-42. https://doi.org/10.1300/J135v02n02_08.

- Fabes, Richard, and Nancy Eisenberg. 1998. "Meta-Analyses of Age and Sex Differences in Children's and Adolescents' Prosocial Behavior." *Handbook of Child Psychology* 3: 1-29.
- Ford, Cassie, Ha Yeon Kim, Lindsay Brown, J. Lawrence Aber, and Margaret Sheridan. 2019. "A Cognitive Assessment Tool Designed for Data Collection in the Field in Low- and Middle-Income Countries." *Research in Comparative and International Education* 14 (1): 141-57. <https://doi.org/10.1177/1745499919829217>.
- Furrer, Carrie, and Ellen Skinner. 2003. "Sense of Relatedness as a Factor in Children's Academic Engagement and Performance." *Journal of Educational Psychology* 95 (1): 148-62. <https://doi.org/10.1037/0022-0663.95.1.148>.
- Gest, Scott D., Janet A. Welsh, and Celene E. Domitrovich. 2005. "Behavioral Predictors of Changes in Social Relatedness and Liking School in Elementary School." *Journal of School Psychology* 43 (4): 281-301. <https://doi.org/10.1016/j.jsp.2005.06.002>.
- Goldman-Rakic, Patricia. 1996. "Regional and Cellular Fractionation of Working Memory." *Proceedings of the National Academy of Sciences* 93 (24): 13473-80. <https://doi.org/10.1073/pnas.93.24.13473>.
- Hamoudi, Amar, and Margaret Sheridan. 2015. "Unpacking the Black Box of Cognitive Ability: A Novel Tool for Assessment in a Population-Based Survey." Manuscript under review, Young Lives Study.
- Hartup, Willard. 1996. "The Company They Keep: Friendships and Their Developmental Significance." *Child Development* 67 (1): 1-13. <https://doi.org/10.1111/j.1467-8624.1996.tb01714.x>.
- Hayes, Andrew, and Jacob Coutts. 2020. "Use Omega Rather than Cronbach's Alpha for Estimating Reliability, But..." *Communication Methods and Measures* 14 (1): 1-24. <https://doi.org/10.1080/19312458.2020.1718629>.
- Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411-82. <https://doi.org/10.1086/504455>.

- Hu, Li-tze, and Peter Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6 (1): 1-55. <https://doi.org/10.1080/10705519909540118>.
- Hughes, Claire, and Rosie Ensor. 2011. "Individual Differences in Growth in Executive Function across the Transition to School Predict Externalizing and Internalizing Behaviors and Self-Perceived Academic Success at 6 Years of Age." *Journal of Experimental Child Psychology* 108 (3): 663-76. <https://doi.org/10.1016/j.jecp.2010.06.005>.
- Jacob, Robin, and Julia Parkinson. 2015. "The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement: A Review." *Review of Educational Research* 85 (4): 512-52. <https://doi.org/10.3102/0034654314561338>.
- Javdani, Shabnam, Sukhmani Singh, and Corianna Sichel. 2017. "Negotiating Ethical Paradoxes in Conducting a Randomized Controlled Trial: Aligning Intervention Science with Participatory Values." *American Journal of Community Psychology* 60 (3-4): 439-49. <https://doi.org/10.1002/ajcp.12185>.
- Jones, Damon, Mark Greenberg, and Max Crowley. 2015. "Early Social-Emotional Functioning and Public Health: The Relationship Between Kindergarten Social Competence and Future Wellness." *American Journal of Public Health* 105 (11): 2283-90. <https://doi.org/10.2105/AJPH.2015.302630>.
- Jones, Stephanie, Rebecca Bailey, and Sophie Barnes. 2015. *Classroom Executive Function Survey (CEFS): A Measure of Executive Function and Self-Regulation Skills in Everyday Classroom Behavior*. Cambridge, MA: Harvard University Press.
- Jordans, Mark, Hugo Pigott, and Wietse Tol. 2016. "Interventions for Children Affected by Armed Conflict: A Systematic Review of Mental Health and Psychosocial Support in Low- and Middle-Income Countries." *Current Psychiatry Reports* 18 (1): 1-15. <https://doi.org/10.1007/s11920-015-0648-z>.
- Keresteš, Gordana. 2006. "Children's Aggressive and Prosocial Behavior in Relation to War Exposure: Testing the Role of Perceived Parenting and Child's Gender." *International Journal of Behavioral Development* 30 (3): 227-39. <https://doi.org/10.1177/0165025406066756>.

- Kim, Ha Yeon, Lindsay Brown, Carly Tubbs Dolan, Kalina Gjicali, Rena Deitz, Sol Prieto Bayona, and J. Lawrence Aber. "Comprehensive Social Emotional Learning Intervention with Syrian Refugee Children: Impact Variation by Pre- and Post-Migration Conflict Experiences." Unpublished manuscript, 2021.
- Margaret Sheridan, and John Lawrence Aber. 2020. "Post-Migration Risks, Developmental Processes, and Learning among Syrian Refugee Children in Lebanon." *Journal of Applied Developmental Psychology* 69: 101142. <https://doi.org/10.1016/j.appdev.2020.101142>.
- Kochanska, Grazyna, Kathleen Murray, and Elena Harlan. 2000. "Effortful Control in Early Childhood: Continuity and Change, Antecedents, and Implications for Social Development." *Developmental Psychology* 36 (2): 220-32. <https://doi.org/10.1037/0012-1649.36.2.220>.
- Koth, Christine W., Catherine P. Bradshaw, and Philip J. Leaf. 2009. "Teacher Observation of Classroom Adaptation—Checklist: Development and Factor Structure." *Measurement and Evaluation in Counseling and Development* 42 (1): 15-30. <https://doi.org/10.1177/0748175609333560>.
- Kourkounasiou, Mari, and Emmanouil Skordilis. 2014. "Validity and Reliability Evidence of the TOCA-C in a Sample of Greek Students." *Psychological Reports* 115 (3): 766-83. <https://doi.org/10.2466/08.11.PR0.115c31z5>.
- LaFontana, Kathryn, and Antonius Cillessen. 2002. "Children's Perceptions of Popular and Unpopular Peers: A Multimethod Assessment." *Developmental Psychology* 38 (5): 635-47. <https://doi.org/10.1037/0012-1649.38.5.635>.
- Lei, Pui-Wa. 2009. "Evaluating Estimation Methods for Ordinal Data in Structural Equation Modeling." *Quality and Quantity* 43 (3): 495. <https://doi.org/10.1007/s11135-007-9133-z>.
- Lumley, Vicki, Cheryl McNeil, Amy Herschell, and Alisa Bahl. 2002. "An Examination of Gender Differences among Young Children with Disruptive Behavior Disorders." *Child Study Journal* 32 (2): 89-101.
- McCoy, Dana Charles, Stephanie Zuilkowski, Hirokazu Yoshikawa, and Günther Fink. 2017. "Early Childhood Care and Education and School Readiness in Zambia." *Journal of Research on Educational Effectiveness* 10 (3): 482-506. <https://doi.org/10.1080/19345747.2016.1250850>.

- McDonald, Roderick P. 1999. *Test Theory: A Unified Treatment*. New Jersey: Lawrence Erlbaum Associates.
- McKown, Clark, and Rhona Weinstein. 2008. "Teacher Expectations, Classroom Context, and the Achievement Gap." *Journal of School Psychology* 46 (3): 235-61. <https://doi.org/10.1016/j.jsp.2007.05.001>.
- McRae, Kateri, Kevin Ochsner, Iris Mauss, John Gabrieli, and James Gross. 2008. "Gender Differences in Emotion Regulation: An FMRI Study of Cognitive Reappraisal." *Group Processes & Intergroup Relations* 11 (2): 143-62. <https://doi.org/10.1177/1368430207088035>.
- Muthén, Bengt, and Linda Muthén. 2014. MPlus (version 7.2). Computer software. Los Angeles: Muthén and Muthén.
- Raudenbush, Stephen, and Sally Sadoff. 2008. "Statistical Inference When Classroom Quality Is Measured with Error." *Journal of Research on Educational Effectiveness* 1 (2): 138-54. <https://doi.org/10.1080/19345740801982104>.
- Raver, Cybele, Stephanie Jones, Christine Li-Grining, Fuhua Zhai, Kristen Bub, and Emily Pressler. 2011. "CSRP's Impact on Low-Income Preschoolers' Preacademic Skills: Self-Regulation as a Mediating Mechanism." *Child Development* 82 (1): 362-78. <https://doi.org/10.1111/j.1467-8624.2010.01561.x>.
- Reise, Steven P., Keith F. Widaman, and Robin H. Pugh. 1993. "Confirmatory Factor Analysis and Item Response Theory: Two Approaches for Exploring Measurement Invariance." *Psychological Bulletin* 114 (3): 552.
- Revelle, William, and Richard Zinbarg. 2009. "Coefficients Alpha, Beta, Omega, and the Glb: Comments on Sijtsma." *Psychometrika* 74 (1): 145.
- Rothbart, Mary K., and M. Rosario Rueda. 2005. "The Development of Effortful Control." In *Developing Individuality in the Human Brain: A Tribute to Michael I. Posner*, edited by Ulrich Mayr, Edward Awh, and Steven W. Keele, 167-88. Washington, DC: American Psychological Association. <https://doi.org/10.1037/11108-009>.
- Shah, Ritesh. 2017. *Improving Children's Wellbeing: An Evaluation of NRC's Better Learning Programme in Palestine*. Oslo: Norwegian Refugee Council. <https://www.nrc.no/globalassets/pdf/evaluations/nrc-blp-palestine-full-report.pdf>.

- Simon, Richard, and Alan Rudell. 1967. "Auditory S-R Compatibility: The Effect of an Irrelevant Cue on Information Processing." *Journal of Applied Psychology* 51 (3): 300. <https://doi.org/10.1037/h0020586>.
- Smith-Donald, Radiah, Cybele Raver, Tiffany Hayes, and Breeze Richardson. 2007. "Preliminary Construct and Concurrent Validity of the Preschool Self-Regulation Assessment (PSRA) for Field-Based Research." *Early Childhood Research Quarterly* 22 (2): 173-87. <https://doi.org/10.1016/j.ecresq.2007.01.002>.
- Valiente, Carlos, Nancy Eisenberg, R. G. Haugen, Tracy Spinrad, Claire Hofer, Jeffrey Liew, and Anne Kupfer. 2011. "Children's Effortful Control and Academic Achievement: Mediation through Social Functioning." *Early Education & Development* 22 (3): 411-33. <https://doi.org/10.1080/10409289.2010.505259>.
- Van de Mortel, Thea F. 2008. "Faking It: Social Desirability Response Bias in Self-Report Research." *Australian Journal of Advanced Nursing* 25 (4): 40.
- Vandenberg, Robert, and Charles Lance. 2000. "A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research." *Organizational Research Methods* 3 (1): 4-70. <https://doi:10.1177/109442810031002>.
- Zavala, Miguel. 2013. "What Do We Mean by Decolonizing Research Strategies? Lessons from Decolonizing, Indigenous Research Projects in New Zealand and Latin America." *Decolonization: Indigeneity, Education & Society* 2 (1): 55-71.
- Zelazo, Philip David, Stephanie M. Carlson, and Amanda Kesek. 2008. "The Development of Executive Function in Childhood." In *Handbook of Developmental Cognitive Neuroscience*, 2nd ed., edited by Charles A. Nelson and Monica Luciana, 553-74. Cambridge, MA: MIT Press.
- Zimmermann, Peter, and Alexandra Iwanski. 2014. "Emotion Regulation from Early Adolescence to Emerging Adulthood and Middle Adulthood: Age Differences, Gender Differences, and Emotion-Specific Developmental Variations." *International Journal of Behavioral Development* 38 (2): 182-94. <https://doi.org/10.1177/0165025413515405>.

APPENDIX A

TOOLSEL MEASURE ITEM DESCRIPTION AND DESCRIPTIVE STATISTICS

Table A1: TOOLSEL Measure Descriptions

Item	Description
1	TOC1: In the last two weeks [your child]: Concentrates
2	TOC2: In the last two weeks [your child]: Is friendly
3	TOC3: In the last two weeks [your child]: Pays attention
4	TOC4: In the last two weeks [your child]: Breaks rules (removed)
5	TOC5: In the last two weeks [your child]: Is liked by classmates
6	TOC7: In the last two weeks [your child]: Works hard
7	TOC9: In the last two weeks [your child]: Shows empathy & compassion for other's feelings
8	TOC10: In the last two weeks [your child]: Gets angry when provoked by other children
9	TOC11: In the last two weeks [your child]: Stay on task (removed)
10	TOC12: In the last two weeks [your child]: Yells at others
11	TOC14: In the last two weeks [your child]: Is rejected by classmates
12	TOC15: In the last two weeks [your child]: Fights
13	TOC17: In the last two weeks [your child]: Has many friends (removed)
14	TOC20: In the last two weeks [your child]: Teases classmates
15	TOC21: In the last two weeks [your child]: Learns up to ability
16	CEFS1: In the last two weeks [your child]: Remembers lists or items in the correct order
17	SCS2: In the last two weeks [your child]: Can accept things not going his/her way (removed)

Item	Description
18	CEFS2: In the last two weeks [your child]: Follows multiple-step instructions
19	CEFS3: In the last two weeks [your child]: Uses multiple rules to complete a task
20	SCS8: In the last two weeks [your child]: Thinks before acting (removed)
21	CEFS4: In the last two weeks [your child]: Waits to be called on before responding
22	SCS11: In the last two weeks [your child]: Can calm down when excited or all wound up
23	CEFS6: In the last two weeks [your child]: Transitions easily to new activities, tasks, or major parts of the day (e.g., from recess)
24	CEFS8: In the last two weeks [your child]: Uses self-control techniques
25	SCS12: In the last two weeks [your child]: Can wait in line patiently when necessary
26	CEFS9: In the last two weeks [your child]: Waits patiently for her/his turn
27	CEFS10: In the last two weeks [your child]: Uses listening skills
28	SCS18: In the last two weeks [your child]: Controls temper when there is a disagreement

Note: Items labeled starting with TOC are taken from TOCA-C, with original item numbers used in TOCA-C. Similarly, items labeled starting with SCS were taken from the SCS Emotional Regulation Scale, and items labeled starting with CEFS were taken from CEFS. Some items on this list were removed from the final scale, as indicated.

EVIDENCE FOR TOOLSEL IN EIE PROGRAM EVALUATION

Table A2: Descriptive Statistics of Indicators by Proposed Construct

Item	Fall (N=3,254) N	Spring (N=2,952) M	SD	Min	Max	N	M	SD	Min	Max
TOC1	3254	3.632	1.103	1	5	2950	3.536	1.133	1	5
TOC2	3248	3.823	1.015	1	5	2947	3.680	1.055	1	5
TOC3	3246	3.673	1.092	1	5	2942	3.533	1.146	1	5
TOC4	3233	2.467	1.151	1	5	2942	2.359	1.120	1	5
TOC5	3223	3.764	0.966	1	5	2933	3.638	1.025	1	5
TOC7	3227	3.592	1.064	1	5	2924	3.487	1.102	1	5
TOC9	3212	3.597	1.038	1	5	2922	3.505	1.067	1	5
TOC10	3224	2.717	1.248	1	5	2929	2.823	1.207	1	5
TOC11	3204	3.419	1.129	1	5	2912	3.370	1.140	1	5
TOC12	3208	2.112	1.170	1	5	2923	2.259	1.163	1	5
TOC14	3229	1.850	1.065	1	5	2926	1.988	1.086	1	5
TOC15	3231	2.005	1.190	1	5	2940	2.158	1.194	1	5
TOC17	3215	3.485	1.098	1	5	2919	3.487	1.091	1	5
TOC20	3207	2.183	1.242	1	5	2913	2.240	1.223	1	5
TOC21	3208	3.504	1.069	1	5	2924	3.434	1.060	1	5
SCS2	3230	3.540	1.056	1	5	2935	3.430	1.053	1	5
SCS8	3232	3.468	1.094	1	5	2942	3.418	1.103	1	5
SCS11	3221	3.518	1.113	1	5	2932	3.449	1.071	1	5
SCS12	3214	3.530	1.105	1	5	2929	3.457	1.074	1	5

Item	Fall (N=3,254) N	Spring (N=2,952) M	SD	Min	Max	N	M	SD	Min	Max
SCS18	3242	3.535	1.185	1	5	2947	3.447	1.138	1	5
CEFS1	3218	3.534	1.050	1	5	2933	3.462	1.071	1	5
CEFS2	3241	3.609	1.092	1	5	2944	3.518	1.057	1	5
CEFS3	3235	3.348	1.127	1	5	2946	3.382	1.097	1	5
CEFS4	3234	3.521	1.105	1	5	2937	3.454	1.095	1	5
CEFS6	3244	3.587	1.073	1	5	2941	3.533	1.062	1	5
CEFS8	3234	3.443	1.097	1	5	2943	3.430	1.057	1	5
CEFS9	3240	3.535	1.122	1	5	2940	3.473	1.084	1	5
CEFS10	3239	3.606	1.099	1	5	2944	3.545	1.089	1	5

Note: Items labeled starting with TOC are taken from TOCA-C, with original item numbers used in TOCA-C. Similarly, items labeled starting with SCS were taken from SCS Emotional Regulation Scale, and items labeled starting with CEFS were taken from CEFS.

APPENDIX B

MODEL FIT INDICES

Table B1: Model Fit Indices of Confirmatory Factor Analyses of Originally Proposed Subscales (five-factor models)

Wave	k	Chi-sq	df	p	CFI	TLI	RMSEA	SRMR
Fall	150	4068.026	340	0	0.929	0.921	0.082	0.051
Spring	150	4312.336	340	0	0.929	0.921	0.089	0.055

EVIDENCE FOR TOOLSEL IN EIE PROGRAM EVALUATION

Table B2: Model Fit Indices of Exploratory Factor Analyses Four-Factor Models for Fall and Spring

Wave	CFI	TLI	RMSEA	SRMR
Fall	0.965	0.951	0.060	0.025
Spring	0.965	0.951	0.065	0.023

Table B3: Model Fit Indices of Confirmatory Factor Analyses Final Models for Fall and Spring

Wave	k	Chi-sq	df	p	CFI	TLI	RMSEA	SRMR
Fall	123	1529.214	222	0	0.972	0.968	0.06	0.029
Spring	123	1636.506	222	0	0.972	0.968	0.066	0.037

Table B4: Model Fit Indices of Treatment Invariance Models

Model	k	χ^2	df	p	$\Delta\chi^2$	df	p	CFI	TLI	RMSEA	SRMR
Fall											
Configural	246	2255.002	444	0				0.967	0.962	0.05	0.032
Metric	227	1711.771	463	0	40.906	19	0.0025	0.977	0.975	0.041	0.033
Scalar	139	1720.471	551	0	130.661	88	0.0022	0.979	0.98	0.036	0.034
Spring											
Configural	246	2718.794	444	0				0.96	0.955	0.059	0.038
Metric	227	2132.169	463	0	65.362	19	0	0.971	0.968	0.049	0.04
Scalar	139	2183.087	551	0	206.47	88	0	0.972	0.974	0.045	0.041

Table B5: Model Fit Indices of Gender Invariance Models

Model	k	χ^2	df	p	$\Delta\chi^2$	df	p	CFI	TLI	RMSEA	SRMR
Fall											
Configural	246	3170	444	0				0.961	0.956	0.061	0.033
Metric	227	2315	463	0	51.89	19	0	0.973	0.971	0.050	0.035
Scalar	139	2289	551	0	122.28	88	0.001	0.975	0.977	0.044	0.035
Spring											
Configural	246	3582	444	0				0.965	0.960	0.069	0.038
Metric	227	2727	463	0	61.03	19	0	0.975	0.973	0.058	0.039
Scalar	139	2728	551	0	185.34	88	0	0.976	0.978	0.052	0.040

EVIDENCE FOR TOOLSEL IN EIE PROGRAM EVALUATION

Table B6: Model Fit Indices of Age Invariance Models

Model	k	χ^2	df	p	$\Delta\chi^2$	df	p	CFI	TLI	RMSEA	SRMR
Fall											
Configural	492	3488	888	0				0.973	0.969	0.060	0.033
Metric	435	2536	945	0	107.86	57	.0001	0.983	0.982	0.046	0.035
Scalar	171	2696	1209	0	277.49	264	0.272	0.984	0.987	0.039	0.036
Spring											
Configural	492	3837	888	0				0.975	0.972	0.067	0.037
Metric	435	2872	945	0	123.15	57	0	0.984	0.983	0.053	0.039
Scalar	171	3072	1209	0	357.55	264	0	0.984	0.987	0.046	0.040

Table B7: Model Fit Indices of Longitudinal Invariance Models

Model	k	χ^2	df	p	$\Delta\chi^2$	df	p	CFI	TLI	RMSEA	SRMR
Configural	262	3530.362	957	0				0.966	0.963	0.028	0.03
Metric	243	3163.245	976	0	35.621	19	0.0117	0.971	0.969	0.025	0.03
Scalar	155	3292.212	1064	0	211.309	107	0	0.97	0.971	0.025	0.031

CREATING A TOOL TO MEASURE CHILDREN'S WELLBEING: A PSS INTERVENTION IN SOUTH SUDAN

MOSES OLAYEMI, MELISSA TUCKER, MAMOUR CHOUL, TOM PUREKAL,
ARLENE BENITEZ, WENDY WHEATON, AND JENNIFER DEBOER

ABSTRACT

Since 2015, more than 560,000 South Sudanese primary school children have received psychosocial support (PSS) through the USAID-funded Integrated Essential Emergency Education Services program, which is implemented by UNICEF. Several South Sudan-based nongovernmental organizations partnered with UNICEF to train local teachers to implement the PSS activities in child-friendly spaces. To evaluate the impact this intervention had on students' wellbeing and academic performance, a multi-institutional consortium of multidisciplinary partners purposively sampled 2,982 students and 580 teachers in 64 schools from five states in the Republic of South Sudan. Critical to the evaluation's aims was the design of a contextually relevant, rigorously validated instrument to measure students' wellbeing in a region where research on PSS outcomes in education in emergencies is needed. In this article, we first present the process by which these survey instruments were designed, including the collaborative efforts of experts on measuring PSS outcomes in conflict settings and experts on the local context. We then describe how we tested for the construct validity of the resulting instrument and present the results of our confirmatory factor analysis of its three-factor model of social wellbeing, emotional wellbeing, and resilience/coping. Finally, based on our process and the resulting instrument, we make recommendations for future research on PSS outcomes in emergency settings.

Received January 27, 2020; revised February 23, 2021, and October 1, 2021; accepted November 8, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

INTRODUCTION

Incidents of violence and war have caused an extremely high level of displacement of South Sudanese children and youth both within and outside the geographic boundaries of the world's newest nation. Unfortunately, these forced displacements have worsened the country's already struggling education system. In 2015, for example, South Sudan reported an illiteracy rate of 84 percent for females and 73 percent for males (Republic of South Sudan Ministry of Education, Science, and Technology 2015). The 2016 clashes in Juba, South Sudan's capital and largest city, left close to one-third of the country's primary school learning spaces partially or completely destroyed. Data collected and analyzed in 2018 from South Sudan's Education Management Information System (EMIS) revealed a marked increase in the student dropout rate, which was compounded by a decline in enrollment (UNESCO 2018). However, the disruption of the education system is only one of the consequences the war and violence have had for South Sudan's youth population.

The research is clear: forced displacement can have profoundly negative effects on the uprooted populations (Amnesty International 2016; Kamau et al. 2004; UN Children's Fund [UNICEF] 2015a). The literature on the psychological wellbeing of displaced populations paints a concerning picture. Subjects who experience armed conflict and are forced to flee to temporary shelters, such as protection of civilians (POC) sites and refugee camps, are prone to sadness and depression, and have poor social and emotional skills (IASC 2007). Some medical practitioners warn that exposure to adversity can impair children's cognitive, physical, and mental health (Shonkoff, Boyce, and McEwen 2009). This suggests that, left unchecked, the negative consequences of forced displacement and exposure to violence can prevent children and youth from living up to their full potential as fully functional, productive members of the community. It is of critical importance, therefore, not only to seek effective ways to foster the social and emotional wellbeing of children in conflict settings alongside their academic needs, but to do so with a sense of urgency.

This article is divided into five sections. In the first section, we provide a backdrop for the study, including a short summary of the South Sudan context and current efforts at nation-building, followed by a discussion of the theoretical constructs of wellbeing we operationalized in our study, and of interventions that have been implemented and instruments that are used to measure wellbeing in conflict settings. In the second section, we review the purpose of this study and the broader research from which this paper emerged. We also discuss the implementation of psychosocial support (PSS) interventions in South Sudan, the modalities of PSS

activities, and child friendly-spaces (CFS). In the third section, we address our core research question and present the guiding framework for our development of the instrument. In section four we present the results of our analyses, and we conclude with a discussion of the results, policy implications, and limitations of our study, and applications for future work.

THE SOUTH SUDAN CONTEXT

The civil war in South Sudan has caused most South Sudanese to experience some adverse event, such as physical violence or forced displacement, as well as the effects these events have had on their psychological and emotional states. The mandates of President Kiir to dismiss prominent South Sudanese government figures in mid-2013, including the vice president and the secretary-general of the Sudan People's Liberation Movement, led to the outbreak of violence in Juba in December of that year. As conditions became more volatile, security deteriorated and spread from state to state, leading to the internal displacement of a large segment of the population. There is no consensus on the cause of the violence, although the various theories include an attempted coup to silence government officials and weaken state institutions, cause an imbalance of power, and militarize government institutions (International Crisis Group 2014). An inquiry into the cause of the conflict suggested that the violence may have been more ethnically polarized than political (African Union 2014).

In July 2016, the reappointment of the vice president was marked by another outbreak of violence in Juba. While the president and vice president openly condemned the violence sparked by their loyalists and promptly ordered a ceasefire ("South Sudan Clashes" 2016), the conflict spread throughout the country. Sporadic fighting continued for much of 2017 and 2018, and much of the population from the country's southern regions fled to Uganda. In September 2018, Sudan and Uganda brokered a new peace agreement for South Sudan.

Research suggests that the relationship between the causes and effects of conflict can be intricate and complex. For example, conflict has been identified as both a cause and an effect of inequitable access to education (Burde et al. 2017), and those worst affected usually are the most vulnerable members of the population (Centre for Research on the Epidemiology of Disasters 2013). During the war in South Sudan, acts of violence were rampant, including against the elderly, women,

children, and the disabled. As the EMIS report shows, schools and learning centers were greatly affected (UNESCO 2018), with reports of sexual assaults and violence in schools (UNICEF 2015b). The effects this violence has had on South Sudanese students were illustrated in the recent education needs assessments conducted across 400 learning sites in South Sudan, which found both higher dropout rates and lower enrollment and attendance rates than in the prewar years (Education Cluster 2018).

This study is part of an ongoing intervention to attend to the wellbeing of children and youth who are victims of direct violence or have witnessed intercommunal conflict. In the next section, we review the literature on the concept of wellbeing.

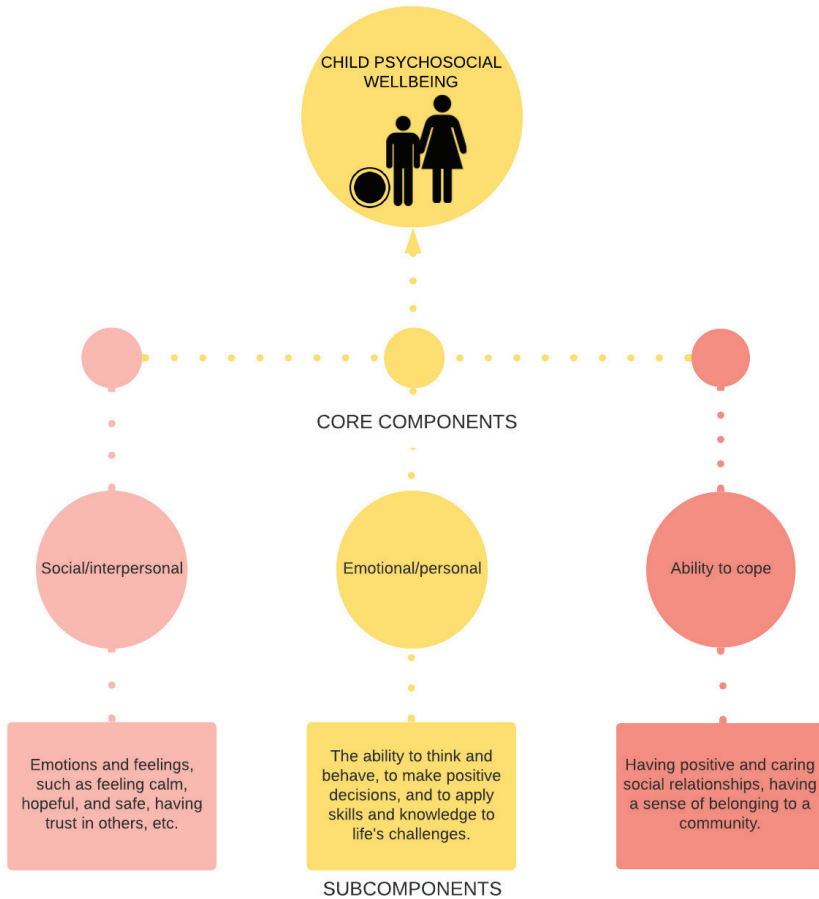
LITERATURE REVIEW

Our review of literature reveals a sparsity of validated quantitative instruments designed to measure wellbeing among sub-Saharan African children (Kabiru, Izugbara, and Beguy 2013). This lack sharply contrasts with the availability of guidance on designing instruments to measure children's psychosocial wellbeing in emergency settings (Bohl, Dzino-Siladjic, and Ryan 2018). We specifically identified the lack of a contextually relevant instrument that could be deployed in a linguistically and culturally diverse setting like South Sudan (Lu, Lim, and Mezzich 1995). Multiple systematic literature reviews of recent developments in interventions for children affected by armed conflict and political violence recommend understanding the effects of these interventions on a range of wellbeing outcomes (O'Sullivan, Bosqui, and Shannon 2016). This paper, which is our response to these recommendations, provides evidence of the systematic process through which we identified the wellbeing outcomes that are most relevant for studies among South Sudanese children and youth. We also describe our process of measuring the change in these wellbeing outcomes in this unique context. To situate our work, we first expound on the literature that helped us to operationalize the concept of wellbeing and the construct of psychosocial wellbeing, and their subconstructs. We then describe the kinds of interventions that typically produce these outcomes, thus situating the intervention we evaluated. We conclude the section with a description of the specific informed decisions we made during this study.

**OPERATIONALIZING WELLBEING, PSYCHOSOCIAL WELLBEING,
AND THEIR SUBCONSTRUCTS**

Wellbeing is generally perceived as a condition of holistic health (Bohl et al. 2018). This perception comprises such a broad range of physical, cognitive, mental, psychological, social, and spiritual states that it often complicates the operationalization of what we mean by wellbeing (Dodge et al. 2012). Therefore, we subscribe to the definition of psychosocial wellbeing found in the guidance on measuring children's psychosocial wellbeing, which states that psychological wellbeing refers to "the inter-connection between psychological sub-components—such as emotions, thoughts, and behaviors, including coping strategies—and social sub-components, such as interpersonal relationships, social roles, norms, values, traditions and community life, that contribute to the overall well-being of a person" (Bohl et al. 2018, 2). The guidelines also suggest that there are three subconstructs of psychosocial wellbeing, namely, emotional wellbeing, social wellbeing, and resilience. We define these as follows: emotional wellbeing is a person's internal state, as demonstrated through their emotions and feelings; social wellbeing is the nature of an individual's interactions with others; and resilience refers to an individual's ability to cope in an uncertain or changing environment, which stems from their sense of agency. In this paper, we focus on these subcomponents of children's psychosocial wellbeing.

Figure 1: Subcomponents of Psychosocial Wellbeing



Source: Extracted from Bohl et al. (2018)

INTERVENTIONS THAT INFLUENCE CHILDREN'S WELLBEING

This study is a component of ongoing interventions to support children and youth who are victims of direct violence or have witnessed intercommunal conflict. A review of the literature on the impact of interventions for children affected by armed conflict (Jordans, Pigott, and Tol 2016) suggests that learning environments can provide nurturing, supportive relationships and a sense of safety, which are key to recovery and can help distressed school-age children and youth acquire social and emotional skills (Alliance for Child Protection in Humanitarian Action 2020). Access to schools and CFS can help children build relationships with teachers, spend time in a safe and protected space, and learn key life skills

(Tol et al. 2011). This is consistent with the view that education mitigates the psychosocial impact of conflict and disasters by providing a sense of normalcy, stability, structure, and hope for the future (Convery, Balogh, and Carroll 2010; Nicolai and Triplehorn 2003; Alliance for Child Protection in Humanitarian Action 2020). In fact, a Global Education Monitoring report (UNESCO 2019) describes learning environments as safe spaces in which teachers observe the psychosocial wellbeing of children and teach coping skills that improve their social and emotional wellbeing by providing access to “therapeutic rapport,” which enables children to express emotion without experiencing a moralistic or judgmental response (Bosqui and Marshoud 2018).

A systematic review found that more than one-fifth of mental health and psychosocial wellbeing interventions in humanitarian settings involved the provision of CFS (Tol et al. 2011). In these environments, children and youth typically experience the programmatic interventions of social and emotional learning (SEL) and PSS. Although the two are often used interchangeably, there is a comprehensive treatise on the difference between SEL and PSS (see INEE 2016).

In this paper, we define SEL as activities that help learners gain the social and emotional competencies that enable them to recognize their feelings and emotions, and to manage them in a way that makes it possible for them to set and achieve positive goals. The aim of SEL is to enable them to do this without losing sight of the societal constraints in which they are situated and the need to be empathetic toward others (Collaborative for Academic, Social, and Emotional Learning 2003). One clear characteristic of SEL is that it is specifically designed to align with academic goals. Many studies have investigated what constitutes quality SEL programs (Aspen Institute 2018), how they affect students’ retention and attrition (Bridgeland, Dilulio, and Morison 2006), the benefits SEL brings to students’ academic performance (Zins et al. 2007; Payton et al. 2008), teachers’ impressions of its impact (Bridgeland, Bruce, and Hariharan 2013), and the benefits of SEL in marginalized settings, in public governance, and in social functioning (USAID 2019).

In contrast, PSS typically refers to a holistic system that recognizes how intrinsically connected people’s internal experiences are with their social perspectives, actions, and interactions with others. This holistic view discourages compartmentalizing the social, attitudinal, and aptitudinal aspects of a person’s wellbeing while emphasizing the need to view them within their broader environmental contexts (Action for the Rights of Children 2009). Thus, PSS has been defined as “processes and actions that promote the holistic wellbeing of people in their social world, including support provided by family and friends” (INEE 2010, 121). There is

strong evidence to support the view that PSS interventions have a positive impact in school-based settings.

The positive role PSS and SEL interventions play in recovery after a crisis is widely acknowledged (INEE 2016), and research suggests that they have both short- (Zins et al. 2007) and long-term (Elias et al. 2002) benefits. This primarily stems from the environments where PSS and SEL activities take place, such as CFS and learning spaces, where distressed students and youth can interact with their peers, communicate with trained instructors and trusted adults, participate in recreational activities designed to help them recognize and cope with their feelings, and, in the process, learn openness and social skills (see, e.g., Global Protection Cluster 2011). Efforts to achieve and measure these effects are described in the following section.

IMPLEMENTING AND MEASURING THE IMPACT OF PSS INTERVENTIONS IN CONFLICT SETTINGS

Conventional concepts of wellbeing have been measured indirectly using proxies based on observable, countable factors, such as personal income, physical health parameters, the local economy, etc. (Mguni and Bacon 2010). The use of indicators such as individual self-reports is now generally accepted, as they are a direct reflection of what people think, feel, and metacognitively reflect on (Steuer and Marks 2008). We identified some construct- and criterion-validated instruments used to measure wellbeing outcomes. Some of the scales and measures used in this study were initially developed for other contexts. They included the Strengths and Difficulties Questionnaire, the KIDSCREEN questionnaires, the Self-Description Questionnaire II, the Child and Youth Resilience Questionnaire, the California Healthy Kids Survey (California Department of Education 2007), the New Philanthropy Capital (2011) wellbeing measure, the UNICEF Actions for the Rights of the Children (ARC) Resource Pack (2009), the Rosenberg Self-Esteem Scale (RSES), and the New Economics Foundation (Thompson and Aked 2009) guide to measuring children's wellbeing.

The Child and Youth Resilience Measure is a measure of the individual, relational, communal, and cultural resources available to individuals to bolster their capacity to sustain wellbeing (Ungar and Liedenberg 2016). Originally designed to be used with 9- to 23-year-old youth, it was developed as part of the International Resilience Project implemented in 14 communities around the world. Three of those communities were in Africa: Serekunda, The Gambia; Njoro, Tanzania; and Cape Town, South Africa. The rest were in North America, Asia, and Europe

(Ungar and Liebenberg 2011). In a separate study, the measure was validated for a Canadian population (Liebenberg, Ungar, and Van de Vijver 2012). The KIDSCREEN questionnaires (Ravens-Sieberer and the KIDSCREEN Group Europe 2016) were developed through a collaborative effort of European pediatric researchers for use in epidemiologic public health surveys, clinical intervention studies, and research projects. To warrant cross-cultural applicability, different versions of the questionnaire were developed simultaneously in 13 European countries (Ravens-Sieberer et al. 2014). Although the 2014 KIDSCREEN technical document affirms the reliability, validity, sensitivity, and conceptually/linguistic appropriateness of this questionnaire in 38 countries/languages, only two African countries were included, Uganda and Kenya (Ravens-Sieberer et al. 2014, 799; Ravens-Sieberer and the KIDSCREEN Group Europe 2016, 46).

The Self-Description Questionnaire II is another well-validated instrument whose normative archive sample comprised Australian high school students who were predominantly ages 12-18 (Marsh et al. 2005). The Children's Hope Scale, used extensively in the United States, measures such constructs as a child's belief in their ability to achieve their goals, and to initiate and sustain movement toward those goals (Snyder et al. 1997). The RSES aims to measure a single dominant factor representing global self-esteem (a measure of how individuals value themselves) by measuring one's positive and negative feelings about oneself (Rosenberg 1965). According to the meta-analytic database from the Longitudinal Internet Studies for the Social Sciences, in studies that used the RSES from 1966 to 2016, only 4 percent of the samples included African contexts; the rest were predominantly in Europe, North America, and Asia (Gnambs, Scharl, and Schroeders 2018). Other scales specifically designed to measure both the feeling and functioning aspects of positive mental wellbeing have been widely reported in population surveys in the UK, including the national health surveys for England (Michaelson, Mahony, and Schifferes 2012) and New Zealand (Medvedev and Landhuis 2018).

Our point is that there is a strong evidence base for construct- and criterion-validated instruments that measure psychosocial wellbeing in high-income national contexts, whereas the vast majority of studies (approximately 92%) of the delivery and implementation of PSS interventions for vulnerable populations in conflict settings are situated in sub-Saharan Africa, the Middle East, and North Africa (Jordans et al. 2016). The modalities of these interventions are varied; most involve randomized control trials (Panter-Brick et al. 2018) that have a significant impact on children's social and emotional wellbeing and their ability to cope. The most frequent interventions were school-based and involved creative, expressive,

and psycho-educational and cognitive-behavioral strategies (O’Sullivan et al. 2016). Creative, expressive approaches in these settings emphasized interactive activities, such as drama, music, role-playing, and drawing (Jordans et al. 2016). PSS was the intervention reported to be delivered most frequently, partly because of its potential to recognize and strengthen resilience and local coping capacities (UNICEF 2009). A significant gap we found in the literature was the lack of reports on the development and effectiveness of contextualized measures in low- and middle-income settings, let alone in conflict settings (Kamali et al. 2020).

We also reviewed the literature for reports of contextualized instruments to measure the impact of PSS interventions on wellbeing outcomes. Several studies based in Northern Uganda described the development of a monitoring and evaluation tool to collect data on the process and outcomes of locally relevant and participatory intervention (Ager, Akesson et al. 2011; Claessens et al. 2012). These studies were deemed necessary, as previous scientifically validated questionnaires were judged to be inaccurate, due to their cultural and linguistic incongruencies with the Northern Ugandan population. The authors of one article suggested that existing instruments not only seemed unsuitable for the character of the interventions, they also failed to respond to practitioners’ need to identify social changes at the individual and group levels (Claessens et al. 2012).

To satisfy the need for cultural adaptability, a participatory research tool developed for a study in South Sudan was based on research conducted in Sri Lanka (Hart et al. 2007) and Sierra Leone (Stark et al. 2012). The tool was used to compare local perceptions of wellbeing and to determine the tool’s relevance and cultural fit (Eiling et al. 2014). A similar approach was used in Kenya (Kostelny, Ondoro, and Wessells 2014). Although these studies satisfied cultural relevance, the need for scientifically validated instruments remained. Impact studies conducted in Burundi (Jordans et al. 2013), Nigeria (Sheikh et al. 2014), and Rwanda (Chauvin, Mugaju, and Comlavi 1998) revealed the consistent challenges researchers face in interpreting cross-cultural validity and in validating measures to evaluate psychosocial wellbeing outcomes in these settings. These findings emphasize the need for instruments that are both culturally adapted and scientifically validated, and thus appropriate for evaluating the impact of PSS provided to children in low-resource communities.

In summary, our study focused on the measurement of psychosocial wellbeing outcomes using the subconstructs of emotional wellbeing, social wellbeing, and resilience. The clear impact school-based interventions have on students’ academic and wellbeing outcomes justify the setting of our study. As we sought to provide

rigorous evidence for the contextualization and validation of the instrument in a setting as unique as South Sudan, we reflected on the limitations of the existing scales and measures identified in the literature. Our work has clear implications for the methodology that should be adopted for the measurement of wellbeing in South Sudan and similar contexts. We demonstrate that how social wellbeing, emotional wellbeing, and resilience are understood is specific to our study setting.

PURPOSE OF THIS STUDY

As our review of the literature demonstrated, the impact of PSS on the social wellbeing, mental wellbeing, and academic outcomes of students living predominantly in North America and Western Europe is well established. While the literature on PSS and SEL in North America provides strong construct- and criterion-validity evidence of wellbeing measurement instruments, these instruments are primarily tested on and validated with sample populations in a nonemergency context. We argue that the study of wellbeing, PSS, and SEL demands careful consideration of the distinct way social wellbeing, emotional wellbeing, and resilience domains can be observed and measured in diverse and complex settings, such as South Sudan. To achieve the overarching objective of this research project—that is, to measure the effectiveness of PSS interventions on children’s wellbeing in South Sudan—a more contextually aligned view of the instrument design, its domains, and its interpretation is needed.

IMPLEMENTING PSYCHOSOCIAL SUPPORT INTERVENTIONS IN SOUTH SUDAN

There is evidence that children in conflict-affected settings are more likely to start school with lower levels of social-emotional skills (Ursache, Blair, and Raver 2012). Additional research suggests that school-based interventions help to build children’s early social-emotional competencies, such as behavior regulation (Blair 2002), attention regulation, and problem-solving (Diamond and Lee 2011), all of which affect their academic outcomes (McCormick et al. 2015). As a result, the US Agency for International Development (USAID) requested an impact evaluation study in South Sudan to test these assumptions and build an evidence base for including PSS in education programs in conflict settings, with a particular emphasis on isolating the impact PSS has on academic skills, such as reading and math.

Since 2015, UNICEF South Sudan has attempted to implement PSS activities in schools throughout the country. A training manual of PSS resources that UNICEF developed for use in CFS, schools, and communities offered instructions for conducting activities at a variety of age levels, which were centered around play, learning, and wellbeing. Due to the wide variation in learning environments and in the age of students participating in the programming, and the fact that implementation in South Sudan is carried out through several subimplementing partners involved in relief interventions, there is no unified curriculum that fits the needs of all learners. Many of the PSS activities were carried out in temporary learning spaces or CFS developed by UNICEF, which trained teachers to implement PSS interventions across much of the country. Several nongovernmental organizations (NGOs) also trained PSS facilitators. Most of these organizations employed consultants who were either South Sudanese or international workers. The psychologists and practitioners engaged by NGOs also were both local and international. However, high attrition rates among teachers resulted in inconsistent implementation of the program.

One of our authors participated in and observed a PSS training session offered by UNICEF. The main PSS activities were designed to help teachers identify children with unique needs and problems and to support them in dealing with their grief, suffering, loss, and a gradual return to normalcy. These activities were categorized into seven themes: creative, imaginative, physical, communicative, manipulative, cultural/traditional, and participatory.

- *Creative activities* were designed to help children express their feelings and ideas. Activities included painting, drawing, clay molding, making dolls, etc.
- *Imaginative activities* were intended to help children develop creative social skills, and to gain an understanding or make sense of what happened or is happening in their lives. Activities included dance, theatre/drama, music/singing, role-play, etc.
- *Physical activities* were conceived to help children develop self-confidence and motor skills, and to facilitate peer interactions. Activities included football, volleyball, outdoor team games, and traditional games.
- *Communicative activities* aimed to help children express their feelings in words and to discuss important issues in their lives. Activities included stories read from books and oral storytelling, reading, and focus group discussions and debates.

- *Manipulative activities* had the goal of improving children’s problem-solving and cognitive skills. Activities included doing puzzles, using building blocks, and molding clay.
- *Cultural/traditional activities* were intended to help children appreciate their own culture and to give them a sense of being part of the community, despite what they had gone through. Activities included dancing, singing, traditional games, storytelling, poetry, etc.
- *Participatory activities* were developed to enhance children’s and youths’ resilience and adaptability, create good relations among the children, give them a sense of civic responsibility, and help them develop cognitive functioning. Activities included learning life skills such as reading and numeracy, landmine awareness, health education, and joining in community events.

All the PSS activities were categorized according to the children’s ages and, where possible, gender.

Teachers were trained to identify and respond to students who were experiencing distress while at school, and to offer what could be termed psychological first aid; they were instructed to refer serious cases to specialists. The implementing agencies and their partners constructed CFS in or near schools where the PSS activities could be carried out. The CFS in South Sudan were largely structured to handle “relatively short to medium-term program responses. They are very often operated from tents and/or temporary structures (e.g. in schools, under a tree or a vacant building)” (Davis and Iltus 2008, 9). Many CFS in South Sudan were set up to enhance what the formal learning spaces offered. This involved providing key PSS/SEL interventions in a context where students and teachers had been under attack by armed forces, and were subjected to sexual and gender-based violence, forced recruitment, and other threats. While regular classroom teachers were not expected to be equipped to deliver PSS/SEL services, teachers working in CFS were uniquely trained to deliver these interventions to children and youth affected by conflict and crisis.

Displacement often brings large numbers of children into local classrooms. CFS were set up to provide spaces where the schools could run double shifts, and thus be able to provide all children with learning opportunities. Communities also created CFS to provide nurturing environments where children could enjoy both free and structured play, recreation, leisure, and learning activities. CFS, which are

designed and operated in a participatory manner, also provide health, nutrition, psychosocial support, and other activities that restore children's sense of normality and continuity. In South Sudan, the children's local language, ethnic make-up, and education level influence which social skills, emotional competencies, and psychosocial supports are provided.

Since 2015, more than 560,000 South Sudanese children and youth have received PSS through the Integrated Essential Emergency Education Services, a USAID-funded program that was implemented by UNICEF with the aim of reaching South Sudan's 2.2 million out-of-school children and youth. The authors of this paper were part of a multi-institutional consortium of multidisciplinary partners who evaluated the impact of these interventions in order to inform the future allocation of resources. The overarching objective of the research project was to investigate the impact of the PSS intervention on the wellbeing and academic outcomes of the South Sudanese children who received it. Thus, our larger research team set out to test the theory that children who receive PSS and SEL interventions will have a greater sense of wellbeing, as well as higher math and literacy outcomes, than their peers who did not receive the intervention. The study we describe in this article, which was developed as a subset of the larger study, specifically aimed to develop a more contextually aligned instrument that we could use to measure the impact of these interventions on local perceptions of wellbeing outcomes.

RESEARCH QUESTIONS

Our study was guided by the following research questions:

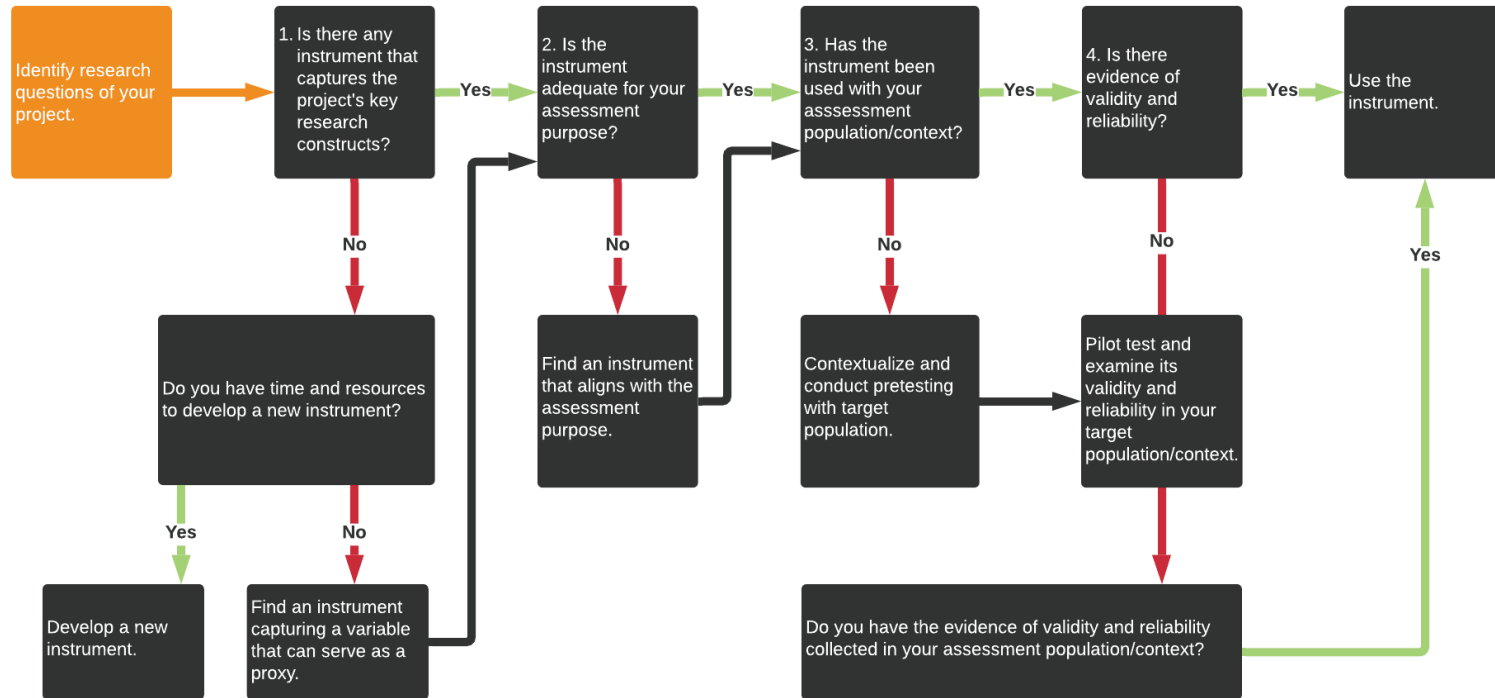
- What is a relevant and inclusive process for teams to follow to identify constructs and questions and to test the adaptation of instruments to measure the wellbeing of students in South Sudan or similar contexts?
- What is the evidence of the validity of an instrument that was adapted to measure the wellbeing of students in South Sudan? What information do confirmatory and exploratory factor analyses provide for understanding wellbeing in conflict settings?
- What is the structure of the wellbeing domains when measured in the specific context of South Sudan?

METHODOLOGY

FRAMEWORK

The instrument we developed for this study follows the recommendations provided in *Measure Guidance: Choosing and Contextualizing Assessment Measures in Educational Contexts* (Diazgranados Ferráns and Lee 2019). The procedure outlined in this guidance document specifies five consecutive steps for instrument development: (1) identify the key research questions, (2) identify an assessment that matches the needs of the research, (3) review evidence of the validity and reliability of existing instruments for the target population in the setting of interest, (4) contextualize the instruments to meet the specific contextual needs, and (5) conduct a validation study. The guidance document also presents a decision tree (Figure 2), with recommendations for alternative steps to take if the requirements of any step are violated.

Figure 2: A Decisionmaking Tree to Guide the Process of Choosing and Contextualizing Measures in Unique Contexts of Conflict and Crisis



Source: Extracted from Diazgranados Ferrás and Lee (2019)

INSTRUMENT DEVELOPMENT

Following the five steps outlined here, we present our approach to the development of the instrument we used in this research (Figure 3). A consortium of researchers and stakeholders from North American and African universities and NGOs, with representatives from USAID, the USAID South Sudan Mission, and UNICEF, met at a workshop in early 2019 to develop a common understanding of the objectives of this research project. At the meeting, we consulted with our team's experts on psychosocial wellbeing measurement to discuss the availability and suitability of instruments to measure specific aspects of child psychosocial wellbeing in emergency settings. Using existing guidance (Bohl et al. 2018; Ager, Ager et al. 2011), we itemized three broad measurement domains: emotional wellbeing (comprising emotions/feelings and behaviors), social wellbeing, and the ability to cope (resilience and skill-building). Researchers and research associates who were indigenous to South Sudan and had intimate knowledge of the population commented on the local conceptualization of these wellbeing outcomes. Their comments were crucial to our final selection of instruments to measure these outcomes.

Eleven measures and instruments were originally presented for consideration:

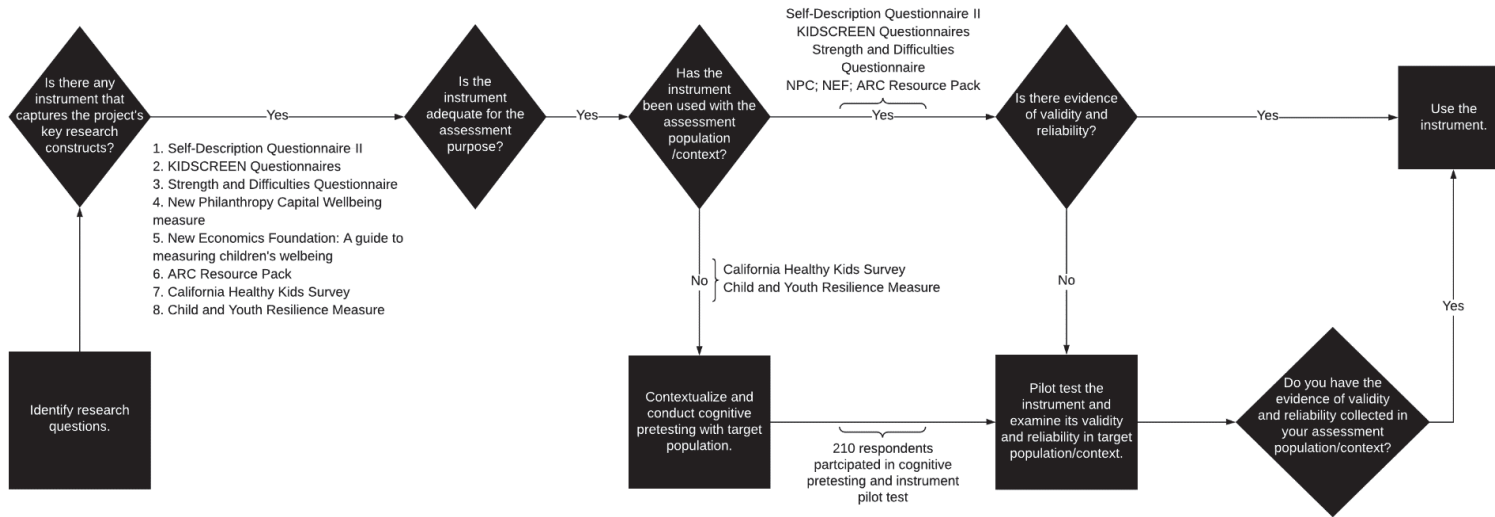
- The California Healthy Kids Survey (California Department of Education 2007)
- The Child and Youth Resilience Questionnaire (Ungar and Liedenberg 2016)
- The KIDSCREEN questionnaires (Ravens-Sieberer and the KIDSCREEN Group Europe 2016)
- The Multidimensional Students' Life Satisfaction Scale MSLSS (Huebner et al. 1998)
- The New Philanthropy Capital wellbeing measure (2011)
- The New Economics Foundation guide to measuring children's wellbeing (Thompson and Aked 2009)

- The Resilience Scale (Wagnild and Young 1993)
- The Rosenberg Self-Esteem Scale (Rosenberg 1965)
- The Self-Description Questionnaire II (Marsh et al. 2005)
- The Strengths and Difficulties Questionnaire (Goodman 1997)
- The UNICEF ARC Resource Pack (2009)

These 11 were chosen because they are the ones most commonly used as quantitative measures with children, for their validity and reliability, and because they measured the subdomains being evaluated. Measures that had not been tested in similar contexts were eliminated. Those that would overlap with the subdomains targeted in the UNICEF PSS activities were selected.

Participants discussed the cultural appropriateness of the different measures and instruments for the South Sudanese population. Three instruments were excluded (MSLSS, RSES, and the Resilience Scale) based on relevance, and on the local experts' and implementing partners' knowledge of and experiences in the South Sudan context. For instance, the RSES focuses mainly on personality, such as an individual's feelings about him- or herself, while the MSLSS looks at children's satisfaction across six subdomains—satisfaction with school, family, friends, living environment, self, and overall life satisfaction. These constructs are captured in the instruments adopted. To avoid duplicating questions and to keep the questionnaire a reasonable length, we decided to exclude the MSLSS, RSES, and the Resilience Scale instruments. Moreover, it's generally known that children in conflict-affected areas often experience abuse and may avoid talking about family and personal matters. Such discussions could bring back painful memories (Bohl et al. 2018), and with the low self-esteem children in such situations often exhibit, it was judged prudent to exclude the three instruments in question. Thus, three measurement outcomes and eight instruments were judged relevant to the study participants. Consequently, the first and second steps of the measure guide were fully satisfied.

Figure 3: Adapted Decisionmaking Tree to Guide the Process of Choosing and Contextualizing Measures of Wellbeing in South Sudan



Source: Adapted from Diazgranados Ferrás and Lee (2019)

The third step of the decision tree resulted in a split response. Some of the instruments identified in step two were being used by the implementation partners as part of their programmatic formative assessment, whereas the other identified instruments had not been used. However, there was no evidence of the validity of either set of instruments for South Sudanese children and youth. Thus, with insights and guidance from the South Sudanese researchers who had a firsthand understanding of the context and its population, and of the challenges that may arise from translations into different languages, the likelihood of survey fatigue from filling out long questionnaires, and the appropriateness or seeming complexity of certain terms and items for different age groups, we proceeded to adapt an instrument from the existing ones.

CONTEXTUALIZATION OF THE INSTRUMENT

South Sudan is a multilingual republic. Most of the residents speak English, Juba Arabic, Nuer, or Dinka, depending on their location and ethnic affiliation. In conducting the instrument designs, our research team also noted other languages; in Juba, for example, some students in the sampled schools spoke Acholi, Balanda, Anuak, and other indigenous languages. However, the majority in the states we covered spoke the seven languages mentioned previously. We translated the instrument into these seven languages and then translated it back to English to see if it retained its intended meaning. We then conducted pretesting and cognitive interviews. Research associates and enumerators asked students to explain to them how they understood the instruments. The pilot testing took one day and was conducted in three sites, one Juba POC school and two Juba non-POC schools. It involved approximately 70 students per school, and 210 students across grades three (P3) and eight (P8). Although students in grades six, seven, and eight were included in the pilot study, the research team experienced some difficulty in finding students in the higher grades across the schools. This was in part due to the prohibitively high dropout rate of students at higher grade levels. Consequently, the final study focused on students in grades 3 (P3) and 6 (P6).

We categorized the final survey items into three sections. Section one (10 items) collected demographic information and measures of students' home resources, like the frequency of meals. Section two (20 items) was a general wellbeing section intended for all respondents, and section three was a six-item section with questions specifically for adolescents. In total, 26 items specific to the measurement of wellbeing were adapted from validated instruments we found in the literature, from UNICEF's monitoring and evaluation tool, and from concepts the South Sudanese researchers identified as relevant to the context (see Table 1 for a full list). All 26

items were worded and scored on a four-point agreement Likert scale (1, “strongly disagree”; 2, “disagree”; 3, “agree”; and 4, “strongly agree”). The questionnaires contained no neutral responses.

With these items developed and contextualized for the specific population, the final step of the decision tree was to investigate whether the newly adapted instrument we had developed actually measured the study’s intended factors of interest. We conducted a factor analysis test to obtain this evidence, and the findings are the focus of this paper.

We employed an iterative set of confirmatory factor analyses (CFA) and exploratory factor analyses (EFA). We began with a CFA, using the constructs of emotional wellbeing, social wellbeing, and resilience, and their associated variables in the surveys they were drawn from (Table 1). Because the model did not fit well in the South Sudanese context, we then conducted an EFA to suggest more fitting models and, finally, a CFA to confirm the new model in this context. Below we describe the specific variables that we moved from their original constructs.

Table 1: Items of the PSS Wellbeing Student Survey

Item	Item Description	Code
<i>Demographics</i>		
School Name	Name of school	None
Class	Which class are you in?	1 = “P3” 2 = “P6”
Gender	Are you a boy or girl?	1 = “Male” 2 = “Female”
Age	How old are you? Please write age in years.	None
Mother Tongue	What is your mother tongue? Please write in the space provided.	
Language of Instruction	In what language do you learn in school?	1 = “English” 2 = “Arabic” 3 = “Other (write here)”
Time Spent in School	How long have you been in this school? Please tick one choice.	1 = “Less than 1 year” 2 = “1 to 3 years” 3 = “3 to 5 years” 4 = “5 or more years”

MEASURING WELLBEING AMONG CHILDREN IN SOUTH SUDAN

Item	Item Description	Code
Breakfasts/ Week	How many times do you eat breakfast in a week?	1 = "Not at all" 2 = "1 to 3 days per week" 3 = "4 to 6 days per week" 4 = "Every day of the week"
Lunches/Week	How many times do you eat lunch in a week?	
Suppers/Week	How many times do you eat supper in a week?	
<i>Wellbeing Questions</i>	<i>In the last two weeks, have you...</i>	
Worry	...been worried about anything?	Self-Description Questionnaire II
Calm	...been able to calm yourself down when you are upset or angry?	CONTEXT
Teacher Listening	...felt that your teacher listened to you and respected your ideas?	KIDSCREEN Questionnaires
Sadness	...felt sad?	Strength and Difficulties Questionnaire
Mood Understanding	...been able to understand your moods or feelings?	California Healthy Kids Survey
Bullying	...been bullied in school?	Strength and Difficulties Questionnaire
Dispute Resolution	...been able to find friendly ways to solve misunderstandings or disputes?	Child and Youth Resilience Measure
Concentration in Class	...been able to concentrate or pay attention in the classroom?	Strength and Difficulties Questionnaire
School Environment	...felt that your school is a nice place to be in?	New Philanthropy Capital
Someone to Trust	...felt that you had someone you trust to help you when you were in need?	New Economic Foundation (NEF)
Do Free Time	...been able to do the things you wanted to do in your free time?	KIDSCREEN Questionnaires
Good Mood	...been in a good mood?	KIDSCREEN Questionnaires; NEF
Time with Friends	...spent time with your friends?	KIDSCREEN Questionnaires

Item	Item Description	Code
Helping Others	...been helpful to others?	Strength and Difficulties Questionnaire
Lost Temper	...gotten angry and lost your temper?	Strength and Difficulties Questionnaire
Feeling Helpless	...been in situations where you felt helpless?	KIDSCREEN Questionnaires
Understanding Others	...tried to understand how others feel?	California Healthy Kids Survey
Parents Listening	...felt that your parents/guardians listened to you and respected your ideas?	ARC Resource Pack
Suggest Games	...suggested activities or games to do with your friends?	KIDSCREEN Questionnaires
Bad Dreams	...had bad dreams?	Strength and Difficulties Questionnaire
Adolescent Questions		
Acting Responsibly	In the last two weeks, have you had opportunities to show others that you can act responsibly?	CONTEXT
Feeling Pressure	In the last two weeks, have you felt under pressure?	KIDSCREEN Questionnaires
Care of Tasks	In the last two weeks, have you done well taking care of your tasks at home or at school?	CONTEXT
Confidence during Hard Times	Do you believe that your confidence or trust helps you to get through hard times?	CONTEXT
Things Me Good	A lot of things about me are good.	New Philanthropy Capital
Friends Stand By	My friends stand by me during difficult times.	Child and Youth Resilience Measure
If I Try	If I really try, I can do almost anything I want to do.	Self-Description Questionnaire II

Note: Self-Description Questionnaire II (Marsh et al. 2005), KIDSCREEN Questionnaires (Ravens-Sieberer and the KIDSCREEN Group Europe 2016), Strength and Difficulties Questionnaire (Goodman 1997), California Healthy Kids Survey (California Department of Education 2007), Child and Youth Resilience Measure (Ungar and Liedenberg 2011), New Philanthropy Capital (2011), New Economics Foundation (Thompson and Aked 2009), ARC Resource Pack (UNICEF 2009)

SAMPLING

We employed a two-stage sampling strategy. First, we purposively sampled 64 schools from 5 states (Central Equatoria, Western Equatoria, Upper Nile, Unity, Jonglei) in the Republic of South Sudan that had the potential to provide both treatment and control schools. To provide a sample that included diverse locations, we chose the five states in consultation with UNICEF, the implementers, and members of the USAID South Sudan Mission. We chose the five states based primarily on the areas where USAID had been implementing its activities for the longest time. Accessibility was another major consideration. Other parts of the country were receiving similar interventions but for a shorter time, so we were concerned that we would not detect an equally strong effect.

In line with the “do no harm” principle of emergency education settings (European Commission 2019), we secured access to the research sites only after the South Sudan Ministry of General Education and Instruction (MoGEI), which served as the review board for the study, had granted permission. With the notable exception of the demographic data described in the previous section, we did not collect any identifying data from participants of the quantitative study.

The final school selection comprised 36 treatment schools that had received the PSS intervention at the time of the study (targeted teacher training by UNICEF and implementing partners) and 28 control schools that had not received it but were located relatively close to the treatment schools or shared a similar profile. It is worth noting that all the POC schools were treatment schools. Teachers in most of the schools (even control schools) had previously been trained in PSS.¹ What differentiated the treatment schools from the control schools was the inclusion of a class or session during the school day where children would have participated in more dedicated and purposeful activities.

Second, we selected 2,982 students and 580 teachers, including on average 10 teachers per school, 15 students randomly sampled from grade P3, and 35 students randomly selected from grade P6. We selected these grades to provide a mix of younger and slightly older students who had the literacy skills to complete the surveys. The students were randomly selected to complete the PSS outcomes survey we review in this article, and the teachers were required to answer a questionnaire.

¹ Although we also gathered data on the teachers, their characteristics are not within the scope of this paper.

Of the 2,982 students randomly sampled for this study, 40.3 percent were in the control schools. Boys accounted for about 60.7 percent of the participants (Table 2). Grade P3 respondents were between ages 8 and 22, with a median and modal age of 13, while grade P6 respondents were between ages 10 and 37, with a median and modal age of 16. In our sample population, 96 percent of the male respondents and 99.5 percent of the female respondents fell within the age range of 8 to 20 (Table 3). As previously noted, in this factor analysis we focused only on data from students' responses to questions about their wellbeing.

Due to the overall age distribution of primary school students in South Sudan and the high rate of overage students, the sample of student respondents spanned a broad range of ages—8 to 34, with a median age of 13 in P3 and 16 in P6. The age breakdown of primary and secondary school students in South Sudan is significantly affected by several factors, including late entry into school, migration or displacement, and frequent school closures or interruptions due to conflict. As of 2016, nearly 90 percent of students in South Sudan's primary schools were considered overage, and 93 percent of secondary students (MoGEI 2017). In the later primary and secondary school years (P6-P8), the population of significantly overage students (defined as more than five years over age) is more than 50 percent for boys. The percentage of significantly overage girls declines at that point, as they become more likely to drop out due to early marriage or pregnancy. While this trend is true across the nation, it is particularly evident in the most vulnerable and conflict-affected states, where conflict frequently interrupts schooling and efforts are made to reintegrate former child soldiers back into school (Skårås 2017). For these reasons, we decided not to drop the overage students' data from our study. However, we suggest that readers interpret the results of this study as tentatively generalizable to this broad age range.

Table 2: Frequency Distribution of Survey Respondents by Gender, Location, and Class

Category	Groups	Frequency	Percentage	Age (in years)		
<i>Students (n=2982)</i>		Min	Max	Median		
<i>Gender</i>	Male	1750	60.7			
	Female	1131	39.3			
<i>Class</i>	P3	986	33.1	8	22	13
	P6	1996	66.9	10	37	16
<i>Intervention Status</i>	Control	1201	40.3			

Category	Groups	Frequency	Percentage	Age (in years)		
<i>Location</i>	Treatment	1781	59.7			
	Unity	1042	34.9			
	Jonglei	780	26.2			
	Upper Nile	251	8.4			
	Western Equatoria	420	14.1			
<i>Schools (n=64)</i>	Central Equatoria	489	16.4			
<i>Intervention Status</i>	Control	28	43.8			
	Treatment	36	56.2			

Table 3: Age Distribution of Survey Respondents by Gender

Gender		Male	Female	Frequency	Percentage
		Frequency	Percentage		
<i>Age Distribution (in years)</i>	8-20	1674	95.7	1125	99.5
	21-30	67	3.8	5	0.4
	30-37	9	0.5	1	0.1

DATA COLLECTION

To facilitate data collection in the field, we printed paper copies of the survey questionnaires. The surveys were administered by research associates (RAs) who were either members of the community or had a general knowledge of the community’s cultural norms. They were supported by enumerators who were native speakers of the indigenous languages spoken in the study sites. The RAs and enumerators were trained twice before each data-collection phase. They performed mock demonstrations during the training sessions and at the piloting stage. No RA or enumerator was allowed to work alone. Each data-collection team consisted of one RA and one or two enumerators. The teams traveled to peri-urban and rural areas to survey treatment schools and corresponding control schools. During the data collection, questions that had been identified as difficult to understand during the pilot testing were explained in the local languages.

Participation was voluntary. Participants were informed that they could choose not to participate in any activity they did not properly understand or felt uncomfortable with. If a student felt like stopping at any time, she or he was permitted to do so without being pressed for their reasons. The aims, objectives, and significance of the study were explained to the participants, and those who did not raise their hands to volunteer were not forced to. Where necessary, teachers from the participant schools were involved to allay students' suspicions, providing their presence did not appear to make the students uneasy. All the supports crucial for the success of the study and to comply with the "do no harm" policy (Bonis Charancle and Lucchi 2018) were strictly followed.

The support, translation, and supervision the enumerators and RAs provided were very effective, and we ultimately achieved the high response rate of 99.16 percent. All data collected in the field were scanned and saved in a secure repository. The scanned files were then entered into Qualtrics (an online survey tool) and reviewed by an assigned data custodian prior to the analysis that was conducted by the multinational team and reviewed by an assigned data custodian.

RESULTS

VALIDITY TESTS

We used both EFA and CFA in this study. EFA is particularly useful in investigating latent variable structures from continuous data (Costello and Osborne 2005), as it examines assumed underlying characteristics of items/variables through correlation patterns between them. However, since we had a theoretical model recommended by our expert on psychosocial wellbeing measurement, we began by conducting a CFA using the three suggested latent constructs: intrapersonal, interpersonal, and ability to cope factors (Henson and Roberts 2006). After determining that the model functioned poorly in this context with this population (results of the original CFA did not meet thresholds for model fit, with RMSEA=0.068; TLI=0.642), we then used EFA to analyze the data collected from the survey instruments for the latent factor structure (see parallel analysis scree plot, Figure 5). This process revealed possible factor models that we evaluated using indices of fit and procedural recommendations for the elimination of poor models (Bandalos 2018). Having eliminated poor factor models, we then used an acceptable model in the next stage of CFA. We employed an iterative process of consulting theory to make informed decisions about our factor structures and to

connect the well-fit models to theoretical support for the psychosocial constructs they were measuring.²

Following recommendations about EFA and CFA data preparation from the methodological literature (Costello and Osborne 2005), we first randomized our data before splitting it into two halves. We used the first half of the data (1,400 cases) for EFA and the second half (1,582 cases) for CFA.

CONFIRMATORY FACTOR ANALYSIS OF THE THEORETICAL THREE-FACTOR MODEL

For CFA, we analyzed two three-factor models for fit and convergence. The first model was our initial structure, which, based on the literature, assumed three factors. Note that these loadings (shown in Figure 4) differed from the EFA-informed three-factor model.

We used a latent variable analysis (Lavaan) package for the analysis. Since the existing factor structure did not meet thresholds for model fit, we then conducted EFA to modify our latent structure.

² While procedures suggested conducting a multivariate and univariate normality test prior to the factor analysis, we observe that our four-point Likert response scales are on an ordinal scale and would not align with this assumption.

Figure 4: Confirmatory Factor Analysis Path Diagram of Theoretical Three-Factor Model

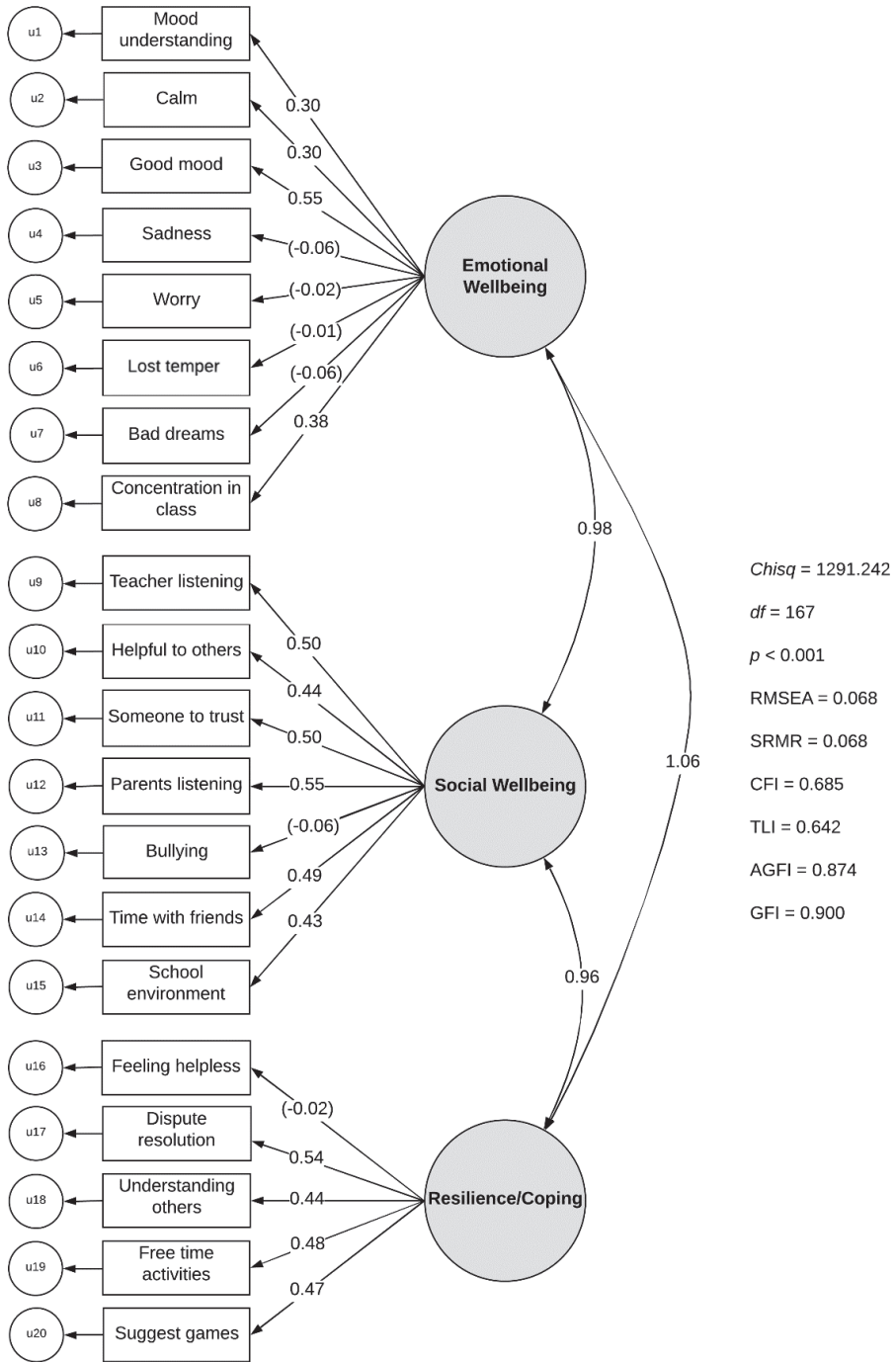


Table 4: Confirmatory Factor Analysis of Theoretical and EFA-Informed Model

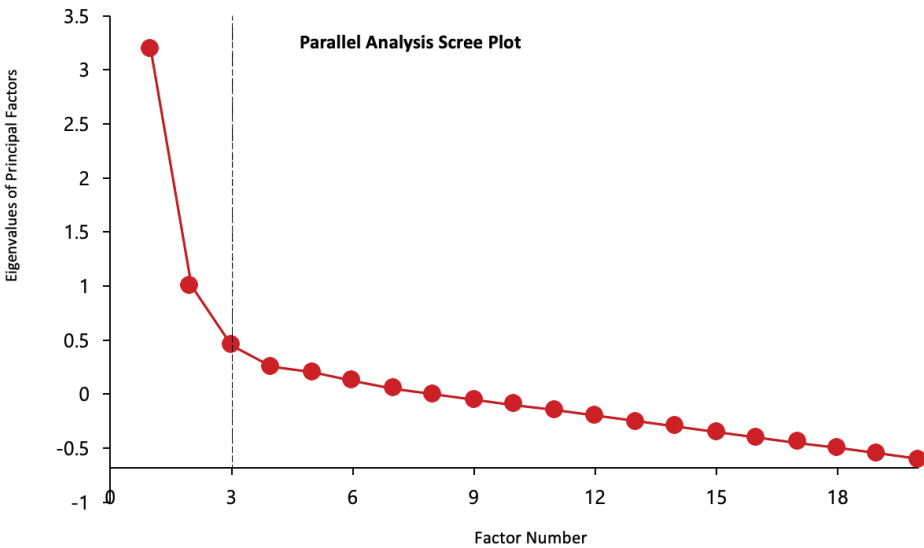
Models	2	df	2/df	p	RMSEA	TLI	CFI	GFI	AGFI
<i>Thresholds for Acceptable Fit</i>			≤ 5.00	≥ 0.05	≤ 0.05	≥ 0.80	≥ 0.85	≥ 0.80	≥ 0.85
Theoretical Model	1291.242	167	7.732	<0.001	0.068	0.642	0.685	0.900	0.874
EFA Model	629.703	167	3.771	<0.001	0.043	0.853	0.870	0.956	0.945

Note: Recommendations for acceptable fit are based on Hu and Bentler (1999). RMSEA, root mean square error of approximation; TLI, Tucker-Lewis index; CFI, comparative-fit index; GFI, goodness-of-fit index; AGFI, adjusted goodness-of-fit index; NFI, normed-fit index; BIC, Bayesian information criterion.

EXPLORATORY FACTOR ANALYSIS

Because the established model did not function well, we explored potential two-, three-, four-, and five-factor solutions to identify a statistically and substantially viable alternative model. Each of these factors was examined for fit. In the two-factor structure, factor 1 comprised 14 items, all with single-factor loadings ranging from 0.3 to 0.5, and factor 2 had 6 items with single-factor loadings from 0.4 to 0.5. However, the fit was poor (RMSEA=0.039; TLI=0.886). The four- and five-factor model analyses revealed insignificant variable loadings and were discarded.

Figure 5: Scree Plot



For the three-factor model (see Table 5, informed by the scree plot), factor 1 was composed of nine items with single-factor loadings from 0.35 to 0.55. Factor 2 comprised five items with factor loadings from 0.36 to 0.51. Factor 3 had five items with loadings from 0.39 to 0.46. The model also had acceptable fit indices (see Table 5; RMSEA=0.029; TLI=0.939). The EFA suggested that a three-factor model was indeed the most fitting structure, although the loading differed slightly and in informative ways from our initial theoretically prescribed model.

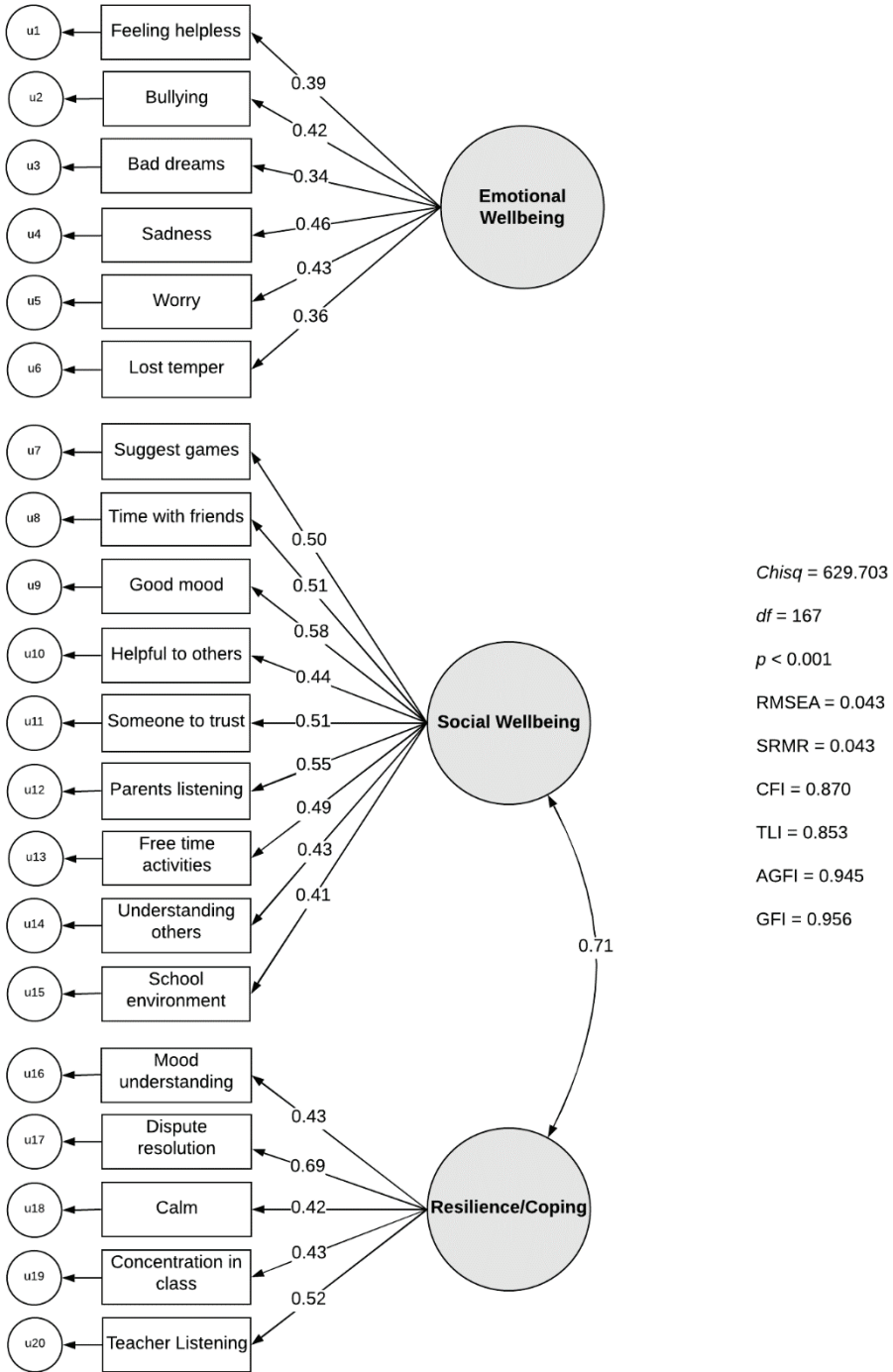
Table 5: Exploratory Factor Analysis of the General Wellbeing Items (RMSEA=0.029; TLI=0.939)

Items	Factor	Dimension		
	1	2	3	
Good Mood	0.52	-0.01	0.06	Social Wellbeing
Do Free Time	0.53	-0.00	0.04	
Time with Friends	0.54	0.03	-0.09	
Someone to Trust	0.55	-0.04	-0.05	
Suggest Games	0.52	0.00	0.02	
Parents Listening	0.45	-0.12	0.10	
Helping Others	0.48	0.03	0.08	
Understanding Others	0.38	0.12	0.13	
School Environment	0.35	-0.09	0.15	
Calm	-0.03	0.03	0.48	Resilience/Coping
Dispute Resolution	0.14	0.02	0.51	
Mood Understanding	-0.04	0.00	0.47	
Concentration in Class	0.21	-0.01	0.36	
Teachers Listening	0.20	-0.05	0.36	
Feeling Helpless	0.15	0.45	-0.12	Emotional Wellbeing
Sadness	-0.11	0.46	0.10	
Lost Temper	0.05	0.44	-0.11	
Bullying	-0.11	0.42	0.09	
Bad Dreams	0.02	0.40	-0.02	
Worry	-0.06	0.39	0.14	

Note: Extraction method; maximum likelihood; Rotation method; Oblimin with Kaiser normalization. Loadings larger than 0.30 are in bold

The EFA-informed three-factor model showed a better model fit (as noted in all fit metrics given in Table 4).

Figure 6: Confirmatory Factor Analysis Path Diagram of EFA-Informed Three-Factor Model



SUMMARY

After our iterative factor analysis, the questions still fell into the three commonly used categories or domains. However, as a result of our CFA of the three-factor model, some of the questions no longer matched with their expected factors. As noted in Figures 4 and 6, three questions that were expected to measure resilience/coping skills (“understanding how others think or feel,” “doing things you want in your free time,” and “suggesting activities with friends”) were instead matched to the social wellbeing factor. This added to the number of items in this factor, and it also suggests that the questions were understood in relation to others, rather than as individual experiences or skills.

The question asking whether a child “had been bullied or picked on” moved from the social wellbeing factor to the emotional wellbeing factor, suggesting that being bullied is perceived much more as having individual impact, rather than as how one relates with others. Additionally, “feeling helpless,” which would reflect a child’s self-efficacy and ability to reach out to others, was moved from the resilience factor to emotional wellbeing. This fits well conceptually, as the question focuses on feelings of helplessness (in other words, a state of being), rather than acting on the feeling, as a behavior or skill.

The third factor, which we labeled resilience, gained “teacher respected you” from the social wellbeing factor, as well as “concentrating in class,” “calming yourself,” and “understanding your mood,” all of which were former subdomains of emotional wellbeing. This shift, combined with the remaining item, “finding friendly ways to solve problems,” transitioned from a more general resilience or skills-for-life domain to a more specific domain, which we now call self-regulation. This self-regulation factor seems to relate more specifically to skills that can be applied in a classroom setting. We will investigate this construct further in future work.

The social wellbeing factor has become more broadly inclusive, perhaps indicating that students related the questions more strongly to their social interactions than to individual actions or reactions. The focus in the resilience domain (skills and knowledge) on self-regulation skills consolidates positive skills that one might learn in PSS activities, with the exception of “teacher listened to me and respected my opinions.” This raises the question of whether teacher behavior toward a student connects to their use of their skills. One possibility is that each question in this factor was understood by students specifically in relation to their school environment, and that teachers are included in that environment. It is also possible

that these are skills teachers desire, and that a teacher is more likely to listen to a student who is able to use them. The inverse is equally possible—that a teacher listening to a student will encourage them to use these skills. Given the intervention’s focus on training teachers in PSS, this would be reasonable. Where the skills are tied to interactions with the teacher, this would suggest that the self-regulation domain is also closely tied to social interaction.

DISCUSSION

The original construction of the instrument drew questions from a set of widely used and tested instruments. These instruments contained questions that reflected the three domains and measured the subdomains that were identified by our multinational, multi-institutional working group as most relevant to the children surveyed (see Table 1). We adapted the questions from the instruments, from the UNICEF IEEF evaluation form, and from context-specific questions we constructed from recommendations made by the South Sudanese researchers on our team, including questions about feeling under pressure and taking care of responsibilities.

Our work has both methodological and conceptual implications for scholarship in the area of measurement in education in emergencies. We demonstrate in this entire study the importance of a methodological approach that requires local leadership, a review and revision of questions even in established instruments before they are distributed, and a validity check when using an instrument in a new environment.

Our work shows that wellbeing factors look slightly different in the South Sudanese context than what we find in the literature. We see that a number of specific coping skills are viewed as relevant to the individual, but also in social interactions and processes. These coping skills influence the way students see themselves relating to others. This suggests the importance of social connections in South Sudanese students’ conceptions of their wellbeing, and of the fact that PSS and the behaviors it supports are taught in the communal context of schools and child-friendly spaces. This adjustment fits with the value children place on types of learning in conflict settings beyond school subjects, including practical knowledge of social norms (Winthrop 2011).

The other major shift we see is greater specificity in the resilience factor in what we now call self-regulation. For students in South Sudan, the broad concept of resilience is understood specifically in the context of the school or CFS as skills that can be applied with knowledge gained from PSS activities. This construct of self-regulation may in fact be more precisely supportable in PSS interventions. Our work clearly shows that the way wellbeing is understood by students in South Sudan is closely related to the context of where their skills are built and practiced, and with whom they are practicing these skills and knowledge.

LIMITATIONS, AND IMPLICATIONS FOR POLICY AND FUTURE WORK

Although the team used well-accepted measures and local experts to inform the construction of the tool, the changes found in the factor analysis may demonstrate that these measures still do not fully capture the nuances of psychosocial wellbeing as understood by the participants. Many additional factors could be considered in further development of the tool, including the culturally situated concepts of wellbeing and the meaning of the subdomains, the linguistic construction of questions, and how location might affect those differences. For example, the concept of resilience may make sense to academics and those who work for NGOs, but it may be less consequential in contexts where shocks or violence are ongoing and bouncing back is less relevant than persistence. The concept of wellbeing may indeed vary across states, security status, identity groups, and languages, thus causing subtleties that threaten the validity of the measure. Without the engagement of students, caregivers, and teachers to help co-construct these definitions, we clearly will not understand these differences.

However, this tool does appear to capture the general aspects of psychosocial wellbeing that are experienced across cultures and contexts, and it seems suitable for looking at large cohorts. We did find the three domains previously validated in other settings, but there were notable differences particular to the population we studied. The tool also captured one aspect of child wellbeing that local experts identified as specific to South Sudanese children: whether or not children are taking care of their responsibilities. This explains that children who are thriving are able to fulfill their responsibilities. While this is captured in the social domain through the questions about helping others, it also appears to be connected to the self-regulation domain through the behaviors expected of children in their role as students. The fulfillment of these behaviors in school also maps onto some of the skills that are central to SEL.

Triangulating the results of the survey with qualitative data will give us greater confirmation of our findings, as well as deeper insight into what remains central to wellbeing across cultures and what varies. Although we discussed overage students in South Sudan, we invite readers to exercise caution in generalizing the results of our study across a broad age range, due to the fact that less than 3 percent of our sample population was older than age 20. As part of the larger study, we conducted a second phase of qualitative data collection between September and November 2019. We selected five intervention schools, one from each state, for the interviews. We conducted two separate structured focus groups in each school with five randomly selected boys and girls from grade P6. We also conducted interviews with the PSS-trained teachers, members of the school management committee, head teachers, students' parents/guardians, in addition to the county education director, the PSS implementing partner, and the UNICEF or implementing partner's child protection officer in the area. In each of the 60 schools previously covered, we randomly selected 10 students from grade P5 to take a modified numeracy and literacy test. These data provided triangulation by allowing our team to study the relationship between PSS outcomes and students' academic performance.

This instrument will serve as a model for measuring the psychosocial wellbeing of learners under pressure from conflict and recurring crises. We will adapt it for use in ongoing South Sudan USAID-funded projects that aim to build resilience and support the recovery of children and youth. We developed a tool that could be used routinely to assess the effectiveness of mental health and psychosocial support programs in education in emergencies and will share it widely among our donors and other implementing partners working on these programs. We will share the tool in particular with the more than 60 education-focused agencies currently working to provide education in conflict-affected areas of South Sudan and will provide information about its use, intent, and results through workshops we will hold for donors, partners, and MoGEI officials. We introduced the MoGEI to this tool and took part in reviewing questions about its usefulness and providing feedback about its adaptation. We are also working with USAID South Sudan to help educators gain access to conflict-affected communities in order to conduct fieldwork on the tool.

PSS programming is mentioned in the education-sector plan led by the MoGEI relative to activities for the most vulnerable and out-of-school populations. The implementing partner, USAID, was at the forefront of the support given the MoGEI in developing its 2017-2022 plan. USAID also convenes education authorities from around the country to attend an annual meeting where all state-level education ministry representatives report on progress of the plan and discuss the challenges they are facing.

CONCLUSION AND RECOMMENDATIONS

We find evidence for the need to establish the reliability and validity of PSS instruments when deploying them in emergency settings. Notably, the questionnaires we used had previously been used extensively, but in very different sociopolitical and cultural contexts. We find overall that the three core domains we measured (emotional wellbeing, social wellbeing, and resilience) emerged as factors in the South Sudanese context, albeit with important changes. The domain of resilience, in particular, is identified as a significant self-regulation factor in South Sudan.

One important outcome of this work is the modified instrument we present in this study. We recommend that the academic and practitioner communities use it as and when appropriate to assess wellbeing outcomes in South Sudan or similar contexts. While our results are broadly generalizable to South Sudan, we would recommend conducting appropriate reliability and validity confirmations if it is implemented in similar contexts.

Another equally important recommendation is the process through which we adapted, implemented, and reassessed the instruments we used to measure wellbeing outcomes. From the beginning of the process, local researchers prioritized domains of interest and modified the questions as appropriate. We strongly recommend that this level of collaboration and local leadership be a core facet of any work on psychosocial support and, more broadly, on the study of education in conflict settings.

ACKNOWLEDGMENTS

The authors would especially like to acknowledge the contributions of our South Sudanese research associates—Akwero Christine Okumu, Batheng Babouth Rut, Sister Georgina Victor, and Grace Chunyua Marious Said—without whom this research would have been impossible. Also, special thanks to Wonki Lee for her insightful methodological suggestions. Thank you to all the LASER PULSE research team members, far too many to count, who made this project a success. This research is supported in part by the USAID South Sudan Mission and USAID/LAB/CDR through the LASER PULSE Consortium, under cooperative agreement No. 7200AA18CA00009.

REFERENCES

- Action for the Rights of Children. 2009. "Foundation Module 7: Psychosocial Support." Geneva: European Commission Humanitarian Aid Department, Swedish Agency for International Development Cooperation. https://resourcecentre.savethechildren.net/sites/default/files/documents/arc-modf7-studymaterial-2009_0.pdf.
- African Union. 2014. *Final Report of the African Union Commission of Inquiry on South Sudan*. Addis Ababa: African Union Commission of Inquiry on South Sudan. <https://www.peaceau.org/uploads/auciss.final.report.pdf>.
- Ager, Alastair, Bree Akesson, Lindsay Stark, Eirini Flouri, Braxton Okot, Faith McCollister, and Neil Boothby. 2011. "The Impact of the School-Based Psychosocial Structured Activities (PSSA) Program on Conflict-Affected Children in Northern Uganda." *Journal of Child Psychology and Psychiatry* 52 (11): 1124-33. <https://doi.org/10.1111/j.1469-7610.2011.02407.x>.
- Ager, Alastair, Wendy Ager, Vivi Stavrou, and Neil Boothby. 2011. *Inter-agency Guide to the Evaluation of Psychosocial Programming in Emergencies*. New York: UNICEF. https://inee.org/system/files/resources/IASC_Guide_to_the_Evaluation_of_Psychosocial_Programming_in_Humanitarian_Crises.pdf.
- The Alliance for Child Protection in Humanitarian Action. 2020. *Position Paper: Collaboration across Child Protection in Humanitarian Action and Education in Emergencies*. New York: INEE and the Alliance for Child Protection in Humanitarian Action. <https://inee.org/system/files/resources/CPHA%20-%20EIE%20Position%20Paper%20ENG%20v01.pdf>.
- Amnesty International. 2016. "Our Hearts Have Gone Dark": *The Mental Health Impact of South Sudan's Conflict*. London: Amnesty International. https://www.amnestyusa.org/files/southsudan_mental_health_report.pdf.
- The Aspen Institute. 2018. *From a Nation at Risk to a Nation at Hope: Recommendations from the National Commission on Social, Emotional, & Academic Development*. Washington, DC: The Aspen Institute. http://nationathope.org/wp-content/uploads/2018_aspen_final-report_full_webversion.pdf.

- Bandalos, L. Deborah. 2018. *Measurement Theory and Applications for the Social Sciences*. New York: The Guilford Press.
- Blair, Clancy. 2002. "School Readiness: Integrating Cognition and Emotion in a Neurobiological Conceptualization of Children's Functioning at School Entry." *American Psychologist* 57 (2): 111-27. <https://doi.org/10.1037/0003-066X.57.2.111>.
- Bohl, Nicole, Velida Dzino-Siladjic, and Michelle Ryan. 2018. *Guidance on Measuring Children's Psychosocial Well-being in Emergency Response and Recovery Programming*. Baltimore: Catholic Relief Services. https://www.crs.org/sites/default/files/tools-research/meal4kids_well-being_guidance_aug_10_lo_res.pdf.
- Bonis Charancle, Jean Martial, and Elena Lucchi. 2018. "Incorporating the Principle of 'Do No Harm': How to Take Action without Causing Harm—Reflections on a Review of Humanity & Inclusion's Practices." London: Humanity & Inclusion. https://www.alnap.org/system/files/content/resource/files/main/donoharm_pe07_synthesis.pdf.
- Bosqui, Tania Josiane, and Bassam Marshoud. 2018. "Mechanisms of Change for Interventions Aimed at Improving the Wellbeing, Mental Health and Resilience of Children and Adolescents Affected by War and Armed Conflict: A Systematic Review of Reviews." *Conflict and Health* 12 (1): 1-17. <https://doi.org/10.1186/s13031-018-0153-1>.
- Bridgeland, John, John J. Dilulio, Jr., and Karen Burke Morison. 2006. *The Silent Epidemic: Perspectives of High School Dropouts*. Washington, DC: Civic Enterprises. <https://files.eric.ed.gov/fulltext/ED513444.pdf>.
- Bridgeland, John, Mary Bruce, and Arya Hariharan. 2013. *The Missing Piece: A National Teacher Survey on How Social and Emotional Learning Can Empower Children and Transform Schools*. Washington, DC: Civic Enterprises. <https://casel.s3.us-east-2.amazonaws.com/the-missing-piece.pdf>.
- Burde, Dana, Amy Kapit, Rachel L. Wahl, Ozen Guven, and Margot Iglan Skarpeteig. 2017. "Education in Emergencies: A Review of Theory and Research." *Review of Educational Research* 87 (3): 619-58. <https://doi.org/10.3102/0034654316671594>.

- California Department of Education. 2007. *California Healthy Kids Survey*. Berkeley, CA: WestEd. https://data.calschls.org/resources/rydm_surveycontent.pdf.
- Centre for Research on the Epidemiology of Disasters. 2013. *People Affected by Conflict: Humanitarian Needs in Numbers 2013*. Brussels: Centre for Research on the Epidemiology of Disasters. <https://reliefweb.int/sites/reliefweb.int/files/resources/PubID303ConflictReport.pdf>.
- Chauvin, Luc, James Mugaju, and Jondoh Comlavi. 1998. "Evaluation of the Psychosocial Trauma Recovery Programme in Rwanda." *Evaluation and Program Planning* 21 (4): 385-92. [https://doi.org/10.1016/S0149-7189\(98\)00034-2](https://doi.org/10.1016/S0149-7189(98)00034-2).
- Claessens, Lotte F., Donatien C. de Graaff, Mark J. D. Jordans, Frits Boer, and Tom van Yperen. 2012. "Participatory Evaluation of Psychosocial Interventions for Children: A Pilot Study in Northern Uganda." *Journal of Mental Health and Psychosocial Support in Conflict Affected Areas* 10 (1): 43-58.
- Collaborative for Academic, Social, and Emotional Learning. 2003. *Safe and Sound: An Educational Leader's Guide to Evidence-Based Social and Emotional Learning (SEL) Programs*. Chicago: Collaborative for Academic, Social, and Emotional Learning. <https://casel.s3.us-east-2.amazonaws.com/Safe-and-Sound-Guide-to-SEL-Programs.pdf>.
- Convery, Ian, Ruth Balogh, and Bob Carroll. 2010. "Getting the Kids Back to School: Education and the Emotional Geographies of the 2007 Hull Floods." *Journal of Flood Risk Management* 3 (2): 99-111. <https://doi.org/10.1111/j.1753-318X.2010.01060.x>.
- Costello, Anna B., and Jason Osborne. 2005. "Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis." *Practical Assessment, Research, and Evaluation* 10 (7): 1-9. <https://doi.org/10.7275/jyj1-4868>.
- Davis, Kimberley, and Selim Iltus. 2008. *A Practical Guide for Developing Child Friendly Spaces*. New York: UNICEF. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/a_practical_guide_to_developing_child_friendly_spaces_-_unicef.pdf.

- Diamond, Adele, and Kathleen Lee. 2011. "Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old." *Science* 333 (6045): 959-64. <https://doi.org/10.1126/science.1204529>.
- Diazgranados Ferráns, Silvia, and Jeongmin Lee. 2019. *Measure Guidance: Choosing and Contextualizing Assessment Measures in Educational Contexts*. New York: International Rescue Committee. https://inee.org/system/files/resources/Measure_Guidance_final_11.19.2019.pdf.
- Dodge, Rachel, Annette P. Daly, Jan Huyton, and Lalage D. Sanders. 2012. "The Challenge of Defining Wellbeing." *International Journal of Wellbeing* 2 (3): 222-35. <https://doi.org/10.5502/ijw.v2i3.4>.
- Education Cluster. 2018. "Education Cluster Assessment South Sudan." Juba: South Sudan Education Cluster and the Republic of South Sudan Ministry of General Education and Instruction. https://reliefweb.int/sites/reliefweb.int/files/resources/nation_wide_assessment_report-_ssec_-_2018-10.pdf.
- Eiling, Ellen, Marianne Van Diggele-Holtland, Tom Van Yperen, and Frits Boer. 2014. "Psychosocial Support for Children in the Republic of South Sudan: An Evaluation Outcome." *Journal of Mental Health and Psychosocial Support in Conflict Affected Areas* 12 (1): 61-75.
- Elias, Maurice J., M. C. Wang, Roger P. Weissberg, Joseph E. Zins, and Herbert J. Walberg. 2002. "The Other Side of the Report Card: Student Success Depends on More than Test Scores." *American School Boards Journal* 189 (11): 28-31.
- European Commission. 2019. *Education in Emergencies in EU-Funded Humanitarian Aid Operations*. Luxembourg: European Civil Protection and Humanitarian Aid Operations. https://ec.europa.eu/echo/files/news/eie_in_humanitarian_assistance.pdf.
- Global Protection Cluster. 2011. "Guidelines for Child Friendly Spaces in Emergencies." Geneva: Global Education Cluster, Global Protection Cluster, INEE, and IASC. https://inee.org/system/files/resources/GEC_GPC_Child_Friendly_Spaces_Guidelines_for_Field_Testing.pdf.
- Gnambs, Timo, Anna Scharl, and Ulrich Schroeders. 2018. "The Structure of the Rosenberg Self-Esteem Scale: A Cross-Cultural Meta-Analysis." *Zeitschrift Für Psychologie* 226 (1): 14-29. <https://doi.org/10.1027/2151-2604/a000317>.

- Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: A Research Note." *Journal of Child Psychology and Psychiatry* 38 (5): 581-86. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>.
- Hart, Jason, Ananda Galappatti, Jo Boyden, and Miranda Armstrong. 2007. "Participatory Tools for Evaluating Psychosocial Work with Children in Areas of Armed Conflict: A Pilot in Eastern Sri Lanka." *Intervention* 5 (1): 41-60. <https://doi.org/10.1097/WTF.0b013e3280be5a34>.
- Henson, Robin K., and J. Kyle Roberts. 2006. "Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice." *Educational and Psychological Measurement* 66 (3): 393-416. <https://doi.org/10.1177/0013164405282485>.
- Hu, Li-tze, and Peter M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6 (1): 1-55. <https://doi.org/10.1080/10705519909540118>.
- Huebner, E. Scott, James E. Laughlin, Chris Ash, and Rich Gilman. 1998. "Further Validation of the Multidimensional Students' Life Satisfaction Scale." *Journal of Psychoeducational Assessment* 16 (2): 118-34. <https://doi.org/10.1177/073428299801600202>.
- IASC (Inter-agency Standing Committee). 2007. *IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings*. Geneva: American Psychological Association.
- INEE (Inter-agency Network for Education in Emergencies). 2010. *Minimum Standards for Education: Preparedness, Response, Recovery*. New York: INEE. https://inee.org/system/files/resources/INEE_Minimum_Standards_Handbook_2010%28HSP%29_EN.pdf.
- INEE (Inter-agency Network for Education in Emergencies). 2016. *Psychosocial Support and Social and Emotional Learning for Children and Youth in Emergency Settings*. New York: INEE. https://inee.org/system/files/resources/INEE_PSS-SEL_Background_Paper_ENG_v5.3.pdf.

- International Crisis Group. 2014. *South Sudan: Jonglei—“We Have Always Been at War.”* Africa Report No. 221. Brussels: International Crisis Group. <https://www.refworld.org/pdfid/54991c5e4.pdf>.
- Jordans, Mark J. D., Hugo Pigott, and Wietse A. Tol. 2016. “Interventions for Children Affected by Armed Conflict: A Systematic Review of Mental Health and Psychosocial Support in Low- and Middle-Income Countries.” *Current Psychiatry Reports* 18 (9): 1-15. <https://doi.org/10.1007/s11920-015-0648-z>.
- Jordans, Mark J. D., Wietse A. Tol, Aline Ndayisaba, and Ian H. Komproe. 2013. “A Controlled Evaluation of a Brief Parenting Psychoeducation Intervention in Burundi.” *Social Psychiatry and Psychiatric Epidemiology* 48 (11): 1851-59. <https://doi.org/10.1007/s00127-012-0630-6>.
- Kabiru, Caroline W., Chimaraoke O. Izugbara, and Donatien Beguy. 2013. “The Health and Wellbeing of Young People in Sub-Saharan Africa: An Under-Researched Area?” *BMC International Health and Human Rights* 13 (11): 1-7. <https://doi.org/10.1186/1472-698X-13-11>.
- Kamali, Mahdis, Mariella Munyuzangabo, Fahad J. Siddiqui, Michelle F. Gaffey, Sarah Meteke, Daina Als, Reena P. Jain et al. 2020. “Delivering Mental Health and Psychosocial Support Interventions to Women and Children in Conflict Settings: A Systematic Review.” *BMJ Global Health* 5 (3): 1-16. <https://doi.org/10.1136/bmjgh-2019-002014>.
- Kamau, Michael, Derrick Silove, Zachary Steel, Ronald Catanzaro, Catherine Bateman, and Solvig Ekblad. 2004. “Psychiatric Disorders in an African Refugee Camp.” *Intervention* 2 (2): 84-89.
- Kostelny, Kathleen, Ken Ondoro, and Michael Wessells. 2014. *A Grounded Approach to the Definition of Population-Based, Child Protection and Well-Being Outcome Areas in Kenya*. London: Interagency Learning Initiative on Community-Based Child Protection Mechanisms and Child Protection Systems. https://resourcecentre.savethechildren.net/sites/default/files/documents/kenya_outcomes_report_final.pdf.
- Liebenberg, Linda, Michael Ungar, and Fons Van de Vijver. 2012. “Validation of the Child and Youth Resilience Measure-28 (CYRM-28) among Canadian Youth.” *Research on Social Work Practice* 22 (2): 219-26. <https://doi.org/10.1177/1049731511428619>.

- Lu, Francis G., Russell F. Lim, and Juan E. Mezzich. 1995. "Issues in the Assessment and Diagnosis of Culturally Diverse Individuals." *American Psychiatric Press Review of Psychiatry* 14: 477-510.
- Marsh, Herbert W., Louise A. Ellis, Roberto H. Parada, Garry Richards, and Bernd G. Heubeck. 2005. "A Short Version of the Self Description Questionnaire II: Operationalizing Criteria for Short-Form Evaluation with New Applications of Confirmatory Factor Analyses." *Psychological Assessment* 17 (1): 81-102. <https://doi.org/10.1037/1040-3590.17.1.81>.
- McCormick, Meghan P., Elise Cappella, Erin E. O'Connor, and Sandee G. McClowry. 2015. "Social-Emotional Learning and Academic Achievement: Using Causal Methods to Explore Classroom-Level Mechanisms." *AERA Open* 1 (3): 1-26. <https://doi.org/10.1177/2332858415603959>.
- Medvedev, Oleg N., and C. Erik Landhuis. 2018. "Exploring Constructs of Well-Being, Happiness and Quality of Life." *PeerJ* 6: e4903. <https://doi.org/10.7717/peerj.4903>.
- Mguni, Nina, and Nicola Bacon. 2010. *Taking the Temperature of Local Communities: The Wellbeing and Resilience Measure (WARM)*. London: The Young Foundation. <https://youngfoundation.org/wp-content/uploads/2012/10/Taking-the-Temperature-of-Local-Communities.pdf>.
- Michaelson, Juliet, Sorcha Mahony, and Jonathan Schifferes. 2012. *Measuring Well-Being: A Guide for Practitioners*. London: New Economics Foundation. https://b3cdn.net/nefoundation/8d92cf44e70b3d16e6_rgm6bpd3i.pdf.
- MoGEI (Ministry of General Education and Instruction). 2017. *National Education Statistics for the Republic of South Sudan, 2016*. Juba: Ministry of General Education and Instruction. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/emis_2016_-_statistics_booklet_-_2017-02.pdf.
- New Philanthropy Capital. 2011. *Well-Being Measure*. London: New Philanthropy Capital.

- Nicolai, Susan, and Carl Triplehorn. 2003. *The Role of Education in Protecting Children in Conflict*. Humanitarian Practice Network Paper 42. London: Overseas Development Institute. <https://odihpn.org/wp-content/uploads/2003/03/networkpaper042.pdf>.
- O’Sullivan, Clodagh, Tania Bosqui, and Ciaran Shannon. 2016. “Psychological Interventions for Children and Young People Affected by Armed Conflict or Political Violence: A Systematic Literature Review.” *Intervention* 14 (2): 142-64.
- Panter-Brick, Catherine, Rana Dajani, Mark Eggerman, Sabrina Hermosilla, Amelia Sancilio, and Alastair Ager. 2018. “Insecurity, Distress and Mental Health: Experimental and Randomized Controlled Trials of a Psychosocial Intervention for Youth Affected by the Syrian Crisis.” *Journal of Child Psychology and Psychiatry* 59 (5): 523-41. <https://doi.org/10.1111/jcpp.12832>.
- Payton, John, Roger P. Weissberg, Joseph A. Durlak, Allison B. Dymnicki, Rebecca D. Taylor, Kriston B. Schellinger, and Molly Pachan. 2008. *The Positive Impact of Social and Emotional Learning for Kindergarten to Eighth-Grade Students: Findings from Three Scientific Reviews*. Chicago: Collaborative for Academic, Social, and Emotional Learning. <https://files.eric.ed.gov/fulltext/ED505370.pdf>.
- Ravens-Sieberer, Ulrike, and the KIDSCREEN Group Europe, eds. 2016. *The KIDSCREEN Questionnaires: Quality of Life Questionnaires for Children and Adolescents: Handbook*, 3rd ed. Lengerich, Germany: Pabst Science Publishers.
- Ravens-Sieberer, Ulrike, Michael Herdman, Janine Devine, Christiane Otto, Monika Bullinger, Matthias Rose, and Fionna Klasen. 2014. “The European KIDSCREEN Approach to Measure Quality of Life and Well-Being in Children: Development, Current Application, and Future Advances.” *Quality of Life Research* 23 (3): 791-803. <https://doi.org/10.1007/s11136-013-0428-3>.
- Republic of South Sudan Ministry of Education, Science, and Technology. 2015. *Policy for Alternative Education Systems*. Juba: Ministry of Education, Science and Technology. <https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/AES%20policy%20%2019.6.14.pdf>.

- Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.
- Sheikh, Taiwo Lateef, Abdulaziz Mohammed, Samuel Agunbiade, Joseph Ike, William N. Ebiti, and Oluwatosin Adekeye. 2014. "Psycho-Trauma, Psychosocial Adjustment, and Symptomatic Post-Traumatic Stress Disorder among Internally Displaced Persons in Kaduna, Northwestern Nigeria." *Frontiers in Psychiatry* 5 (127): 1-6. <https://doi.org/10.3389/fpsy.2014.00127>.
- Shonkoff, Jack P., W. Thomas Boyce, and Bruce S. McEwen. 2009. "Neuroscience, Molecular Biology, and the Childhood Roots of Health Disparities: Building a New Framework for Health Promotion and Disease Prevention." *The Journal of the American Medical Association* 301 (21): 2252-59. <https://doi.org/10.1001/jama.2009.754>.
- Skårås, Merethe. 2017. "Educational and Social Challenges in the Reintegration Process of Former Child Soldiers." *The Power of Resistance* 12: 243-63. <https://doi.org/10.1108/S1479-358X20140000012012>.
- Snyder, C. Richard, Betsy Hoza, William E. Pelham, Michael Rapoff, Leanne Ware, Michael Danovsky, Lori Highberger, Howard Ribinstein, and Kandy J. Stahl. 1997. "The Development and Validation of the Children's Hope Scale." *Journal of Pediatric Psychology* 22 (3): 399-421. <https://doi.org/10.1093/jpepsy/22.3.399>.
- "South Sudan Clashes: Salva Kiir and Riek Machar Order Ceasefire." 2016. *BBC News*, July 11, 2016. <https://www.bbc.com/news/world-africa-36763076>.
- Stark, Lindsay, Mike Wessells, Dora King, David Lamin, and Sarah Lilley. 2012. *A Grounded Approach to the Definition of Population-Based, Child Protection and Well-Being Outcome Areas*. London: Interagency Learning Initiative on Community-Based Child Protection Mechanisms and Child Protection Systems. <https://resourcecentre.savethechildren.net/document/grounded-approach-definition-population-based-child-protection-and-well-being-outcome-0/>.
- Steuer, Nicola, and Nic Marks. 2008. *Local Wellbeing: Can We Measure It?* London: The Young Foundation. <https://youngfoundation.org/wp-content/uploads/2013/02/Local-Wellbeing-Can-we-Measure-it-September-2008.pdf>.

- Thompson, Sam, and Jody Aked. 2009. *A Guide to Measuring Children's Well-Being. Backing the Future*. London: New Economics Foundation. https://neweconomics.org/uploads/files/094c9bd92c79f7129f_w5m6i2zzh.pdf.
- Tol, Wietse A., Corrado Barbui, Ananda Galappatti, Derrick Silove, Theresa S. Betancourt, Renato Souza, Anne Golaz, and Mark van Ommeren. 2011. "Mental Health and Psychosocial Support in Humanitarian Settings: Linking Practice and Research." *The Lancet* 378 (9802): 1581-91. [https://doi.org/10.1016/S0140-6736\(11\)61094-5](https://doi.org/10.1016/S0140-6736(11)61094-5).
- UNESCO (UN Educational, Scientific and Cultural Organization). 2018. *Global Initiative on Out-of-School Children: South Sudan Country Study*. Paris: UNESCO. <http://uis.unesco.org/sites/default/files/documents/global-initiative-out-of-school-children-south-sudan-country-study.pdf>.
- UNESCO (UN Educational, Scientific and Cultural Organization). 2019. *Education as Healing: Addressing the Trauma of Displacement through Social and Emotional Learning*. Policy Paper ED/GEM/MRT/2019/PP/38. Paris: Global Education Monitoring. <https://en.unesco.org/gem-report/Educationhealing>.
- Ungar, Michael, and Linda Liebenberg. 2011. "Assessing Resilience across Cultures Using Mixed Methods: Construction of the Child and Youth Resilience Measure." *Journal of Mixed Methods Research* 5 (2): 126-49. <https://doi.org/10.1177/1558689811400607>.
- Ungar, Michael, and Linda Liedenberg. 2016. "The Child and Youth Resilience Measure (CYRM) Child Version." Halifax, NS, Canada: Resilience Research Centre.
- UNICEF. 2009. *ARC Resource Pack Foundation Module 7: Psychosocial Support*. Stockholm: Swedish Agency for International Development Cooperation and European Commission Humanitarian Aid Department. https://resourcecentre.savethechildren.net/node/8605/pdf/arc-modf7-studymaterial-2009_0.pdf.
- UNICEF. 2015a. *Growing up in Conflict: The Impact on Children's Mental Health and Psychosocial Well-Being*. Report on the symposium. New York: UNICEF. <https://reliefweb.int/sites/reliefweb.int/files/resources/Growing-up-in-Conflict-final.pdf>.

- UNICEF. 2015b. *Conflict Analysis Summary: South Sudan*. Peacebuilding, Education and Advocacy in Conflict-Affected Contexts Programme. New York: UNICEF. <https://reliefweb.int/sites/reliefweb.int/files/resources/Conflict-Analysis-Summary-South-Sudan-Final1.pdf>.
- Ursache, Alexandra, Clancy Blair, and C. Cybele Raver. 2012. "The Promotion of Self-Regulation as a Means of Enhancing School Readiness and Early Achievement in Children at Risk for School Failure." *Child Development Perspectives* 6 (2): 122-28. <https://doi.org/10.1111/j.1750-8606.2011.00209.x>.
- USAID. 2019. "Social and Emotional Learning and Soft Skills: USAID Policy Brief." Washington, DC: USAID. <https://www.edu-links.org/resources/social-and-emotional-learning-and-soft-skills-usaid-education-policy-brief>.
- Wagnild, Gail M., and Heather M. Young. 1993. "Development and Psychometric Evaluation of the Resilience Scale." *Journal of Nursing Measurement* 1 (2): 165-78.
- Winthrop, Rebecca. 2011. "Understanding the Diverse Forms of Learning Valued by Children in Conflict Contexts." In *Educating Children in Conflict Zones: Research, Policy, and Practice for Systemic Change. A Tribute to Jackie Kirk*, edited by Karen Mundy and Sarah Dryden-Peterson, 123-36. New York: Teachers College Press.
- Zins, Joseph E., Michelle R. Bloodworth, Roger P. Weissberg, and Herbert J. Walberg. 2007. "The Scientific Base Linking Social and Emotional Learning to School Success." *Journal of Educational and Psychological Consultation* 17 (2-3): 191-210. <https://doi.org/10.1080/10474410701413145>.

HOW DO WE KNOW IF TEACHERS ARE WELL? THE WELLBEING HOLISTIC ASSESSMENT FOR TEACHERS TOOL

FERNANDA SOARES, NINA MENEZES CUNHA, AND PAUL FRISOLI

ABSTRACT

This article reports on the development, adaptation, and validation of the Wellbeing Holistic Assessment for Teachers (WHAT) tool with a sample of 1,659 Salvadoran teachers. El Salvador is a conflict-affected country marked by high levels of gang-related violence, which interacts with education and directly affects the wellbeing of teachers. Having a contextually grounded and validated tool is imperative to further our understanding of educator wellbeing in El Salvador and other conflict-affected settings, as it enables us to generate evidence that informs policies and interventions. In this article, we describe how we reviewed and selected the measures that comprise the WHAT tool, followed by an initial conceptualization of teacher wellbeing and a description of the experiences and challenges teachers in El Salvador are facing. We describe our process for translating and adapting the selected measures to the Salvadoran context, which included conducting cognitive interviews. The results from our exploratory factor analysis provide construct validity evidence for the internal structure of the individual measures used. The exploratory factor analysis that included all the items for all the measures confirmed that each scale is indeed measuring a different construct. The results from a confirmatory factor analysis confirmed a good model fit. The process of adapting the tool and the results of our psychometric analysis provide evidence of the tool's validity, based on the content of the items in the tool, the internal structure, and its relationship to other variables.

Received March 27, 2020; revised February 12, 2021; accepted August 2, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

INTRODUCTION

Although teaching can be a rewarding profession, it also has been identified as one of the most stressful occupations (Grenville-Cleave and Boniwell 2012; Maslach, Jackson, and Leiter 1997). Teachers in many different contexts face multiple work-related stressors, which may include the expectation that they will manage students with behavioral difficulties; problems with parent-teacher relationships (Skaalvik and Skaalvik 2007); high job demands; a lack of autonomy and planning time; heavy emphasis on accountability measures; and school systems becoming more bureaucratic (Curry and O'Brien 2012). However, in low-income countries and countries affected by crises and conflict, teachers face a unique set of challenges, both in and out of work (Wolf et al. 2015). In these settings, classrooms may be overcrowded and underresourced, teacher-to-pupils ratios may be high. Moreover, while teachers often are expected to accommodate the mental, social, and emotional needs of their students, they themselves may need support in dealing with their personal psychosocial issues (Wessels and Wood 2019; Kirk and Winthrop 2013; Burns and Lawrie 2015). Heavy and often increasing workloads, limited incentives, and low compensation make the situation for teachers even more challenging (Bennell and Akyeampong 2007).

In crises and conflict settings specifically, students often bring the effects of poverty, trauma, and exposure to conflict into the classroom. Regulating their own negative emotional responses when dealing with students' misbehavior, which often is associated with exposure to trauma and violence, can be a major source of stress and burnout for teachers (Carson, Weiss, and Templin 2010; Montgomery and Rupp 2005; Sutton and Wheatley 2003). In such settings, the teachers themselves are also likely to have experienced conflict-related trauma, and schools may be located in high-conflict areas and be targeted for attack (Wolf et al. 2015). Teachers in conflict-affected and insecure contexts tend to play multiple roles with their students: supporting the children's overall academic and social-emotional growth; being a key caregiver, especially for children who have lost their parents to conflict or displacement; addressing children's overall development; and being their mental health provider, which involves tending to their wellbeing (Sommers 2004; INEE 2010; Frisoli 2013). Teacher wellbeing has implications for the quality of teaching, equitable student access to education, student learning and wellbeing, and the retention and sustainability of the teaching workforce (Fullan 2016; Winthrop and Kirk 2005; Gastaldi et al. 2014).

To inform policy and advance research on teacher wellbeing in low-income and crisis- and conflict-affected contexts, we need measurement tools that are reliable, valid, comparable, and feasible, and also contextually relevant. If education systems, school districts, and school-based leadership are to support teachers adequately, they must know what teachers are experiencing in order to determine how they can make that experience better. Several measurement tools with strong psychometric properties have been developed and validated in the Western context to capture different dimensions of wellbeing, but we do not know if they are adequate for collecting information about teachers' wellbeing in low-income and crisis-affected countries. Researchers and practitioners often use measurement tools that were developed for use in developed countries, with little adaptation. This raises the question of whether the tool can accurately capture the intended construct in a different context. Few studies have focused on adapting and examining the psychometric properties of wellbeing measurement tools with teacher samples in low-income and crisis-affected countries (exceptions include Aboagye et al. 2018).

With this study, we contribute to the literature and to education practice in crisis- and conflict-affected contexts by developing, contextualizing, and validating the Wellbeing Holistic Assessment for Teachers (WHAT) tool, which provides a new multidimensional measure of teacher wellbeing in El Salvador. Given the constraints on resources, logistics, and time that policymakers and practitioners often face when working in conflict-affected settings, our goal was to develop a short, self-administered tool that teachers could complete independently with minimum assessor support. In this article, we first provide a brief overview of the Salvadoran context and how current levels of violence interact with education, followed by a conceptualization of teacher wellbeing and the constructs selected for the WHAT tool. After establishing the contextual background and conceptual underpinnings, we then highlight the purposes and structure of the study.

THE SALVADORAN CONTEXT

The rivalry in El Salvador between two local gangs, Mara Salvatrucha and Barrio 18, which emerged in the aftermath of the 1979-1992 Salvadoran civil war, has been the primary driver of the high level of violence the country has been experiencing for nearly 30 years (ECCN 2016).¹ Gangs in El Salvador consider certain neighborhoods

1 The phenomenon of local gangs, or *pandillas*, was influenced by a US policy launched in 1992 that led to massive deportation of Salvadorian immigrants with criminal records. Youth who had left El Salvador to flee the civil war and had gotten involved in gang violence in the United States (most notably in Los Angeles) brought the US gang culture to their homeland (Borgh and Savenije 2019). "Their deportation back to El Salvador ultimately transformed the pre-existing local *pandilla* culture and organization into the *mara* model" (ECCN 2016, 13).

their territory—mostly poor, marginalized, urban communities but also rural areas (Van Der Borgh and Savenije 2019). They use violence to defend their territory against rival gangs and to control who can cross and live within their domains (ECCN 2016). The gangs oblige local businesses, bus drivers, and other local residents to pay extortion money, which is one of their main sources of income (Van Der Borgh and Savenije 2019). Salvadoran gangs are notorious for their extremely violent behavior, which instills fear and anxiety in the local residents (ECCN 2016). While gangs are dispersed throughout El Salvador, the violence is concentrated in the municipalities; while a substantial portion of the country is homicide free, some municipalities have extremely high homicide rates (Ingram and Curtis 2014).²

The current high level of violence in El Salvador has a direct impact on education and teacher wellbeing. According to the country's ministry of education (MINED 2015), gangs directly threaten the internal security of 24 percent of the nation's schools, and 65 percent of schools are affected by a gang presence in their community. Gangs usually consider the schools in their territory to be their "property"; they often threaten and extort school staff members and prevent students from crossing into their territory to attend school (ECCN 2016). Teachers in particular are directly affected by gang violence: in 2015, 3.55 percent reported receiving threats from gangs and 2.36 percent experienced extortion; in fact, teachers in 7.35 percent of the nation's schools were extorted in or around their school (MINED 2015). The Rapid Education and Risk Analysis (RERA) ECCN conducted in El Salvador in 2016 revealed that working in an environment of intimidation is stressful for teachers and that they express a need for psychological support. The study also revealed that teachers feel overwhelmed and underequipped to handle the emotional needs of their students, who come to school burdened by the effects of violence, threats, and family difficulties. Teachers also reported feeling threatened and being afraid to teach and discipline students who are, or are related to, gang members.

CONCEPTUAL FRAMEWORK

General wellbeing, like occupational wellbeing, is a multidimensional phenomenon comprising affect, cognition, motivations, behaviors, and physical health (Van Horn et al. 2004; Klusmann et al. 2008).³ Affective wellbeing, which is commonly identified as the most central aspect of overall wellbeing, is defined as experiencing

² Forty percent of the homicides occurred in only 5 of the country's 262 municipalities in 2019: San Salvador (12%), San Miguel (10%), Santa Ana (7%), Apopa (7%), and Mejicanos (5%) (Asociación Civil Diálogos 2020).

³ Teacher occupational wellbeing specifically "encompasses teachers' affections, attitudes, and evaluations of their work" (Falk et al. 2019, 2).

low levels of stress and burnout, high job satisfaction, and positive emotions. Affective wellbeing is also usually the focus of interventions that aim to promote teachers' wellbeing. While certain wellbeing dimensions, such as motivation, might be difficult to teach in short-term interventions, there is a growing interest in teaching self-regulation as a protective factor against teacher stress and burnout (Mattern and Bauer 2014). The hypothesis is that teachers with more effective emotion-regulation strategies may be better equipped to deal with the emotional demands of their work that tend to increase stress and may lead to emotional exhaustion.

Specifically, self-awareness and emotion regulation have been linked to having more positive emotions and fewer negative emotions, as well as less stress and burnout (Mattern and Bauer 2014; Roeser et al. 2013; Montgomery and Rupp 2005; Chang 2009; Brackett et al. 2010). Emotion regulation is a key skill that enables teachers to maintain their desirable emotions and reduce or modify unwanted ones (Brackett et al. 2010). As Brackett et al. (2010) explain, "when managing feelings, one must be able to monitor, discriminate among, and label feelings accurately; select and employ strategies that will alter the feelings; and assess the effectiveness of these chosen strategies" (407). In a meta-analysis of 65 studies published between 1998 and 2003 that used quantitative approaches to investigate teacher stress, Montgomery and Rupp (2005) found that emotion regulation is key to preventing stress among teachers. Teachers who have the self-awareness and self-regulatory strategies that are critical to managing stress will experience less stress and burnout and will be more effective teachers.

Several studies also found a strong negative relationship between burnout and teacher self-efficacy (Aloe, Amo, and Shanahan 2014; Skaalvik and Skaalvik 2010; Tsouloupas et al. 2010; Betoret 2009; Bümen 2010; Brouwers and Tomic 2000). Although there is no conclusive empirical evidence on the direction of this relationship, Skaalvik and Skaalvik (2007) hypothesize that the relationship between self-efficacy and teacher burnout may be reciprocal. On the one hand, self-efficacy has been considered a protective factor against burnout. Building on Bandura's (1977) self-efficacy theory, Skaalvik and Skaalvik (2007) argue that people with low self-efficacy may dwell on their coping deficiencies and construe new situations as threatening, and may resort to an escapist mode of coping that can increase anxiety and stress and lead to burnout. However, burnout may also affect self-efficacy. Building again on Bandura's (1977) self-efficacy theory, Brouwers and Tomic (2000) hypothesize that enactive mastery experiences may decrease as a consequence of emotional exhaustion; that is, "the more emotionally exhausted teachers are, the poorer their performances will generally be" (248). Since self-efficacy beliefs are based heavily on experience (Skaalvik and Skaalvik

2010), self-efficacy in classroom management may decline as a consequence of diminished performance (Brouwers and Tomic 2000).

One of the main causes of burnout among teachers is students' misbehavior and the need to discipline them; both have been linked with the emotional exhaustion dimension of teacher burnout (Chan 2006; Evers et al. 2004; Sutton and Wheatley 2003). The classroom-management dimension of teacher self-efficacy is of key importance, as teachers who feel they have less ability to organize and manage students' behavior when dealing with discipline issues could be more susceptible to feeling stressed, which may lead to emotional exhaustion (Aloe et al. 2014). Brouwers and Tomic (2000), for instance, found that perceived self-efficacy in classroom management has an effect on emotional exhaustion, and they conclude that "it is important to take perceived self-efficacy in classroom management into consideration when devising interventions to prevent and to treat teacher burnout" (249).

CHOICE OF CONSTRUCTS

In order to keep the final WHAT instrument short, our goal was to select no more than four constructs for the final tool, each captured by individual scales or subscales. The final constructs we selected were classroom-management self-efficacy (CMSE), stress, emotional exhaustion, and emotion regulation.

The choice of constructs was largely informed by our conceptual framework, and by the experiences of teachers in El Salvador, as per the findings of the ECCN (2016) RERA study, which highlighted gang threats against teachers; extortion; challenges in disciplining students who are, or are related to, gang members; and teaching in an environment of intimidation. These constant stressors, if not resolved or coped with adaptively, can lead to various negative outcomes for teachers, including high levels of stress and burnout. Indeed, the RERA report found that the teachers in their sample often experienced feelings of stress, anxiety, and fear. Given the stressors that educators constantly face in crisis- and conflict-affected settings, and the potential for these stressors to negatively affect wellbeing, we prioritized the measurement of stress and burnout in our tool. With the goal of keeping our tool short and able to generate clear results that can be interpreted easily, we focused on the emotional exhaustion dimension of burnout. It has been argued that emotional exhaustion is the central element of burnout and that the additional weight of depersonalization and reduced personal accomplishment is limited (Betoret 2009; Shirom 1989).

For the purposes of developing this tool, we considered teacher wellbeing to be a broad phenomenon that involves nonaffective aspects. We not only adapted and validated measures of stress and emotional exhaustion, we also included the construct of emotion regulation. This is due to the consensus in the literature that emotion regulation is directly relevant to teachers' stress and burnout (Brackett et al. 2010; Gross 2002), and to the implementation of teacher wellbeing interventions by I/NGOs in crisis- and conflict-affected countries that promote this construct (e.g., through mindfulness programs and socioemotional learning interventions; Falk et al. 2019). Individual differences in emotion regulation may explain differences in teachers' responses to external stressors, which influence feelings of stress and burnout. In other words, teachers will not necessarily feel stressed and burned out when facing the same source of stress—for example, dealing with students' disruptive behavior—but they may feel so depending on how they perceive, appraise, and reinterpret the situation. We also included classroom management self-efficacy as a key construct, as it may prevent teachers' stress and burnout (teachers' low self-efficacy may result in stress and emotional exhaustion) at the same time they are influenced by it (emotional exhaustion may reduce classroom accomplishments and mastery of experiences, which negatively affects self-efficacy). This multidimensional approach enabled us to assess teachers' current level of wellbeing and the protective factors (emotion regulation and self-efficacy) that may influence their responses to different stressors.

The constructs we selected reflect both negative indicators of wellbeing, such as stress and emotional exhaustion, and positive indicators, which focus on protective factors such as self-efficacy and emotion regulation. However, we left important positive aspects of subjective wellbeing, such as positive emotions and cognition, out of our conceptualization and choice of constructs. We recognize that a lack of stress and emotional exhaustion does not equate with a flourishing, successful, and healthily functioning teacher. Nevertheless, given the constant violence-related stressors teachers in El Salvador face, and the importance of protective factors in contributing to our understanding of the nature and causes of teacher wellbeing, we decided to prioritize the negative indicators of wellbeing.

PURPOSES OF THE PRESENT STUDY

The overarching purpose of this study was to develop and establish the psychometric properties of the multidimensional WHAT tool with a sample of Salvadoran teachers. The specific subpurposes of this study were to (1) review and select measures for the WHAT tool (see Appendix A); (2) translate and adapt

these measures to the Salvadoran context through cognitive interviews; (3) collect new data from a sample of Salvadoran primary and secondary education teachers in order to evaluate the psychometric properties of the measures included in the tool; and (4) assess the concurrent relationship of each measure with other variables. Following this process, we selected, translated, and contextualized four measures to El Salvador: the Perceived Stress Scale (PSS); the Emotion Regulation Questionnaire (ERQ); the Emotional Exhaustion subscale of the Maslach Burnout Inventory (MBI-EE); and the self-efficacy for classroom management subscale of the Ohio State Teacher Efficacy (OSTES-CM). Results from our psychometric analysis provided validity evidence for content, internal structure, and concurrent relation to other variables for each of the four translated and contextualized measures comprising the tool. Exploratory factor analysis (EFA), which included all the items of the tool, showed six latent factors, and the confirmatory factor analysis (CFA) confirmed a good model fit. Having a validated tool will help meet the call for research that helps to increase understanding of current levels of educator wellbeing in low-income and crises- and conflict-affected contexts, and for an exploration of the factors and interventions that can influence teacher wellbeing.

METHODS

SELECTING MEASUREMENT TOOLS

To select the measures that are part of the WHAT tool, we first developed an inventory of available measures that assess each of the constructs identified: perceived stress, emotional exhaustion, emotion regulation, and CMSE (see Appendix A). Subsequently, for each construct we selected measures from the inventory based on five criteria adapted from Saloviita (2015). First, we considered the extent to which the scale encompassed themes considered critical for the construct. Second, we assessed the brevity of the scale, as it would be incorporated into a questionnaire with several others. Our goal was to keep the overall questionnaire brief in order to limit responder fatigue and poor response rates. However, we were careful not to achieve brevity at the cost of reduced coverage of the target construct. Third, we looked at internal consistency as a measure of reliability, as scales containing items with high intercorrelation indicate that they measure the same construct. Fourth, we prioritized unidimensional scales to make analysis simpler and more understandable. Fifth, we aimed to include scales with items that are easy to understand.

Based on the five criteria, we selected the following measures.

Emotion Regulation Questionnaire (Gross and John 2003). The ERQ assesses two emotion regulation strategies: cognitive reappraisal (e.g., “When I’m faced with a stressful situation, I make myself *think about it* in a way that helps me stay calm.”) and expressive suppression (e.g., “I keep my emotions to myself.”). Cognitive reappraisal consists of examining and reinterpreting a challenging and difficult situation by reconstruing the situation in nonemotional terms and reducing its emotional impact (Gross 2002). Emotion suppression, on the other hand, involves inhibiting outward signs of inner feelings, which has been shown to increase stress and impair wellbeing (Jennings et al. 2017; Gross 2002). The ERQ consists of ten items measured on a seven-point Likert-type scale (1, “strongly disagree,” to 7, “strongly agree”).

Perceived Stress Scale (Cohen et al. 1983). The PSS assesses the level of perceived stress during the previous month. Its short form consists of ten items (e.g., “How often have you felt that you were unable to control the important things in your life?”) with a five-point response scale (1, “never,” to 5, “very often”). Higher scores correspond to greater levels of perceived stress.

Emotional Exhaustion subscale of the Maslach Burnout Inventory-Educators’ Survey (Maslach et al. 1997). This subscale measures teachers’ feelings of being emotionally overextended and exhausted (e.g., “I feel emotionally drained from my work.”) on a seven-point Likert-type scale (1, “never,” to 7, “every day”). A higher score indicates a higher level of burnout.

Self-efficacy on the classroom-management subscale of the Ohio State Teacher Efficacy Scale (Tschannen-Moran and Hoy, 2001). This eight-item subscale asked how much teachers can do when responding to various classroom-management challenges (e.g., “How much can you do to control disruptive behavior in the classroom?”). Items were rated on a nine-point Likert scale (1, “nothing,” to 9, “a great deal”).

TRANSLATION AND ADAPTATION

One person translated the selected measures from English to Spanish following the ITC Guidelines for Translating and Adapting Tests (Gregoire 2018). Two additional translators verified the translation to ensure that items on the different scales held meaning similar to that in English. The translation was a critical step, as items must be translated well linguistically to maintain an accurate comparison

of responses across cultures (Beaton et al. 2000), and to ensure that they are capturing the intended underlying construct.

In addition to being translated well, items must be adapted culturally in order to maintain evidence of validity based on content across different cultures (Beaton et al. 2000). As part of the adaptation process, the research team conducted cognitive interviews, which have been recognized in the literature as an essential technique for instrument development (Groves et al. 2011; Fowler 1995). Cognitive interviewing makes it possible to verify whether “respondents are able to understand the questions being asked, that questions are understood in the same way by all respondents, and that respondents are willing and able to answer such questions” (Collins 2003, 229-38). Cognitive interviews also provide additional evidence of validity based on content by assessing whether the respondents understand the items in the way intended on the original instrument.

The research team trained four Salvadorans to conduct the cognitive interviews and selected two schools to participate. The interviewers recruited 25 primary and secondary education teachers from the two schools to participate in the interviews, which were conducted at the school sites. In order to identify poorly worded or ambiguous items, the interviewers read each item in the questionnaire to the participating teachers, asking them to verbalize their understanding of the item, comment on the wording, and reveal their response strategy. Following a standardized template, they took detailed notes on each item. The teachers who participated in the interviews were enthusiastic about the tool and expressed appreciation for having their voices heard.

After the interviews were completed, the interviewers transcribed the notes in a central database, categorized by item and type of response. The research team then undertook an item-by-item review of participants’ understanding and wording suggestions. We maintained items that respondents found to be comprehensible and consistently interpreted; we slightly modified the wording of others.⁴ One item from the PSS (“felt stressed and nervous”) was split into two, as some teachers reported feeling stressed but not nervous or the other way around. The translation of two items from the PSS (“unable to control the important things” and “on top of things”) was similar in Spanish (“*en control de las cosas*”), so one was excluded to avoid repetition.

⁴ For instance, the word *acabada/o* was changed to *exhausta/o*. The phrase “*Cuando quiero sentir una emoción menos negativa (tal como tristeza o enojo), cambio en lo que estoy pensando*” was changed to “*Cuando quiero..., trato de cambiar mi pensamiento.*”

Given the resource constraints in crisis- and conflict-affected settings, our goal was to develop a self-administered tool that teachers could complete independently in either a paper or electronic format. With the objective of having a simple and user-friendly questionnaire that respondents could easily understand and answer, we adjusted all measures selected for the tool to fit a five-point Likert scale and standardized the recall periods to two weeks. Specifically, the PSS and MBI-EE items were rated from 1 (“never or almost never”) to 5 (“every day”). ERQ items were rated from 1 (“strongly disagree”) to 5 (“strongly agree”). The OSTES-CM rating scale was adjusted in accordance with the cognitive interviews and items were rated from 1 (“not at all”) to 5 (“completely”). For all the scales in this report, higher scores reflect higher values on the underlying construct. For example, higher scores on the emotion regulation cognitive reappraisal measure indicate greater cognitive reappraisal and higher scores on the perceived stress measure indicate more stress.

PARTICIPANTS AND PROCEDURES

The research team administered a self-reported paper questionnaire to in-service teachers who were participating in the first day of a teacher socioemotional workshop in El Salvador. The workshop was implemented by Family Health International 360 in eight departments—Sonsonate, La Libertad, San Salvador, La Paz, Usulután, San Miguel, La Unión, and Ahuachapán—with funding from the Millennium Challenge Corporation and FOMILENIO II.⁵ It was open to all teachers within the Salvadoran Integrated Systems of Full-Time Inclusive Schools.⁶ The questionnaire was administered to collect baseline data from teachers at the beginning of the workshop series. Before the workshop began, all participants received the questionnaire and were informed that participation in the study was anonymous and voluntary, and that they could refuse to participate or opt out at any time once they began. The paper-based questionnaire was self-reported. The procedure for administering the questionnaire and seeking consent was approved by

5 Seven of these eight departments ranked among the eight with the highest homicide rates per 100,000 people in El Salvador in 2019 (Asociación Civil Diálogos 2020). Nonetheless, it is important to note that department-level analysis of violence obscures municipal-level variations: municipalities with high homicide rates coexist within the same department with municipalities with no homicides (Ingram and Curtis 2014).

6 In 2018, the Millennium Challenge Fund of El Salvador began implementing the SI-EITP intervention in eight departments of the coastal region with funding from the Millennium Challenge Corporation. The SI-EITP model organizes neighboring schools of all grade levels into an integrated system (or cluster of schools) and implements six components (of which the socioemotional workshops are a part) in 45 integrated systems: (1) provide professional development for specialists and teachers to strengthen content and pedagogical knowledge and technological and social and emotional competencies; (2) strengthen and provide timely and effective technical assistance on pedagogy and school management through the hiring and training of 30 technical education assistants; (3) improve the governance of the 45 integrated systems; (4) strengthen English-language teaching for third-cycle and secondary schools; (5) develop reading communities; and (6) construct and rehabilitate one school in each integrated system.

the Family Health International 360 protection of human subjects committee. Out of 2,204 teachers participating in the workshop, 1,659 returned the questionnaire.

The questionnaires were processed through optical reading: 65.3 percent of the respondents were female, 71 percent were age 40 or older, and 57 percent were married. The majority of participating teachers (98%) reported having a teaching degree, a bachelor's degree, or higher; 17 percent reported teaching preschool, 58 percent elementary school, 33 percent middle school, and 14 percent high school. Unfortunately, we do not have data on the wellbeing of teachers who chose not to participate in the workshop.

STATISTICAL ANALYSES

To evaluate the psychometric properties of the constructs, we first conducted an EFA for each construct individually, followed by an internal reliability analysis. We then explored validity evidence based on relations to other variables. Next, we conducted an EFA that included all the constructs in the same model. Finally, we performed a CFA to test whether the data fit the hypothesized measurement model.

While EFA is commonly used to explore the nature of scales and the interrelationships of items and CFA is frequently used to test hypotheses and confirm ideas, the appropriateness of EFA versus CFA remains a central question for researchers when developing instruments; no clear consensus has been reached.⁷ Because this study was the first to implement and test the proposed teacher wellbeing constructs in El Salvador, we first fit an EFA to explore how the items of each proposed scale form a coherent factor, and then fit a CFA model to confirm our hypothesized model. In the next five sections, we describe the steps and specifications we used during the instrument validation process.

EXPLORATORY FACTOR ANALYSIS OF EACH SCALE

We conducted an EFA separately for each of the four scales—emotion regulation (10 items), emotional exhaustion (9 items), perceived stress (10 items), classroom management (8 items)—in order to explore the latent dimensions of each. Factor loadings represent a particular item's relative contribution to an underlying factor.

⁷ The CFA approach to fixing many or all cross-loadings at zero might be problematic and result in a more parsimonious model specification than is suitable for the data (Asparouhov and Muthén 2009). MacCallum, Roznowski, and Necowitz (1992) criticize the tendency to rely on extensive model modification to find well-fitting models using CFA, and Browne (2001) advocates for using EFA rather than CFA for exploratory purposes. Browne (2001) argues that the discovery of mis-specified loadings is more direct through a rotation of the factor matrix than through the examination of model modification indices.

They are similar to correlation coefficients and range between -1.00 and 1.00 (Raykov and Marcoulides 2008). Typically, a factor loading greater than 0.30 or less than -0.30 is used as a guideline for an acceptable factor loading, or an association between the item and the underlying factor (Raykov and Marcoulides 2008). In the case of multiple factors, items may load less than -0.30 or greater than 0.30 on one or more factors, which is called a cross-loading. In such cases, a rotation can facilitate the interpretability of factor loadings.⁸ We conducted an EFA on a polychoric correlation matrix and applied an oblique (promax) rotation when more than one factor was retained.⁹ To determine the number of factors to be retained, we used the Kaiser's criterion of eigenvalues greater than 1, the Cattell scree test, and parallel analysis.^{10, 11}

INTERNAL RELIABILITY: CRONBACH'S ALPHA

To further investigate the psychometric properties of each scale, we calculated means, standard deviations, reliability coefficients, and total item correlation. Internal consistency reliability was assessed using Cronbach's alpha coefficient. Even though there is no consensus on acceptable levels of alpha, many authors agree that reliability above 0.70 is acceptable and 0.8 or greater is preferred (Cortina 1993). Coefficients closer to one indicate that the items on the scale are highly correlated and, therefore, measure the variable of interest more accurately.¹² We also conducted a reliability analysis on each of the items for each measure by assessing the scale alpha coefficient when an item was deleted from it.

⁸ Unrotated results from a factor analysis are not easy to interpret, and rotation was developed to help researchers clarify and simplify the results of a factor analysis. While orthogonal rotation forces the factors to be uncorrelated, oblique rotation allows for correlation between factors if that is optimal for the solution. Given that constructs tend to be at least marginally correlated, we opted for oblique rotation by applying the Promax method by Thompson (2004), considered the more desirable oblique rotation choice (Costello and Osborne 2005).

⁹ Polychoric correlation matrix was used to account for the ordinal structure of the data. Data from rating-scale (Likert) responses are commonly treated as continuous (Norman 2010), even though it is proper to treat such data as ordered categorical (Muthén and Kaplan 1985). EFA is conventionally based on a Pearson correlation matrix, which has been found to underestimate the strength of relationships between ordinal items and generate biased factor loadings (Baglin 2014).

¹⁰ The Kaiser criteria and scree plot have been shown to overestimate the number of dimensions in the data (Baglin 2014). Therefore, we also implement parallel analysis (conducted on a polychoric correlation matrix) using the software FACTOR (Lorenzo-Seva and Ferrando 2006). Except for the EFA that included all scales, parallel analysis yielded the same results as the Kaiser criteria and scree plot for all the constructs analyzed.

¹¹ According to Osborne and Costello (2004), the most common guideline for the ratio of sample size to the number of variables included (participant-to-item ratio) should be at least 10 to 1. The participant-to-item ratio for this analysis was approximately 45 to 1, where sample size was 1,659 and the number of variables was 37. This indicates that our sample size was sufficient to produce reliable results.

¹² However, the coefficient is sensitive to the number of items in the scale and a larger number of items can result in a larger coefficient.

VALIDITY EVIDENCE BASED ON RELATIONS TO OTHER VARIABLES

We assessed validity evidence based on relations to other variables by analyzing how each of the four selected measures correlates with other measures that have shown previous significant association in the literature, which also is referred to in the literature as concurrent validity. To conduct this analysis, we collected data using the following scales.

Patient Health Questionnaire Depression Scale (PHQ-8; Kroenke et al. 2010). This scale, which measures depressive symptoms (e.g., “feeling down, depressed, or hopeless”), consists of eight items rated on a four-point Likert-type scale (1, “not at all,” to 4, “nearly every day”). The alpha coefficient for the PHQ-8 in the current sample was 0.82.

Generalized Anxiety Disorder Scale (GAD-7; Spitzer et al. 2006). This scale measures generalized anxiety symptoms (e.g., “feeling nervous, anxious, or on edge”) on a four-point Likert-type scale (1, “not at all,” to 4, “nearly every day”). The alpha coefficient for the GAD-7 in our sample was 0.88.

Positive and Negative Affect Rating Short Form (PANAS; Thompson 2007). This brief ten-item measure asks participants to rate how they “felt during the past few weeks” on ten emotions using a five-point Likert-type scale (1, “very little or not at all,” to 5, “extremely”). The alpha coefficients for positive and negative affect subscales were 0.80 and 0.88, respectively.

We also assessed the intercorrelation among the measures in the tool, and correlations between the four measures and key sociodemographic variables: gender, age, marital status, socioeconomic status (SES), and educational level taught.¹³ Assessing these correlations helped us understand whether the scales are indeed measuring what they are supposed to be measuring, while also expanding our knowledge on how background demographics might be associated with teacher wellbeing measures.¹⁴

¹³ Pearson’s correlation was implemented using `pwcorr` in Stata. `Pwcorr` has the option of showing statistical significance of the correlation and handles missing values by pairwise deletion (all available observations are used to calculate each pairwise correlation).

¹⁴ In the case of missing responses, the following rule was used: If surveys were missing less than 50 percent of responses, the average score for the item was used. This rule was applied to the main constructs and to the additional scales.

EXPLORATORY FACTOR ANALYSIS ON ALL SCALES

To further investigate whether each scale is indeed measuring different constructs, we conducted an EFA that pooled all scales of the tool. We included all 37 items of the tool using a polychoric correlation matrix and oblique (promax) rotation. We used Kaiser's criterion of eigenvalues, a Cattell scree test, and parallel analysis to determine the number of factors to be retained. We discuss our decision to maintain or exclude items that demonstrated poor performance throughout the validation process.

CONFIRMATORY FACTOR ANALYSIS

Finally, we conducted a CFA to test whether the data fit the hypothesized measurement model. Different estimators can be used for model fitting with categorical data, but here we used the diagonally weighted least squares estimator (WLSMV in Mplus), which was appropriate in this case (Muthén, Muthén, and Asparouhov 2015).

Because the Chi-square statistic is sensitive to sample size and may reject well-fitting models, our model fit assessment put more emphasis on the other statistics. We reported the root mean square error of approximation (RMSEA), the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the standardized root mean square residual (SRMR).¹⁵ We assessed these fit statistics to determine whether the models are providing a good fit to the data. Following recommendations from Hu and Bentler (1999), we adopted the following cutoff values as a guide for establishing whether the models fit the data well, as follows: SRMR ideally below 0.08 and at most 0.10; RMSEA ideally below 0.06 and at most 0.10; and CFI and TLI ideally above 0.95, with a minimum of 0.90.

RESULTS

DESCRIPTIVE STATISTICS, EFA, INTERNAL CONSISTENCY, AND VALIDITY FOR EACH SCALE

Table 1 and Table 2 show (1) descriptive statistics for each item and scale (on a five-point Likert scale); (2) Cronbach's alpha coefficient for each scale; (3) the eigenvalue

¹⁵ The SRMR and RMSEA are measures of absolute fit, with values closer to zero indicating a better fit. The CFI and TLI are, in turn, a measure of relative fit, with values closer to one indicating a better fit.

and percentage of variance explained by the factor retained in the EFA;¹⁶ (4) total item correlation for each item (r);¹⁷ (5) EFA loading for each item; and (6) the Cronbach's alpha coefficient the scale possessed when each item was deleted from it.¹⁸ Table 3 provides Pearson correlations and significant levels.¹⁹ Taken together, these statistics give us a measure of how strong and consistent these constructs are, where higher values for alphas (>0.70), total item correlation (>0.20), and factor loadings magnitude (less than -0.30 or greater than 0.30) indicate that the constructs are solid and working well. Below we describe the results for each measurement in more detail.

Table 1: Descriptive Statistics, Total Item Correlation, Internal Reliability, and EFA Loadings²⁰

Item	Obs.	Mean	Std.	r	Alpha	EFA
Emotion regulation reappraisal	1624	3.99	0.62			
When I want to feel more positive emotion (such as joy or amusement), I change what I'm thinking about.	1589	3.89	1.01	0.36	0.71	0.60
When I want to feel less negative emotion (such as sadness or anger), I change what I'm thinking about.	1548	3.95	1.03	0.42	0.69	0.69
When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm.	1568	4.14	0.93	0.43	0.69	0.68
When I want to feel more positive emotion, I change the way I'm thinking about the situation.	1587	4.03	0.94	0.53	0.66	0.77
I control my emotions by changing the way I think about the situation I'm in.	1541	3.96	0.89	0.49	0.67	0.76

16 The eigenvalue represents the total amount of variance explained by the factor and is calculated as the sum of squared factor loadings across all items for each factor. The eigenvalues are higher when there are at least some variables with high factor loadings, and lower when there are mostly low loadings. In practice, only factors with eigenvalues of 1 or higher are analyzed, although other approaches are also used to select the ideal number of factors.

17 Total item correlation, or item rest correlation, is the correlation between an item and the scale that is formed by all other items.

18 Before conducting an EFA, we assessed its suitability. An inspection of the polychoric correlation between the items showed that correlations were high and did not go below 0.28 in any case (Table A1). The Kaiser-Meyer-Olkin measure was greater than 0.73 for each measurement, which according to Kaiser are acceptable classifications (Kaiser and Rice 1974). Bartlett's test of sphericity was statistically significant ($p < .000$) for each scale, which indicates that the data were suitable for EFA.

19 Pearson correlation was implemented using `pwcorr` in Stata. `Pwcorr` has the option of showing statistical significance of the correlation and handles missing values by pairwise deletion (all available observations are used to calculate each pairwise correlation).

20 Items on the emotional exhaustion scale are from the Maslach Burnout Inventory, copyright (c) 1996 Wilmar B. Schaufeli, Michael P. Leiter, Christina Maslach, and Susan E. Jackson. All rights reserved. The inventory may not be used without permission of the publisher, Mind Garden, Inc. www.mindgarden.com.

Item	Obs.	Mean	Std.	r	Alpha	EFA
When I want to feel less negative emotion, I change the way I'm thinking about the situation.	1592	4.01	0.88	0.48	0.67	0.75
Alpha					0.72	
Eigenvalue (Factor 1)						3.44
% Variance (Factor 1)						0.34
Emotion regulation suppression	1508	3.04	0.94			
I keep my emotions to myself.	1390	3.19	1.24	0.54	0.68	0.79
When I am feeling positive emotions, I am careful not to express them.	1412	2.72	1.24	0.56	0.66	0.82
I control my emotions by not expressing them.	1464	2.84	1.24	0.60	0.64	0.84
When I am feeling negative emotions, I make sure not to express them.	1587	3.38	1.22	0.41	0.75	0.65
Alpha					0.74	
Eigenvalue (Factor 2)						2.09
% Variance (Factor 2)						0.21
Cumulative Variance (Factor 1 + Factor 2)						0.55
Emotional exhaustion	1630	1.55	0.5			
Drained	1618	1.86	0.77	0.63	0.84	0.79
Used up	1606	2.14	0.94	0.63	0.84	0.77
Fatigued	1598	1.51	0.74	0.65	0.84	0.80
End of my rope	1563	1.64	0.81	0.64	0.84	0.80
Burned out	1614	1.35	0.68	0.72	0.83	0.89
Frustrated	1590	1.18	0.49	0.53	0.85	0.74
Working too hard	1581	1.63	0.81	0.47	0.86	0.68
Stress	1598	1.42	0.66	0.53	0.85	0.73
Strain	1594	1.23	0.53	0.58	0.85	0.78
Alpha					0.86	
Eigenvalue						5.44
% Variance						0.60

Note: Negative items were reverse scored; “r” shows total-item correlation and “alpha” shows internal reliability for the scale if the item is excluded.

Table 2: Descriptive Statistics, Total Item Correlation, Internal Reliability, and EFA Loadings (cont.)

Item	Obs.	Mean	Std.	r	Alpha	EFA
Perceived stress I (In the past two weeks, how often have you...)	1628	1.69	0.51			
been upset because of something that happened unexpectedly?	1611	1.67	0.65	0.44	0.71	0.69
felt "stressed"?	1597	1.93	0.82	0.56	0.67	0.80
felt nervous?	1589	1.57	0.74	0.51	0.69	0.74
found that you could not cope with all the things that you had to do?	1588	1.81	0.95	0.34	0.75	0.64
been angered because of things that were outside of your control?	1613	1.62	0.74	0.47	0.70	0.75
felt difficulties were piling up so high that you could not overcome them?	1612	1.52	0.72	0.54	0.68	0.77
Alpha					0.74	
Eigenvalue (Factor 1)						3.75
% Variance (Factor 1)						0.37
Perceived stress II (In the past two weeks, how often have you...)	1625	2.76	1.09			
felt confident about your ability to handle your personal problems?	1576	2.62	1.44	0.63	0.79	0.83
felt that things were going your way?	1598	3.12	1.25	0.56	0.82	0.80
been able to control irritations in your life?	1591	2.75	1.38	0.69	0.76	0.86
felt that you were on top of things?	1578	2.55	1.31	0.71	0.75	0.85
Alpha					0.82	
Eigenvalue (Factor 2)						2.38
% Variance (Factor 2)						0.24
Cumulative Variance (Factor 1 + Factor 2)						0.61
Classroom management	1608	4	0.59			
How much can you do to control disruptive behavior in the classroom?	1603	4.05	0.79	0.60	0.87	0.76
How much can you do to get children to follow classroom rules?	1598	4.05	0.71	0.67	0.86	0.82
How much can you do to calm a student who is disruptive or noisy?	1589	4.02	0.78	0.69	0.86	0.84

Item	Obs.	Mean	Std.	r	Alpha	EFA
To what extent can you establish a classroom-management system with each group of students?	1581	3.95	0.80	0.65	0.86	0.80
To what extent can you keep a few problem students from ruining an entire lesson?	1590	4.02	0.84	0.69	0.86	0.82
To what extent can you respond to defiant students?	1587	3.83	0.87	0.60	0.87	0.75
To what extent can you make your expectations clear about student behavior?	1585	4.09	0.77	0.66	0.86	0.80
To what extent can you establish routines to keep activities running smoothly?	1597	4.02	0.79	0.61	0.87	0.76
Alpha					0.88	
Eigenvalue						5.04
% Variance						0.63

Note: (i) Negative items were reversed scored; “r” shows total-item correlation and “alpha” shows internal reliability for the scale if the item is excluded. (ii) All scales are measures on a five-point Likert scale and recall periods are standardized to two weeks. PSS and MBI-EE items were rated from 1 (“never or almost never”) to 5 (“every day”). ERQ items were rated from 1 (“strongly disagree”) to 5 (“strongly agree”). OSTES-CM rating scale was adjusted based on the cognitive interviews and items rated from 1 (“not at all”) to 5 (“completely”).

Table 3: Pearson Correlations between Tool Measures, External Measures, and Demographics²¹

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) Emo. Reg. Reap.	1					
(2) Emo. Reg. Sup.	0.26*	1				
(3) Emo. Exhaustion	-0.07*	0.06*	1			
(4) Stress I	-0.07*	0.11*	0.59*	1		
(5) Stress II	-0.12*	0.18*	0.13*	0.17*	1	
(6) Class Management	0.18*	0	-0.16*	-0.17*	-0.26*	1
(7) Positive Affect	0.21*	-0.07*	-0.16*	-0.15*	-0.32*	0.27*
(8) Negative Affect	-0.07*	0.07*	0.39*	0.52*	0.17*	-0.14*
(9) Depression	-0.09*	0.15*	0.55*	0.62*	0.25*	-0.17*
(10) Anxiety	-0.06*	0.14*	0.55*	0.71*	0.21*	-0.15*
(11) Female	0.02	0.03	0	0.11*	0.13*	-0.09*
(12) Age	0.03	0.05	0.01	0.04	0.04	-0.01
(13) Married	-0.03	-0.03	0.02	-0.01	0.02	-0.02
(14) SES	0.02	-0.02	0.02	0.03	-0.02	0.09*
(15) Preschool	0.02	0.05	0.02	0.04	0.08*	-0.07*
(16) Elem. School	0	0.08*	0	0	0.03	-0.01
(17) Middle School	-0.03	-0.04	0.05	0.03	-0.09*	0.06*
(18) High School	-0.02	-0.07*	-0.01	-0.03	-0.07*	0.02

Note: * shows significance at the .05 level

EMOTION REGULATION QUESTIONNAIRE

Descriptive statistics for the items and the scales show that teachers scored relatively high on the emotion regulation reappraisal subscale: the mean for each item ranged from 3.89 to 4.14, and the scale mean was 3.99 out of 5.²² The average was lower for the emotion regulation suppression: item means ranged from 2.72 to 3.38, and the scale mean was 3.04.²³ This is consistent with other

²¹ Each cell shows Pearson correlation between the two variables in the row and in the column. Pearson correlation was implemented using `pwcorr` in Stata. `Pwcorr` has the option of showing statistical significance of the correlation and handles missing values by pairwise deletion (all available observations are used to calculate each pairwise correlation).

²² Minimum value is 1 and maximum value is 5.

²³ Negative items of all constructs were reverse coded.

studies that found teachers more prone to using cognitive reappraisal than expressive suppression strategies (Tsouloupas et al. 2010), meaning that they tend to reinterpret challenging situations in nonemotional terms rather than inhibit signs of their inner feelings when regulating their emotions.

EFA yielded a two-factor solution, which accounted for 55 percent of the variance that matched the exact original structure of the ERQ proposed by Gross and John (2003) and was subsequently replicated by EFA and CFA in the student samples (Chen 2010; D'Argembeau and Van der Linden 2006; Balzarotti, John, and Gross 2010). EFA yielded one factor for the emotion regulation reappraisal subscale, with eigenvalue of 3.02 accounting for 34 percent of variance; the item loadings ranged from 0.69 to 0.71. The second factor that emerged from the exploratory factor analysis for the suppression subscale showed eigenvalue of 2.42, which accounts for 21 percent of variance. The item loadings ranged from 0.64 to 0.75.

The Cronbach's alpha for the emotion regulation reappraisal subscale was 0.72, and there was no item that, if excluded, would increase this reliability. Item total correlation was high, ranging from 0.36 to 0.53. The emotion regulation suppression subscale showed a total reliability of 0.74, which increased by 0.01 if the fourth item of the scale was excluded. Item total correlation ranged from 0.41 to 0.60.

Table 3 presents clear evidence that emotional suppression is significantly associated with increased depression, anxiety, perceived stress, and negative affect. It also supports an association between reappraisal and more positive emotion. Overall, these findings are consistent with previous studies (Spaapen et al. 2014; Wiltink et al. 2011; Gross and John 2003) that analyzed correlations between the ERQ and other scales. No significant age correlations were observed for suppression or reappraisal, which is in keeping with findings from Spaapen et al. (2014) and Wiltink et al. (2011) but contrary to previous studies showing a positive association between older people and cognitive reappraisal (Charles and Carstensen 2007). Contrary to expectations, we also did not find any association between suppression and gender. While no gender differences have been observed in the use of reappraisal, several studies have shown that males tend to suppress emotions more than females (Spaapen et al. 2014; Wiltink et al. 2011; Balzarotti et al. 2010; Gross and John 2003). Suppression is also positively associated with teaching elementary school and negatively associated with teaching high school, meaning that teachers at the elementary level tend to suppress their emotions more. Finally, we did not find evidence for associations between reappraisal or suppression and marital or SES status, which have not been previously explored in the literature.

Emotional Exhaustion (EE) Subscale. Descriptive statistics show that teachers' scores on each item ranged from 1.18 to 2.14, and the overall scale mean was 1.55 out of 5. The results of EFA indicated a dominant one-factor solution with eigenvalue of 5.44, accounting for 60 percent of variance, and the item loadings ranged from 0.68 to 0.89. While the reliability and validity of the MBI educators survey have been established in previous articles (Aluja, Blanch, and García 2005; Boles et al. 2000), we are not aware of any studies that separately validate the emotional exhaustion subscale for educators. Given the original three-factor structure of the MBI, we would expect all nine items of the MBI-EE, which form one factor in the original MBI, to load well into only one factor.

The MBI-EE subscale showed an overall Cronbach's alpha of 0.86, and this reliability would not increase by excluding any of the items. Item total correlation ranged from 0.49 to 0.94.

Results shown in Table 3 indicate that emotional exhaustion is associated with increased stress, negative affect, depression, and anxiety, and with decreased CMSE. This is consistent with the previous literature that found significant correlations between emotional exhaustion and depression and anxiety (Calvete and Villa Sánchez 1999; Schonfeld and Bianchi 2016), and emotional exhaustion and CMSE (Dicke et al. 2014; Aloe et al. 2014) among teacher samples. In contrast to previous literature that reported higher levels of emotional exhaustion in females than in males (Grayson and Alvarez 2008; Lau, Yuen, and Chan 2005), we did not find any association between gender and emotional exhaustion in the Salvadoran teachers sample. We found no associations between marital status, SES, age, or education level taught. The lack of association with education level taught is somewhat unexpected, given the degree of control and influence gangs have in the functioning of secondary schools in El Salvador (ECCN 2016).

Perceived Stress Scale. Teachers scored higher on the positive scale than the negative scale, which is similar to the pattern we found for emotion regulation: item averages for the positive scale ranged from 2.55 to 3.12, and the scale mean was 2.75, whereas item averages for the negative scale ranged from 1.52 to 1.93, and the scale mean was 1.69 out of 5.

EFA yielded two factors with eigenvalues of 3.75 and 2.38, accounting for 61 percent of variance. The percentage of variance explained by the two factors was higher than what was found by previous studies, most of which have shown that the two-factor structure accounts for less than 50 percent of the total variance (Lee 2012). Factor 1, which accounted for 37 percent of variance, consisted of six

items representing “negative feelings”; factor 2, which accounted for 24 percent of variance, consisted of four items representing “positive feelings.” This factorial structure is in line with findings from Lee (2012) that show that a two-factor structure of the PSS-10 predominate in different validation studies of its English version. This is also consistent with validations of the PSS-14 in Spanish-speaking samples (González 2006; Ramírez and Hernández 2007), which confirmed the two-factor structure. The item loadings ranged from 0.64 to 0.86.

Cronbach’s alpha reliability coefficients were 0.74 for the “negative feelings” factor and 0.82 for the “positive feelings” factor. Reliability of the “negative feelings” scale would increase by 0.01 if item four was excluded. Item total correlation ranged from 0.34 to 0.71.

Table 3 indicates that both perceived stress I (“negative feelings”) and II (“positive feelings”) are strongly correlated with increased depression, anxiety, emotional exhaustion, and negative affect. However, the magnitude of these correlations is stronger for stress I. These results are consistent with an extensive review of articles conducted by Lee (2012) that related to the psychometric properties of the PSS. Lee (2012) found that the PSS was either moderately or strongly correlated with the hypothesized emotional variables, such as depression or anxiety and emotional exhaustion. This is also in line with Cohen et al.’s (1983) expectation that “there is some overlap between what is measured by depressive symptomatology scales and measured by the PSS, since the perception of stress may be a symptom of depression” (391). In our sample, females were associated with increased perceived stress. This may be related to cultural expectations outside the school environment that deem it appropriate for childcare and household tasks to be performed by females. Perceived stress II (“positive feelings”) showed a positive association with teaching preschool and negative associations with teaching middle and high school. This was unexpected in the Salvadoran context; given the heavier presence and influence of gangs in the secondary schools, we would expect middle and high school teachers to show higher levels of stress. Surprisingly, we found no associations with other demographic variables, such as marital status, SES, and age. This is inconsistent with previous studies, which have found that PSS scores were lower for young, married, and higher-paid respondents (Lee 2012).

Classroom Management Subscale. Descriptive statistics shows that teachers scored high on classroom management: the scale mean was 4 out of 5, and the item mean ranged from 3.83 to 4.09. EFA yielded one factor solution with eigenvalue of 5.04, which accounted for 63 percent of variance, and the item loadings ranged from 0.75 to 0.84. The reliability of the classroom management subscale was 0.88; excluding items would not increase the overall reliability. Item total correlation ranged from 0.60 to 0.69.

We found that CMSE is significantly associated with decreased anxiety, depression, negative affect, perceived stress, and emotional exhaustion. This is consistent with the previous literature that demonstrates that positive teaching self-efficacy may result in improved psychological wellbeing and lower levels of stress and burnout (Zee and Koomen 2016; Aloe et al. 2014; Tsouloupas et al. 2010). Classroom-management self-efficacy is negatively associated with teaching preschool and positively associated with teaching middle school. This is not surprising, as previous studies have shown lower CMSE among preschool teachers, which is possibly related to disciplinary difficulties and to teaching appropriate classroom behavior to pupils who are attending school for the first time (Cocca et al. 2018). Table 3 shows that classroom management is negatively associated with being female, but no association was found with the demographic variables of marital status, SES, and age.

EFA-All Scales. Table 4 displays the loadings for the seven factors retained by the EFA, including all 37 items of the proposed WHAT tool. Table B2 shows eigenvalues and statistics for the EFA, Figure B1 shows the scree plot, and Table B3 shows parallel analysis (see Appendix B). While the Kaiser's criterion of eigenvalues suggests the retention of seven factors, the scree plot suggests that six or seven factors could be retained, and the parallel analysis suggests that five could be retained.²⁴ Table 4 shows that only one item (fourth item of the perceived stress I scale) had a high loading on the seventh factor, and the item also shows a high loading for factor 3.²⁵ Given the separation in a seventh factor and the cross-loading, we decided to exclude this item from the tool. Excluding this item increases the perceived stress I scale reliability from 0.74 to 0.75.

²⁴ There are many guidelines for how to decide the number of factors to extract from the analysis. The Kaiser's criteria, the scree plot, and the parallel analysis are three of the most common methods used in this decision of factor extraction. In practice, results for factor extraction using these different methods identify the optimal number of factors to be extracted from the data. In other words, it identifies how many different constructs are being measured by the data.

²⁵ The item with a high loading on factor 7 corresponds to item 4 of the perceived stress I scale. As we saw in the previous section, this item, if excluded, would increase the reliability of the perceived stress I scale.

Table 4: EFA All Scales–Loadings

	F1	F2	F3	F4	F5	F6	F7	Uniqueness
Emotion regulation reappraisal								
(1) When I want to feel more positive emotion (such as joy or amusement)...	-0.03	-0.02	0.12	-0.06	0.44	0.24	-0.32	0.52
(2) When I want to feel less negative emotion (such as sadness or anger)...	0.12	0.02	0.02	0.00	0.65	0.00	-0.20	0.50
(3) When I'm faced with a stressful situation, I make myself think about it...	-0.01	0.03	-0.08	-0.07	0.67	-0.03	0.00	0.50
(4) When I want to feel more positive emotion, I change the way...	0.01	-0.01	-0.04	0.05	0.81	-0.03	0.03	0.39
(5) I control my emotions by changing the way I think about the situation I'm in.	0.01	0.04	-0.04	0.07	0.82	0.00	0.17	0.37
(6) When I want to feel less negative emotion, I change the way I'm thinking...	-0.02	-0.05	0.01	0.00	0.80	-0.09	0.13	0.42
Emotion regulation suppression								
(1) I keep my emotions to myself.	0.01	0.04	0.13	-0.02	-0.02	0.77	-0.21	0.36
(2) When I am feeling positive emotions, I am careful not to express them.	0.04	-0.01	-0.09	0.06	-0.12	0.84	-0.07	0.31
(3) I control my emotions by not expressing them.	-0.02	0.00	-0.02	-0.04	-0.01	0.83	0.11	0.30
(4) When I am feeling negative emotions, I make sure not to express them.	-0.10	-0.04	-0.01	-0.02	0.21	0.59	0.07	0.56

DEVELOPING THE WELLBEING HOLISTIC ASSESSMENT FOR TEACHERS

	F1	F2	F3	F4	F5	F6	F7	Uniqueness
Emotional exhaustion								
(1) Drained	0.75	0.01	0.15	0.01	0.03	-0.08	-0.17	0.33
(2) Used up	0.80	0.05	0.12	-0.08	-0.04	-0.08	-0.25	0.28
(3) Fatigued	0.76	-0.02	0.10	0.06	-0.02	-0.02	-0.16	0.33
(4) End of my rope	0.88	-0.06	-0.14	0.02	0.06	-0.02	-0.04	0.33
(5) Burned out	0.85	-0.05	0.00	0.02	0.08	0.00	0.15	0.20
(6) Frustrated	0.57	-0.07	0.13	0.12	-0.03	0.11	0.09	0.43
(7) Working too hard	0.72	0.05	-0.05	-0.15	0.00	0.03	0.16	0.48
(8) Stress	0.73	0.02	-0.03	-0.03	-0.02	0.04	0.13	0.43
(9) Strain	0.69	0.05	0.01	0.04	-0.04	0.12	0.35	0.26
Perceived stress I (In the past two weeks, how often have you...)								
(1) been upset because of something that happened unexpectedly?	0.04	-0.08	0.71	0.00	-0.01	-0.01	0.17	0.42
(2) felt "stressed"?	0.44	0.01	0.53	-0.02	0.03	-0.03	0.00	0.32
(3) felt nervous?	0.13	-0.01	0.64	0.09	-0.08	0.04	0.10	0.43
(4) found that you could not cope with all the things that you had to do?	0.01	0.03	0.47	-0.09	0.14	-0.06	0.78	0.30
(5) been angered because of things that were outside of your control?	0.07	-0.03	0.73	-0.04	-0.03	-0.05	0.17	0.39
(6) felt difficulties were piling up so high that you could not overcome them?	0.16	0.08	0.64	0.09	-0.03	0.06	0.34	0.34

	F1	F2	F3	F4	F5	F6	F7	Uniqueness
Perceived stress II								
(In the past two weeks, how often have you...)								
(1) felt confident about your ability to handle your personal problems?	0.05	0.03	-0.07	0.84	0.02	0.03	0.01	0.30
(2) felt that things were going your way?	0.00	0.07	-0.12	0.85	0.06	-0.05	-0.08	0.36
(3) been able to control irritations in your life?	0.01	-0.03	0.08	0.83	0.01	0.03	-0.05	0.27
(4) felt that you were on top of things?	-0.06	-0.06	0.19	0.82	-0.02	-0.03	-0.02	0.23
Classroom management								
(1) How much can you do to control disruptive behavior in the classroom?	-0.09	0.74	0.10	-0.02	0.02	-0.06	-0.05	0.39
(2) How much can you do to get children to follow classroom rules?	0.00	0.83	-0.05	0.03	-0.01	0.01	0.01	0.32
(3) How much can you do to calm a student who is disruptive or noisy?	0.09	0.87	-0.11	0.02	-0.03	0.00	0.05	0.27
(4) To what extent can you establish a classroom-management system?	-0.07	0.81	-0.03	0.05	0.02	-0.01	0.09	0.33
(5) To what extent can you keep a few problem students from ruining an entire lesson?	0.06	0.86	-0.02	0.01	-0.02	0.03	0.06	0.29
(6) To what extent can you respond to defiant students?	0.03	0.74	0.01	-0.05	-0.02	0.00	-0.08	0.43
(7) To what extent can you make your expectations clear about student behavior?	-0.03	0.74	0.08	-0.05	0.05	-0.02	-0.11	0.37
(8) To what extent can you establish routines to keep activities running smoothly?	-0.07	0.73	0.02	-0.02	0.00	0.06	0.04	0.44

Note: All scales are measures on a five-point Likert scale and recall periods are standardized to two weeks. PSS and MBI-EE items were rated from 1 (“never or almost never”) to 5 (“every day”). ERQ items were rated from 1 (“strongly disagree”) to 5 (“strongly agree”). OSTES-CM rating scale was adjusted based on the cognitive interviews and items rated from 1 (“not at all”) to 5 (“completely”).

Excluding the seventh factor, the other six factors exactly match the structure of the tool we described in the previous section: factor 1 aggregated items of the emotional exhaustion scale, factor 2 aggregated items from the classroom-management scale, factors 3 and 4 aggregated items from the perceived stress I and II scales, respectively, and factors 5 and 6 aggregated items of the emotion regulation reappraisal and suppression scales, respectively. The parallel analysis suggests that five factors could be retained, where items of the emotional exhaustion scale and stress I (negative) scale would be part of the same factor, while the other factors remain the same. This is an interesting finding, given that these two scales showed the highest correlation (0.59) in Table 3. Following our hypothesized model, we decided to retain six factors, maintaining exhaustion and stress I as separate scales. We fit a CFA to test whether the data fit the hypothesized measurement model, which we describe in the next section.

CONFIRMATORY FACTOR ANALYSIS

Table 5 shows CFA fit statistics for the five CFA models we estimated: (1) emotion regulation, containing two factors (reappraisal and suppression); (2) emotional exhaustion; (3) perceived stress, containing two factors (positive and negative); (4) classroom management; and (5) a model containing all the constructs. Except for the RMSE of 0.115 in the classroom-management model, all the fit statistics were below the acceptable thresholds, which indicates a good model fit for all the proposed models.

Table 5: CFA Fit Indices

	(1)	(2)	(3)	(4)	(5)
	Emo. Reg.	Emo. Exhaust.	Stress	Class Mgmt.	All
RMSE	0.097	0.097	0.072	0.115	0.4
CFI	0.924	0.96	0.978	0.968	0.96
TLI	0.9	0.946	0.971	0.955	0.957
SRMR	0.054	0.043	0.039	0.028	0.045

Note: Following Hu and Bentler (1999), we use SRMR ideally below 0.08 and at most 0.10; RMSEA ideally below 0.06 and at most 0.10; and CFI and TLI ideally above 0.95, with a minimum of 0.90.

DISCUSSION

The results of testing the WHAT tool with Salvadoran in-service teachers support the validity and internal consistency reliability of the tool's individual measures—ERQ, MBI-EE, PSS, and OSTES-CM. The cognitive interviews we conducted provided validity evidence based on the contents of the items, in that they matched what participants expressed in the interviews. The EFA verified the unidimensionality of the emotional exhaustion and CMSE subscales and confirmed the two-factor structure of the ERQ and PSS, as found in the existing literature. Each item loaded well on the proposed factor within the different scales and subscales, which provided validity evidence for the internal structure. The results also provide evidence of reliability of the Salvadoran version of the different measurement tools. Internal consistency coefficients were good for all four measures. The intercorrelation among the measures in the tool, and those with other external measures, are in the expected direction, which provides validity evidence based on relations to other variables. Surprisingly, correlations among the four measures and key sociodemographic variables, such as age, marital status, and SES, are often not significant. The EFA that included all the tool items confirmed that each scale is indeed measuring a different construct; the CFA confirmed a good model fit.

Findings from this study support the inclusion of all four of the adapted and translated measures—ERQ, MBI-EE, PSS, and OSTES-CM—into the WHAT tool. We recommend keeping all of the translated and adapted items included under each measure, except for the fourth item of the perceived stress I subscale. Overall, this study confirms that the adapted and translated measures that comprise the WHAT tool are valid and reliable and can be used with Spanish-speaking Salvadoran in-service teachers.

The WHAT tool can be used to identify who experiences stress and emotional exhaustion—that is, which teachers exhibit these feelings and what their background characteristics are. However, it does not identify which contextual and organizational factors may be influencing teachers' level of stress and emotional exhaustion at the national, community, school, and classroom level. As such, differences across schools in terms of the average level of teacher stress and burnout should not automatically be attributed to stressful and challenging contextual characteristics. As Chang (2009) points out, teacher burnout is often a result of an interaction between individual and organizational-contextual factors. Individual characteristics may influence how different teachers respond to the same environmental stressor. Emotion regulation is considered a protective

factor and may help explain differences in teachers' levels of stress and emotional exhaustion; that is, whether high levels of stress are associated with low cognitive reappraisal and a high degree of suppression. On the other hand, CMSE may both explain and be a result of high levels of stress and emotional exhaustion.

In the specific case of this study in El Salvador, descriptive statistics demonstrate that teacher wellbeing is generally positive. The teachers did not experience a high level of emotional exhaustion at work or overall perceived stress. However, the low level of stress observed may be related to the fact that the PSS assesses global perceived stress but does not address how perceptions of stress may be linked to specific contexts (Stress Measurement Network 2018). Many teachers in El Salvador face chronic social adversity, such as living in low socioeconomic neighborhoods where there is a presence of gangs and related violence. The lower stress scores may suggest that habituation normalizes the environment and thus tends to reduce stress. This means that teachers may respond resiliently or adaptively when facing chronic social adversity, thereby minimizing its impact on stress levels. The low average stress levels also may be explained by the geographic concentration of gang-related violence in specific Salvadoran municipalities. Unfortunately, we do not have data to compare teacher wellbeing in the municipalities with low and high levels of violence.

The statistics also show that teachers tend to have a high degree of confidence in their ability to manage disruptive behavior in the classroom. Although the statistics show that teachers employ cognitive reappraisal strategies, they also commonly suppress their emotions. This is concerning, given that suppression has been associated with negative wellbeing outcomes such as depression and pessimism (Barsade and Gibson 2007; Côté and Morgan 2002). Since the ERQ includes general questions about engaging in emotion regulation, it is not possible to know if teachers specifically apply cognitive reappraisal or suppression when handling events in the classroom (e.g., student misbehavior).

The WHAT instrument shows promise in helping education authorities and researchers to measure teachers' wellbeing in the Salvadoran context, and in other conflict- and crisis-affected contexts. District- or national-level education authorities can identify the geographic regions, school clusters, and individual schools where teachers exhibit high levels of perceived stress and/or emotional exhaustion and may need immediate support. Information gathered on emotion regulation and CMSE can provide further understating of the role these variables play as protective factors, and may lead to policy interventions that focus on these

skills in teacher education or during in-service training. For instance, evidence suggests that emotion regulation can be taught through short-term interventions; for example, mindfulness training has been found to improve teachers' self-regulation (Frank et al. 2013) and adaptive emotion regulation (Jennings et al. 2017).

LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Here we highlight several limitations and suggestions for future research. First, a conceptual framework and the Salvadoran RERA (ECCN 2016) informed the selection of constructs for the WHAT tool. Asking Salvadoran educators what they consider important to their wellbeing would have improved the selection of constructs for a contextually appropriate tool. Future adaptations of the existing tool and other measurement research on teacher wellbeing should incorporate local definitions and an understanding of what constitutes teacher wellbeing that is based on teachers' input. Second, our sample includes only teachers who voluntarily participated in the workshop series. While we did not find that attending more sessions was related to any of the teacher wellbeing measures, it is still possible that our sample differs from the population of teachers in targeted departments in ways that are not identifiable. If that is the case, our results are not representative of the population of interest—that is, all teachers in the eight targeted Salvadoran departments. Third, the tool was not validated for program evaluation purposes. Further research is needed to assess whether the measures included are sensitive to program interventions of short duration and are able to detect change over time. Fourth, all the measures included in the tool are self-reported, which assumes that participants reported truly and accurately. Future studies may use performance-based observation measures or biomarkers (e.g., cortisol) to provide further validity evidence for the tool, based on relations to other variables. Fifth, further research is needed to determine the degree of measurement invariance of the measures in the tool to ascertain if they could be used for comparisons across different cultures and geographic locations. Sixth, more systematic research is needed to examine the interrelationship and pathways between the constructs in the tool (e.g., does emotion regulation influence the experience of perceived stress and emotional exhaustion when considering classroom-management challenges?). Seventh, the global nature of the PSS and ERQ may not capture context-specific perceived stress and emotion regulation strategies used in the classroom. Further exploration of context-specific measures is warranted. Finally, we highly recommend conducting further reliability testing of the measures in the tool, such as test-retest.

Further research is needed on alternative uses of the WHAT tool and how it can be adapted for different purposes. The alternative uses we envision for the instrument include individual self-assessments by teachers to support their reflections on their wellbeing and inform their self-improvement, and administration of the tool by school principals and academic leaders to further understand their teachers' sense of wellbeing and to develop supportive actions and professional development opportunities accordingly.

CONCLUSION

This study selected measures for the WHAT tool and examined the validity and reliability of the four measures selected: ERQ, PSS, MBI-EE, and OSTES-CM. The original tools were translated, reviewed, and cognitively tested with a small sample of Salvadoran teachers before being finalized and applied to a larger sample of 1,659 teachers at the preschool, elementary, and secondary levels. To establish the validity and internal consistency of the different tools comprising the instrument, we employed EFA, CFA, and reliability analysis, and assessed their relationship to other external variables.

Results from the psychometric analysis conducted during this study provide validity evidence for content, internal structure, and concurrent relation to other variables for the adapted and contextualized Spanish version of the measures that comprise the WHAT tool (see Appendix C). The four measures also showed good internal reliability levels, and the CFA results confirmed a good model fit. We conclude that the WHAT tool can be used to measure wellbeing among in-service teachers in El Salvador when conceptualized around the constructs of emotion regulation, perceived stress, emotional exhaustion, and CMSE.

For use in other low-income and crisis- and conflict-affected settings, we recommend a rigorous contextual adaptation process with the WHAT tool, including contextual translation, back translation, cognitive interviewing, and pilot testing. Consideration also should be given to the normative nature of teacher wellbeing during the adaptation process, as understanding and definitions of wellbeing may change in keeping with the norms of specific cultural and societal contexts. Given its self-reported nature, the WHAT tool is not resource intensive and does not require intense assessor training. It can be applied in a group setting—that is, to several teachers at a time. This facilitates its application in crisis and conflict settings, which usually are constrained in terms of resources and time.

ACKNOWLEDGMENTS

We thank Annette Brown, Carina Omoeva, Daniel Santos, and Patrick Kyllonen for providing valuable comments and feedback. This study was funded by the Ward Cates Emerging Scientific Leader Award and the Wellspring Philanthropic Fund in support of the Education Equity Research Initiative.

REFERENCES

- Aboagye, Michael Osei, Jinliang Qin, Abdul Qayyum, Collins Opoku Antwi, Yasin Jababu, and Emmanuel Affum-Osei. 2018. "Teacher Burnout in Pre-Schools: A Cross-Cultural Factorial Validity, Measurement Invariance and Latent Mean Comparison of the Maslach Burnout Inventory, Educators Survey (MBI-ES)." *Children and Youth Services Review* 94: 186-97. <https://doi.org/10.1016/j.chilyouth.2018.09.041>.
- Aloe, Ariel M., Laura C. Amo, and Michele E. Shanahan. 2014. "Classroom Management Self- Efficacy and Burnout: A Multivariate Meta-Analysis." *Educational Psychology Review* 26 (1): 101-26. <https://doi.org/10.1007/s10648-013-9244-0>.
- Aluja, Anton, Angel Blanch, and Luis F. García. 2005. "Dimensionality of the Maslach Burnout Inventory in School Teachers." *European Journal of Psychological Assessment* 21 (1): 67-76. <https://doi.org/10.1027/1015-5759.21.1.67>.
- Amirkhan, James H. 2012. "Stress Overload: A New Approach to the Assessment of Stress." *American Journal of Community Psychology* 49 (1-2): 55-71. <https://doi.org/10.1007/s10464-011-9438-x>.
- Armor, David, Patricia Conry-Oseguera, Millicent Cox, Nicelma King, Lorraine McDonnell, Anthony Pascal, Edward Pauly, Gail Zellman. 1976. *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools*. Report No. R-2007-LAUSD. Santa Monica, CA: Rand Corporation. <https://www.rand.org/content/dam/rand/pubs/reports/2005/R2007.pdf>.

- Asociación Civil Diálogos. 2020. "Informe Sobre La Violencia Homicida en El Salvador Diciembre de 2019." Ciudad de Guatemala, Guatemala: Observatorio de la Violencia Asociación Civil Diálogos. <https://dialogos.org.gt/wp-content/uploads/2020/02/Informe-sobre-la-violencia-homicida-en-El-Salvador-al-31-dic-2019.pdf>.
- Asparouhov, Tihomir, and Bengt Muthén. 2009. "Exploratory Structural Equation Modeling." *Structural Equation Modeling: A Multidisciplinary Journal* 16 (3): 397-438. <https://doi.org/10.1080/10705510903008204>.
- Baglin, James. 2014. "Improving your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR." *Practical Assessment, Research, and Evaluation* 19 (1): 5. <https://doi.org/10.7275/dsep-4220>.
- Balzarotti, Stefania, Oliver P. John, and James J. Gross. 2010. "An Italian Adaptation of the Emotion Regulation Questionnaire." *European Journal of Psychological Assessment* 26 (1): 61-67. <https://doi.org/10.1027/1015-5759/a000009>.
- Bandura, Albert. 1977. "Self-Efficacy: Toward a Unifying Theory of Behavioral Change." *Psychological Review* 84 (2): 191-215. <https://doi.org/10.1037/0033-295X.84.2.191>.
- Bandura, Albert. 1997. *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman.
- Barsade, Sigal G., and Donald E. Gibson. 2007. "Why Does Affect Matter in Organizations?" *Academy of Management Perspectives* 21 (1): 36-59. <https://doi.org/10.5465/AMP.2007.24286163>.
- Beaton, Dorcas E., Claire Bombardier, Francis Guillemin, and Marcos Bosi Ferraz. 2000. "Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures." *Spine* 25 (24): 3186-91. <https://doi.org/10.1097/00007632-200012150-00014>.
- Bennell, Paul, and Kwame Akyeampong. 2007. *Teacher Motivation in Sub-Saharan Africa and South Asia*. No. 71. London: Department for International Development.

- Betoret, Fernando Doménech. 2009. "Self-Efficacy, School Resources, Job Stressors and Burnout among Spanish Primary and Secondary School Teachers: A Structural Equation Approach." *Educational Psychology* 29 (1): 45-68. <https://doi.org/10.1080/01443410802459234>.
- Boles, James S., Dwane H. Dean, Joe M. Ricks, Jeremy C. Short, and Guangping Wang. 2000. "The Dimensionality of the Maslach Burnout Inventory across Small Business Owners and Educators." *Journal of Vocational Behavior* 56 (1): 12-34. <https://doi.org/10.1006/jvbe.1999.1689>.
- Brackett, Marc A., Raquel Palomera, Justyna Mojsa-Kaja, Maria Regina Reyes, and Peter Salovey. 2010. "Emotion-Regulation Ability, Burnout, and Job Satisfaction among British Secondary-School Teachers." *Psychology in the Schools* 47 (4): 406-17. <https://doi.org/10.1002/pits.20478>.
- Brotheridge, Céleste M., and Raymond T. Lee. 2003. "Development and Validation of the Emotional Labour Scale." *Journal of Occupational and Organizational Psychology* 76 (3): 365-79. <https://doi.org/10.1348/096317903769647229>.
- Brouwers, André, and Welko Tomic. 2000. "A Longitudinal Study of Teacher Burnout and Perceived Self-Efficacy in Classroom Management." *Teaching and Teacher Education* 16 (2): 239-53. [https://doi.org/10.1016/S0742-051X\(99\)00057-8](https://doi.org/10.1016/S0742-051X(99)00057-8).
- Brown, Janice M., William R. Miller, and Lauren A. Lawendowski. 1999. "The Self-Regulation Questionnaire". In *Innovations in Clinical Practice: A Sourcebook*, edited by Leon Van deCreek, Samuel Knapp and Thomas. L. Jackson, 281-92. Sarasota, FL: Professional Resource Press/Professional Resource Exchange.
- Browne, Michael W. 2001. "An Overview of Analytic Rotation in Exploratory Factor Analysis." *Multivariate Behavioral Research* 36 (1): 111-50. https://doi.org/10.1207/S15327906MBR3601_05.
- Buja, Andreas, and Nermin Eyuboglu. 1992. "Remarks on Parallel Analysis." *Multivariate Behavioral Research* 27 (4): 509-40. https://doi.org/10.1207/s15327906mbr2704_2.
- Bümen, Nilay T. 2010. "The Relationship Between Demographics, Self-Efficacy, and Burnout Among Teachers." *Egitim Arastirmalari-Eurasian Journal of Educational Research* 40: 16-35.

- Burns, Mary, and James Lawrie. 2015. *Where It's Needed Most: Quality Professional Development for All Teachers*. New York: Inter-agency Network for Education in Emergencies.
- Calvete Zumalde, Esther, and Aurelio Villa Sánchez. 1999. "Estrés y Burnout Docente: Influencia de Variables Cognitivas." *Revista de Educación* 319: 291-303.
- Carson, Russell L., Howard M. Weiss, and Thomas J. Templin. 2010. "Ecological Momentary Assessment: A Research Method for Studying the Daily Lives of Teachers." *International Journal of Research and Method in Education* 33: 165-82. <https://doi.org/10.1080/1743727X.2010.484548>.
- Catanzaro, Salvatore J., and Jack Mearns. 1990. "Measuring Generalized Expectancies for Negative Mood Regulation: Initial Scale Development and Implications." *Journal of Personality Assessment* 54 (3-4): 546-63. <https://doi.org/10.1080/00223891.1990.9674019>.
- Chan, David W. 2006. "Emotional Intelligence and Components of Burnout among Chinese Secondary School Teachers in Hong Kong." *Teaching and Teacher Education* 22 (8): 1042-54. <https://doi.org/10.1016/j.tate.2006.04.005>.
- Chang, Mei Lin. 2009. "An Appraisal Perspective of Teacher Burnout: Examining the Emotional Work of Teachers." *Educational Psychology Review* 21 (3): 193-218. <https://doi.org/10.1007/s10648-009-9106-y>.
- Charles, Susan Turk, and Laura L. Carstensen. 2007. "Emotion Regulation and Aging." In *Handbook of Emotion Regulation*, edited by James J. Gross, 307-30. New York: Guilford Publications.
- Chen, Zhibin. 2010. "Assessing Two Emotion Regulation Processes in Chinese Adolescents." *2010 IEEE International Conference on Advanced Management Science* 3: 392-95. <https://doi.org/10.1109/ICAMS.2010.5553216>.
- Cocca, Michaela, Armando Cocca, Elizabeth A. A. Martínez, and María G. Rodríguez Bulnes. 2018. "Correlation Between Self-Efficacy Perception and Teaching Performance: The Case of Mexican Preschool and Primary School Teachers." *Arab World English Journal* 9: 56-70. <https://doi.org/10.2139/ssrn.3150710>.

- Cohen, Sheldon, Tom Kamarck, and Robin Mermelstein. 1983. "A Global Measure of Perceived Stress." *Journal of Health and Social Behavior* 24 (4): 385-96. <https://doi.org/10.2307/2136404>.
- Collie, Rebecca J., Jennifer D. Shapka, Nancy E. Perry, and Andrew J. Martin. 2015. "Teacher Well-Being: Exploring Its Components and a Practice-Oriented Scale." *Journal of Psychoeducational Assessment* 33 (8): 744-56. <https://doi.org/10.1177/0734282915587990>.
- Collins, Debbie. 2003. "Pretesting Survey Instruments: An Overview of Cognitive Methods." *Quality of Life Research* 12 (3): 229-38. <https://doi.org/10.1023/A:1023254226592>.
- Cortina, Jose M. 1993. "What Is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology* 78 (1): 98-104. <https://doi.org/10.1037/0021-9010.78.1.98>.
- Costello, Anna B., and Jason Osborne. 2005. "Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis." *Practical Assessment, Research, and Evaluation* 10 (1): 7. <https://doi.org/10.7275/jyj1-4868>.
- Côté, Stéphane, and Laura M. Morgan. 2002. "A Longitudinal Analysis of the Association between Emotion Regulation, Job Satisfaction, and Intentions to Quit." *Journal of Organizational Behavior* 23 (8): 947-62. <https://doi.org/10.1002/job.174>.
- Curry, Jennifer R., and Elizabeth R. O'Brien. 2012. "Shifting to a Wellness Paradigm in Teacher Education: A Promising Practice for Fostering Teacher Stress Reduction, Burnout Resilience, and Promoting Retention." *Ethical Human Psychology and Psychiatry* 14 (3): 178-91. <https://doi.org/10.1891/1559-4343.14.3.178>.
- D'Argembeau, Arnaud, and Martial Van der Linden. 2006. "Individual Differences in the Phenomenology of Mental Time Travel: The Effect of Vivid Visual Imagery and Emotion Regulation Strategies." *Consciousness and Cognition* 15 (2): 342-50. <https://doi.org/10.1016/j.concog.2005.09.001>.

- Dellinger, Amy B., Jacqueline J. Bobbett, Dianne F. Olivier, and Chad D. Ellett. 2008. "Measuring Teachers' Self-Efficacy Beliefs: Development and Use of the TEBS-Self." *Teaching and Teacher Education* 24 (3): 751-66.
- Dicke, Theresa, Herbert W. Marsh, Philip D. Parker, Mareike Kunter, Annett Schmeck, and Detlev Leutner. 2014. "Self-Efficacy in Classroom Management, Classroom Disturbances, and Emotional Exhaustion: A Moderated Mediation Analysis of Teacher Candidates." *Journal of Educational Psychology* 106 (2): 569-83. <https://doi.org/10.1037/a0035504>.
- Dolan, Emily D., David Mohr, Michele Lempa, Sandra Joos, Stephan D. Fihn, Karin M. Nelson, and Christian D. Helfrich. 2015. "Using a Single Item to Measure Burnout in Primary Care Staff: A Psychometric Evaluation." *Journal of General International Medicine* 30 (5): 582-87. <https://doi.org/10.1007/s11606-014-3112-6>.
- Dorsey, Achshah, Elissa Scherer, Randy Eckhoff, and Robert D. Furberg. 2020. "Measurement of Human Stress: A Multidimensional Approach." Research Triangle Park, NC: RTI International. <https://doi.org/10.31234/osf.io/es2p9>.
- ECCN (Education in Conflict and Crisis Network). 2016. "Rapid Education and Risk Assessment El Salvador: Final Report." <https://www.eccnetwork.net/resources/rapid-education-and-risk-analysis-rera-el-salvador>.
- Emmer, Edmund T., and Julia Hickman. 1991. "Teacher Efficacy in Classroom Management and Discipline." *Educational and Psychological Measurement* 51 (3): 755-65. <https://doi.org/10.1177/0013164491513027>.
- Evers, Will J. G., Welko Tomic, and Andre Brouwers. 2004. "Burnout Among Teachers: Students' and Teachers' Perceptions Compared." *School Psychology International* 25: 131-48. <https://doi.org/10.1177/0143034304043670>.
- Falk, Danielle, Emily Varni, Julia Finder Johna, and Paul Frisoli. 2019. *Landscape Review: Teacher Well-Being in Low Resource, Crisis, and Conflict-affected Settings*. Washington, DC: Education Equity Research Initiative https://app.mhps.net/?get=57/twb-landscape-review_august-2019_0.pdf.
- Fowler, Floyd J. Jr. 1995. *Improving Survey Questions: Design and Evaluation*. Newbury Park, CA: Sage.

- Frank, Jennifer L., Diane Reibel, Patricia Broderick, Todd Cantrell, and Stacie Metz. 2013. "The Effectiveness of Mindfulness-Based Stress Reduction on Educator Stress and Well-Being: Results from a Pilot Study." *Mindfulness* 6 (2): 208-16. <https://doi.org/10.1007/s12671-013-0246-2>.
- Friedman, Isaac A., and Efrat Kass. 2002. "Teacher Self-Efficacy: A Classroom Organization Conceptualization." *Teaching and Teacher Education* 18 (6): 675-86. [https://doi.org/10.1016/S0742-051X\(02\)00027-6](https://doi.org/10.1016/S0742-051X(02)00027-6).
- Frisoli, Paul St. John. 2013. "Teachers' Experiences of Professional Development in (Post) Crisis Katanga Province, South Eastern Democratic Republic of Congo: A Case Study of Teacher Learning Circles." PhD diss., University of Massachusetts Amherst.
- Fullan, Michael. 2016. *The New Meaning of Educational Change*, 5th ed. New York: Teachers College Press.
- Gastaldi, Francesca Giovanna Maria, Tiziana Pasta, Claudio Longobardi, Laura Elvira Prino, and Rocco Quaglia. 2014. "Measuring the Influence of Stress and Burnout in Teacher-Child Relationship." *European Journal of Education and Psychology* 7 (1): 17-28. <https://doi.org/10.30552/ejep.v7i1.99>.
- Gibson, Sherri, and Myron H. Dembo. 1984. "Teacher Efficacy: A Construct Validation." *Journal of Educational Psychology* 76: 569-82. <https://doi.org/10.1037/0022-0663.76.4.569>.
- González, Mónica-Teresa. 2006. *Confirmación de un Modelo Predictivo para la Salud, a través de Ecuaciones Estructurales*. Madrid: Universidad Nacional de Educación a Distancia.
- Gratz, Kim L., and Lizabeth Roemer. 2004. "Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale." *Journal of Psychopathology and Behavioral Assessment* 26 (1): 41-54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>.
- Grayson, Jessica L., and Heather K. Alvarez. 2008. "School Climate Factors Relating to Teacher Burnout: A Mediator Model." *Teaching and Teacher Education* 24 (5): 1349-63. <https://doi.org/10.1016/j.tate.2007.06.005>.

- Gregoire, Jacques. 2018. "ITC Guidelines for Translating and Adapting Tests." *International Journal of Testing* 18 (2): 101-34.
- Grenville-Cleave, Bridget, and Ilona Boniwell. 2012. "Surviving or Thriving? Do Teachers Have Lower Perceived Control and Well-Being Than Other Professions?" *Management in Education* 26 (1): 3-5. <https://doi.org/10.1177/0892020611429252>.
- Gross, James J. 1998a. "The Emerging Field of Emotion Regulation: An Integrative Review." *Review of General Psychology* 2 (3): 271-99.
- Gross, James J. 1998b. "Antecedent-and Response-Focused Emotion Regulation: Divergent Consequences for Experience, Expression, and Physiology." *Journal of Personality and Social Psychology* 74 (1): 224.
- Gross, James J. 2002. "Emotion Regulation: Affective, Cognitive, and Social Consequences." *Psychophysiology* 39 (3): 281-91. <https://doi.org/10.1017/S0048577201393198>.
- Gross, James J., and Oliver P. John. 2003. "Individual Differences in Two Emotion Regulation Processes: Implications for Affect, Relationships, and Well-Being." *Journal of Personality and Social Psychology* 85 (2): 348-62. <https://doi.org/10.1037/0022-3514.85.2.348>.
- Groves, Robert. M., Floyd. J. Fowler Jr., Mick. P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2011. *Survey Methodology*, Vol. 561. New York: John Wiley & Sons.
- Guskey, Thomas R. 1981. "Measurement of the Responsibility Teachers Assume for Academic Successes and Failures in the Classroom." *Journal of Teacher Education* 32 (3): 44-51. <https://doi.org/10.1177/002248718103200310>.
- Halbesleben, Jonathon R. B., and Evangelia Demerouti. 2005. "The Construct Validity of an Alternative Measure of Burnout: Investigating the English Translation of the Oldenburg Burnout Inventory." *Work and Stress* 19 (3): 208-20. <https://doi.org/10.1080/02678370500340728>.
- Hu, Li-tze, and Peter M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling* 6 (1): 1-55. <https://doi.org/10.1080/10705519909540118>.

- INEE (Inter-agency Network for Education in Emergencies). 2010. *Guidance Notes on Teaching and Learning*. New York: INEE. https://inee.org/system/files/resources/INEE_Guidance_Notes_on_Teaching_and_Learning_EN.pdf.
- Ingram, Mathew C., and Karise M. Curtis. 2014. *Homicide in El Salvador's Municipalities: Spatial Clusters and the Causal Role of Neighborhood Effects, Population Pressures, Poverty, and Education*. Washington, DC: Wilson Center. <https://www.wilsoncenter.org/publication/homicide-el-salvadors-municipalities-spatial-clusters-and-the-causal-role-neighborhood>.
- Jennings, Patricia A., Joshua L. Brown, Jennifer L. Frank, Sebrina Doyle, Yoonkyung Oh, Regin Davis, Damira Rasheed et al. 2017. "Impacts of the CARE for Teachers Program on Teachers' Social and Emotional Competence and Classroom Interactions." *Journal of Educational Psychology* 109 (7): 1010-28. <https://doi.org/10.1037/edu0000187>.
- John, Oliver P., and Joshua Eng. 2014. "Three Approaches to Individual Differences in Affect Regulation: Conceptualizations, Measures, and Findings." In *Handbook of Emotion Regulation*, edited by James J. Gross, 321-45. New York: Guilford Publications.
- Kaiser, Henry F., and John Rice. 1974. "Little Jiffy, Mark IV." *Educational and Psychological Measurement* 34 (1): 111-17.
- Kirk, Jackie, and Rebecca Winthrop. 2013. "Teaching in Contexts of Emergency and State Fragility." In *More and Better Teachers for Quality Education for All: Identity and Motivation, Systems and Support*, edited by Jackie Kirk, Martial Dembélé, and Sandra Baxter, 121-39. https://moreandbetterteachers.files.wordpress.com/2013/09/more-and-better-teaching_september-2013.pdf.
- Klusmann, Uta, Mareike Kunter, Ulrich Trautwein, Oliver Lüdtke, and Jürgen Baumert. 2008. "Teachers' Occupational Well-Being and Quality of Instruction: The Important Role of Self-Regulatory Patterns." *Journal of Educational Psychology* 100 (3): 702-15. <https://doi.org/10.1037/0022-0663.100.3.702>.
- Kristensen, Tage S., Marianne Borritz, Ebbe Villadsen, and Karl B. Christensen. 2005. "The Copenhagen Burnout Inventory: A New Tool for the Assessment of Burnout." *Work and Stress* 19 (3): 192-207. <https://doi.org/10.1080/02678370500297720>.

- Kroenke, Kurt, Robert L. Spitzer, Janet B. W. Williams, and Bernd Löwe. 2010. "The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A Systematic Review." *General Hospital Psychiatry* 32 (4): 345-59. <https://doi.org/10.1016/j.genhosppsy.2010.03.006>.
- Lau, Patrick S. Y., Man Tak Yuen, and Raymond M. C. Chan. 2005. "Do Demographic Characteristics Make a Difference to Burnout among Hong Kong Secondary School Teachers?" In *Quality-of-Life Research in Chinese, Western and Global Contexts*, edited by Daniel T. L. Shek, Ying Keung Chan, and Paul S. N. Lee, 491-516. Dordrecht, The Netherlands: Springer Science & Business Media. https://doi.org/10.1007/1-4020-3602-7_17.
- Lee, Eun Hyun. 2012. "Review of the Psychometric Evidence of the Perceived Stress Scale." *Asian Nursing Research* 6 (4): 121-27. <https://doi.org/10.1016/j.anr.2012.08.004>.
- Leiter, Michael P., and Christina Maslach. 2016. "Latent Burnout Profiles: A New Approach to Understanding the Burnout Experience." *Burnout Research* 3 (4): 89-100. <https://doi.org/10.1016/j.burn.2016.09.001>.
- Lorenzo-Seva, Urbano, and Pere J. Ferrando. 2006. "FACTOR: A Computer Program to Fit the Exploratory Factor Analysis Model." *Behavior Research Methods* 38: 88-91. <https://doi.org/10.3758/BF03192753>.
- MacCallum, Robert C., Mary Roznowski, and Lawrence B. Necowitz. 1992. "Model Modifications in Covariance Structure Analysis: The Problem of Capitalization on Chance." *Psychological Bulletin* 111 (3): 490-504. <https://doi.org/10.1037/0033-2909.111.3.490>.
- Malach-Pines, Ayala. 2005. "The Burnout Measure, Short Version." *International Journal of Stress Management* 12 (1): 78-88. <https://doi.org/10.1037/1072-5245.12.1.78>.
- Maslach, Christina, Susan E. Jackson, and Michael P. Leiter. 1997. *Maslach Burnout Inventory Manual*, Vol. 4. Palo Alto, CA: Consulting Psychologists Press.
- Maslach, Christina, and Michael P. Leiter. 2016. "Understanding the Burnout Experience: Recent Research and Its Implications for Psychiatry." *World Psychiatry* 15 (2): 103-11. <https://doi.org/10.1002/wps.20311>.

- Mattern, Jessica, and Johannes Bauer. 2014. "Does Teachers' Cognitive Self-Regulation Increase Their Occupational Well-Being? The Structure and Role of Self-Regulation in the Teaching Context." *Teaching and Teacher Education* 43: 58-68. <https://doi.org/10.1016/j.tate.2014.05.004>.
- Mayer, John D., Peter Salovey, David R. Caruso, and Gill Sitarenios. 2003. "Measuring Emotional Intelligence with the MSCEIT V2.0." *Emotion* 3 (1): 97-105. <https://doi.org/10.1037/1528-3542.3.1.97>.
- MINED. 2015. "Observatorio MINED 2015 sobre los Centros Educativos Públicos de El Salvador." San Salvador, El Salvador: MINED. <https://www.mined.gob.sv/EstadisticaWeb/observatorio/OBSERVATORIO%20MINED%202015.pdf>.
- Montgomery, Cameron, and André A. Rupp. 2005. "A Meta-Analysis for Exploring the Diverse Causes and Effects of Stress in Teachers." *Canadian Journal of Education* 28 (3): 458-86. <https://doi.org/10.2307/4126479>.
- Muthén, Bengt, and David Kaplan. 1985. "A Comparison of Some Methodologies for the Factor Analysis of Non-Normal Likert Variables." *British Journal of Mathematical and Statistical Psychology* 38 (2): 171-89. <https://doi.org/10.1111/j.2044-8317.1985.tb00832.x>.
- Muthén, Bengt O., Linda Muthén, and Tihomir Asparouhov. 2015. *Estimator Choices with Categorical Outcomes*. MPLUS Technical Appendix, 8. Los Angeles, CA: Muthén & Muthén. <https://www.statmodel.com/download/EstimatorChoices.pdf>.
- Norman, Geoff. 2010. "Likert Scales, Levels of Measurement and the 'Laws' of Statistics." *Advances in Health Science Education* 15 (5): 625-32. <https://doi.org/10.1007/s10459-010-9222-y>.
- OECD. 2013. *OECD Guidelines on Measuring Subjective Well-Being*. Paris: OECD. <https://www.oecd.org/statistics/oecd-guidelines-on-measuring-subjective-well-being-9789264191655-en.htm>.
- OECD. 2017. *How's Life? Measuring Well-Being*. Paris: OECD. https://doi.org/10.1787/how_life-2017-en.
- OECD. 2018. *TALIS: The OECD Teaching and Learning International Survey*. Paris: OECD. <http://www.oecd.org/edu/talis>.

- O'Neill, Sue Catherine, and Jennifer Stephenson. 2011. "The Measurement of Classroom Management Self-Efficacy: A Review of Measurement Instrument Development and Influences." *Educational Psychology* 31: 261-99. <https://doi.org/10.1080/01443410.2010.545344>.
- Osborne, Jason W., and Anna B. Costello. 2004. "Sample Size and Subject to Item Ratio in Principal Components Analysis." *Practical Assessment, Research and Evaluation* 9 (1): 1-15. <https://doi.org/10.7275/ktzq-jq66>.
- Ramírez, Mónica Teresa González, and René Landero Hernández. 2007. "Factor Structure of the Perceived Stress Scale (PSS) in a Sample from Mexico." *The Spanish Journal of Psychology* 10 (1): 199-206. <https://doi.org/10.1017/S1138741600006466>.
- Raykov, Tenko, and George A. Marcoulides. 2008. *An Introduction to Applied Multivariate Analysis*. New York: Routledge.
- Renshaw, Tyler L., Anna C. J. Long, and Clayton R. Cook. 2015. "Assessing Teachers' Positive Psychological Functioning at Work: Development and Validation of the Teacher Subjective Wellbeing Questionnaire." *School Psychology Quarterly* 30 (2): 289-306.
- Roeser, Robert W., Kimberly A. Schonert-Reichl, Amishi Jha, Margaret Cullen, Linda Wallace, Rona Wilensky, Eva Oberle, Kimberly Thomson, Cynthia Taylor, and Jessica Harrison. 2013. "Mindfulness Training and Reductions in Teacher Stress and Burnout: Results from Two Randomized, Waitlist-Control Field Trials." *Journal of Educational Psychology* 105 (3): 787-804. <https://doi.org/10.1037/a0032093>.
- Rotter, Julian B. 1966. "Generalized Expectancies for Internal versus External Control of Reinforcement." *Psychological Monographs: General and Applied* 80 (1): 1-28. <https://doi.org/10.1037/h0092976>.
- Saarni, Carolyn. 1999. *The Development of Emotional Competence*. New York: Guilford Press.
- Saarni, Carolyn. 2011. "Emotional Development in Childhood." In *Encyclopedia on Early Childhood Development*, edited by Doris Fromberg and Leslie Williams, 1-7. London: Routledge.

- Salmela-Aro, Katariina, Johanna Rantanen, Katriina Hyvönen, Kati Tilleman, and Taru Feldt. 2011. "Bergen Burnout Inventory: Reliability and Validity among Finnish and Estonian Managers." *International Archives of Occupational and Environmental Health* 84 (6): 635-45. <https://doi.org/10.1007/s00420-010-0594-3>.
- Salovey, Peter, John D. Mayer, Susan Lee Goldman, Carolyn Turvey, and Tibor P. Palfai. 1995. "Emotional Attention, Clarity, and Repair: Exploring Emotional Intelligence Using the Trait Meta-Mood Scale." In *Emotion, Disclosure, & Health*, edited by James W. Pennebaker, 125-54. Washington, DC: American Psychological Association. <https://doi.org/10.1037/10182-006>.
- Saloviita, Timo. 2015. "Measuring Pre-Service Teachers' Attitudes Towards Inclusive Education: Psychometric Properties of the TAIS Scale." *Teaching and Teacher Education* 52: 66-72. <https://doi.org/10.1016/j.tate.2015.09.003>.
- Schaufeli, Wilmar B., and Toon W. Taris. 2005. "The Conceptualization and Measurement of Burnout: Common Ground and Worlds Apart." *Work and Stress* 19 (3): 256-62. <https://doi.org/10.1080/02678370500385913>.
- Schonfeld, Irvin Sam, and Renzo Bianchi. 2016. "Burnout and Depression: Two Entities or One?" *Journal of Clinical Psychology* 72 (1): 22-37. <https://doi.org/10.1002/jclp.22229>.
- Shirom, Arie. 1989. "Burnout in Work Organizations." In *International Review of Industrial and Organizational Psychology*, edited by C. L. Cooper and I. T. Robertson, 25-48. Chichester, UK: Wiley.
- Shirom, Arie, and Samuel Melamed. 2006. "A Comparison of the Construct Validity of Two Burnout Measures in Two Groups of Professionals." *International Journal of Stress Management* 13 (2): 176-200. <https://doi.org/10.1037/1072-5245.13.2.176>.
- Skaalvik, Einar M., and Sidsel Skaalvik. 2007. "Dimensions of Teacher Self-Efficacy and Relations with Strain Factors, Perceived Collective Teacher Efficacy, and Teacher Burnout." *Journal of Educational Psychology* 99 (3): 611-25. <https://doi.org/10.1037/0022-0663.99.3.611>.

- Skaalvik, Einar M., and Sidsel Skaalvik. 2010. "Teacher Self-Efficacy and Teacher Burnout: A Study of Relations." *Teaching and Teacher Education* 26 (4): 1059-69. <https://doi.org/10.1016/j.tate.2009.11.001>.
- Sommers, Marc. 2004. *Education in Emergencies: Critical Questions and Challenges*. Washington, DC: Creative Associates International. http://www.beps.net/publications/Education_Emergency_Changes.pdf.
- Spaapen, Devon L., Flavie Waters, Laura Brummer, Lusia Stopa, and Romola S. Bucks. 2014. "The Emotion Regulation Questionnaire: Validation of the ERQ-9 in Two Community Samples." *Psychological Assessment* 26 (1): 46-54. <https://doi.org/10.1037/a0034474>.
- Spitzer, Robert L., Kurt Kroenke, Janet B. Williams, and Bernd Löwe. 2006. "A Brief Measure for Assessing Generalized Anxiety Disorder: the GAD-7." *Archives of Internal Medicine* 166 (10): 1092-97.
- Stress Measurement Network. 2018. "Stress Measurement Toolbox." San Francisco, CA: University of California, National Institute on Aging. https://stresscenter.ucsf.edu/sites/stresscenter.ucsf.edu/files/Stress%20Measurement%20Toolbox_3%2022%2018.pdf.
- Sutton, Rosemary E., and Karl F. Wheatley. 2003. "Teachers' Emotions and Teaching: A Review of the Literature and Directions for Future Research." *Educational Psychology Review* 15: 327-58. <https://doi.org/10.1023/A:1026131715856>.
- Thompson, Bruce. 2004. *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. Washington, DC: American Psychological Association.
- Thompson, Edmund R. 2007. "Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS)." *Journal of Cross-Cultural Psychology* 38 (2): 227-42. <https://doi.org/10.1177/0022022106297301>.
- Timmerman, Marieke E., and Urbano Lorenzo-Seva. 2011. Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. *Psychological Methods* 16 (2): 209-20. <https://doi.org/10.1037/a0023353>.

- Tschannen-Moran, Megan, and Anita Woolfolk Hoy. 2001. "Teacher Efficacy: Capturing an Elusive Construct." *Teaching and Teacher Education* 17 (7): 783-805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1).
- Tsouloupas, Costas N., Russell L. Carson, Russell Matthews, Matthew J. Grawitch, and Larissa K. Barber. 2010. "Exploring the Association between Teachers' Perceived Student Misbehaviour and Emotional Exhaustion: The Importance of Teacher Efficacy Beliefs and Emotion Regulation." *Educational Psychology* 30 (2): 173-89. <https://doi.org/10.1080/01443410903494460>.
- Van Der Borgh, Chris, and Wim Savenije. 2019. "The Politics of Violence Reduction: Making and Unmaking the Salvadorean Gang Truce." *Journal of Latin American Studies* 51 (4): 905-28. <https://doi.org/10.1017/S0022216X19000890>.
- Van Horn, Joan E., Toon W. Taris, Wilmar B. Schaufeli, and Paul J. G. Schreurs. 2004. "The Structure of Occupational Well-Being: A Study among Dutch Teachers." *Journal of Occupational and Organizational Psychology* 77 (3): 365-75. <https://doi.org/10.1348/0963179041752718>.
- Wessels, Elsabé, and Lesley Wood. 2019. "Fostering Teachers' Experiences of Well-Being: A Participatory Action Learning and Action Research Approach." *South African Journal of Education* 39 (1): 1-10. <https://doi.org/10.15700/saje.v39n1a1619>.
- Wiltink, Jörg, Heide Glaesmer, Marco Canterino, Klaus Wölfling, Achim Knebel, Henrik Kessler, and Manfred E. Beutel. 2011. "Emotion Regulation in the General Population-Population Characteristics of the Emotion Regulation Questionnaire (ERQ)." *Psychotherapie Psychosomatik Medizinische Psychologie* 61 (2): 117-18.
- Winthrop, Rebecca, and Jackie Kirk. 2005. "Teacher Development and Student Well-Being." *Forced Migration Review* 22: 18-21. <http://www.fmreview.org/FMRpdfs/FMR22/FMR2209.pdf>.
- Wolf, Sharon, Catalina Torrente, Marissa McCoy, Damira Rasheed, and J. Lawrence Aber. 2015. "Cumulative Risk and Teacher Well-Being in the Democratic Republic of the Congo." *Comparative Education Review* 59 (4): 717-42. <https://doi.org/10.1086/682902>.

Zee, Marjolein, and Helma M. Y. Koomen. 2016. "Teacher Self-Efficacy and Its Effects on Classroom Processes, Student Academic Adjustment, and Teacher Well-Being: A Synthesis of 40 Years of Research." *Review of Educational research* 86 (4): 981-1015. <https://doi.org/10.3102/0034654315626801>.

Zelkowitz, Rachel L., and David A. Cole. 2016. "Measures of Emotion Reactivity and Emotion Regulation: Convergent and Discriminant Validity." *Personality and Individual Differences* 102: 123-32.

APPENDIX A

REVIEW OF EXISTING MEASURES

To select the measures that are part of the WHAT tool, we developed an inventory of available measures that assess each of the constructs identified: perceived stress, emotional exhaustion, emotion regulation, and CMSE. To develop this inventory, we first identified a set of measures that have been widely used in our professional context to evaluate teacher wellbeing interventions, and tools whose development and validation have been described in articles published in well-known journals. We then conducted an unsystematic literature review (rather than a systematic comprehensive database search) to identify additional measures for each construct of interest. Table A1 provides an inventory of the main instruments being developed and used in the field, which we identified through our previous knowledge of existing tools and a complementary literature review; it does not provide an exhaustive list of the measurement tools available.

Table A1: Measurement Review

Scale	N of Items	Reliability	Structure	Source
Emotion Regulation				
Generalized Expectancy for Negative Mood Regulation Scale	30	0.86-0.90	2 factors	Catanzaro and Mearns (1990)
Trait Meta-Mood Scale	48	0.82-0.87	3 factors	Salovey et al. (1995)
Self-Regulation Questionnaire	63	0.91	7 factors	Brown et al. (1999)
Managing emotions subscale: Mayer-Salovey-Caruso Emotional Intelligence Test V2.0	29	0.64-0.69	1 factor (2 tasks)	Mayer et al. (2003)
Emotional Labor Scale	15	0.74 -0.91	6 factors	Brotheridge and Lee (2003)
Emotion Regulation Questionnaire	10	0.73-0.79	2 factors	Gross and John (2003)
Difficulties in Emotion Regulation Scale	36	> 0.80	6 factors	Gratz and Roemer (2004)
Perceived Stress				
Perceived Stress Scale (PSS)	14, 10, 4	0.72-0.86	-	Cohen et al. (1983)
Stress Overload Scale (SOS)	30	0.96	2 factors	Amirkhan (2012)
Stress in Context (SIC) Questionnaire	N/A	N/A	N/A	Stress Measurement Network (2018)
Emotional Exhaustion				
Maslach Burnout Inventory–Emotional Exhaustion subscale	9	0.90	1 factor	Maslach et al. (1997)
Bergen Burnout Inventory–Exhaustion at work subscale	3	0.67-0.75	1 factor	Salmela-Aro et al. (2011)
Oldenburg Burnout Inventory–Exhaustion subscale	8	0.74	1 factor	Halbesleben and Demerouti (2005)
The Shirom-Melamed Burnout Measure Emotional Exhaustion subscale	5	N/A	1 factor	Shirom and Melamed (2006)

DEVELOPING THE WELLBEING HOLISTIC ASSESSMENT FOR TEACHERS

Scale	N of Items	Reliability	Structure	Source
The Copenhagen Burnout Inventory	19	0.85-0.87	3 factors	Kristensen et al. (2005)
The Burnout Measure	10	0.85-0.92	3 factors	Malach-Pines (2005)
Single-item measure of burnout	1	-	-	Dolan et al. (2015)
Classroom Management Self-Efficacy				
Classroom-Management and Discipline Self-Efficacy subscale	14	0.81	1 factor	Emmer and Hickman (1991)
Classroom-Management subscale of the Ohio State Teacher Efficacy Scale (CM-OSTES)	8, 4	0.90	1 factor	Tschannen-Moran and Hoy (2001)
Discipline subscale of the Norwegian Teacher Self-Efficacy Scale	4	0.90	1 factor	Skaalvik and Skaalvik (2007)
Discipline Control subscale of Teacher Professional Capability Scale	3	-	1 factor	Friedman and Kass (2002)
Teacher-perceived self-efficacy in classroom management	4	0.87	1 factor	Betoret (2009)
Classroom-Management subscale of the Teachers' Efficacy Beliefs System-Self	10	0.85–0.87	1 factor	Dellinger et al. (2008)

We acknowledge that there are existing frameworks that include broader aspects of wellbeing, such as material conditions, physical health, social support, and civic engagement, and that measure subjective wellbeing. For instance, the Organisation for Economic Co-operation and Development (OECD 2013, 2017) framework for measuring wellbeing and progress and the Gallup World Poll both measure life satisfaction and assess past experiences of negative and positive emotions, such as anger, sadness, stress, and enjoyment, as a measure of affective wellbeing. In addition, recent measures that focused on teachers have included positive indicators of wellbeing. The OECD Talis survey (2018) measures teacher wellbeing by focusing on self-efficacy and job satisfaction. The Teacher Subjective Wellbeing Questionnaire captures school connectedness and teaching efficacy (Renshaw, Long, and Cook 2015), whereas the Teacher Well-Being Scale measures three factors of wellbeing—workload wellbeing, organizational wellbeing, and student interaction wellbeing—by tapping into the determinants of these constructs (Collie et al. 2015).

These existing frameworks tend to focus on positive indicators of wellbeing and do not directly respond to the individual constructs identified in the conceptual framework we proposed in this study. While some of them include measures of self-efficacy, they all lack individual measures of the specific elements of our teacher wellbeing framework that are important to education in emergencies settings, such as occupational burnout and stress. They also do not measure emotion regulation as a key protective factor. Given these gaps, we constructed a multidimensional tool of educator wellbeing that captured the specific constructs of interest.

MEASURING EMOTION REGULATION

Specific definitions of emotion regulation and the corresponding measurement approaches vary by theory, with the conceptions proposed by Gross (1998a, 1998b) and Saarni (1999, 2011) being the most popular. Gross adopts a process-oriented conception, which defines emotion regulation as “processes by which individuals influence which emotions they have, when they have them, and how they experience and express these emotions” (275). This model distinguishes emotion regulation strategies as antecedent or response focused, “referring to when these cognitive events occur along the timeline of information processing” (Spaapen et al. 2014, 46). One of the most commonly used measures of emotion regulation, the Emotion Regulation Questionnaire (ERQ) (Gross and John 2003), was based on this model (Zelkowitz and Cole 2016). The ERQ has been translated into 33 languages and widely applied, but most studies to date have only analyzed the ERQ’s factor structure with university student populations (Spaapen et al. 2014).

Saarni (1999, 2011) postulates a competency-focused model that delineates skills considered prerequisites of emotional competency. John and Eng (2014) argue that the emotional competency approach is much broader than the specific emotion regulation approach, as it includes a host of processes, skills, and competences that do not directly regulate emotions but relate to individual behavior that is considered socioemotionally appropriate. Saarni (1999, 2011) defines emotion regulation as (1) awareness of one's emotional state; (2) skill in discerning and understanding others' emotions; (3) skill in using the vocabulary of emotion and expression; (4) capacity for empathic and sympathetic involvement in others' emotional experiences; (5) skill in realizing that an individual's inner emotional state need not correspond to their outer expression; (6) skill in modulating emotional reactions; (7) awareness that the structure or nature of relationships is defined in part by how emotion is communicated; and (8) a capacity for emotional self-efficacy (Saarni 2011). Although not all eight of these dimensions are agreed to by all researchers, a review by John and Eng (2014) states that measures under this tradition include the Difficulties in Emotion Regulation Scale (Gratz and Roemer 2004), the Generalized Expectancy for Negative Mood Regulation Scale (Catanzaro and Mearns 1990), the Mayer-Salovey-Caruso Emotional Intelligence Test V2.0 (Mayer et al. 2003), and the Trait Meta-Mood Scale (Salovey et al. 1995). The Negative Mood Regulation and Trait Meta-Mood scales have been criticized for equating emotion regulation with emotional avoidance, and for not including all relevant dimensions of the competency-based approach (John and Eng 2014; Gratz and Roemer 2004). Concerns have been raised about the Mayer-Salovey-Caruso Test over validity evidence based on content and relationship to other variables, and the measure has been criticized for measuring individuals' capacity to reason about emotion regulation rather than capturing individual differences in affective regulation. We consider these criticisms of measures that adopt a competency-focused model in the step of scale selection.

While most measures we reviewed present acceptable reliability, they vary greatly in the number of items included (10 to 63) and the number of factors (1 to 6). Measures that adopt a competency-focused model tend to be longer (i.e., have more items) than those that adopt a process-oriented conception and to measure more than one latent factor.

MEASURING PERCEIVED STRESS

Dorsey and colleagues (2020) define stress as a "multi-dimensional construct that is comprised of exposure to events, perceptions of stress, and physiological responses to stress" (2). We reviewed measures that have been specifically developed to assess perceptions of stress. Measures commonly used in the field include the

Perceived Stress Scale (PSS) and the Stress Overload Scale. The PSS (4-, 10-, and 14-item versions) measures the degree to which an individual perceives his/her life as “unpredictable, uncontrollable, and overloading” (Cohen, Kamarck, and Mermelstein 1983, 387) within the past month. It assesses global stress perceptions, and as such can be used with any population and in any context. The Stress Overload Scale measures stress overload, a state that occurs when demands overwhelm resources (Amirkhan 2012). It is comprised of 30 items and 2 subscales, personal vulnerability and event load.

MEASURING EMOTIONAL EXHAUSTION

Emotional exhaustion is one of three dimensions of burnout, also described as “wearing out, loss of energy, depletion, debilitation, and fatigue” (Leiter and Maslach 2016, 89-100). Burnout more broadly is a psychological syndrome defined as a “prolonged response to chronic interpersonal stressors on the job” (103). The three key dimensions of burnout include exhaustion, cynicism and detachment, and a lack of accomplishment at work. Different measures have been developed to either assess several dimensions of burnout or assess the sole dimension of exhaustion. Albeit debated, burnout in some cases has been simplified to a one-dimensional construct focused on exhaustion, as exhaustion is often considered its primary element and potentially a suitable proxy for the burnout construct (Maslach and Leiter 2016).

Our review found three measures of burnout that included an exhaustion subscale: the Bergen Burnout Inventory assesses exhaustion at work, the Oldenburg Burnout Inventory measures physical, affective, and cognitive exhaustion; and the MBI includes an emotional exhaustion subscale. The MBI is the most common measure of burnout (Aloe et al. 2014) and is considered the gold standard (Schaufeli and Taris 2005), given that it is used in more than 90 percent of the studies conducted on the syndrome (Shirom and Melamed 2006). It has been translated into and validated in many languages. Three burnout measures focus on exhaustion alone, although they assess different facets. The Shirom-Melamed Burnout Measure conceptualizes burnout as the depletion of energetic resources and makes a distinction between physical, emotional, and cognitive exhaustion. The Copenhagen Burnout Inventory makes a distinction between physical and psychological exhaustion. The Burnout Measure distinguishes between physical, emotional, and mental exhaustion.

Most of the emotional exhaustion scales or subscales we reviewed have a low number of items (10 or fewer) and only measure one latent factor. While most present decent reliability, the MBI emotional exhaustion subscale has the strongest reliability (0.90) of all the measures reviewed.

MEASURING CLASSROOM-MANAGEMENT SELF-EFFICACY

The conceptualization of teacher self-efficacy is based on two foundational literature strands (Tschannen-Moran and Hoy 2001); one uses Rotter's (1966) social learning theory and the locus of control concept as a theoretical foundation; the other is based on Bandura's (1977) social cognitive theory and his construct of self-efficacy. While Rotter's (1966) informed earlier measurement efforts (Armor et al. 1976; Guskey 1981), later attempts to measure the construct of self-efficacy drew from Bandura's (1977) conceptualization (Gibson and Dembo 1984; Emmer and Hickman 1991). Bandura defines perceived self-efficacy as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (3). While teacher self-efficacy was originally conceived as a single construct, later studies recognized its multidimensional nature (Aloe et al. 2014).

Classroom-management self-efficacy is a domain of teacher self-efficacy that is broadly defined as a teacher's perceived competency in organizing a classroom, maintaining order, proactively managing disruptions, and gaining the participation and attention of all students (Aloe et al. 2014; Tschannen-Moran and Hoy 2001; Brouwers and Tomic 2000; Emmer and Hickman 1991). However, this construct has been conceptualized in different ways, which has led to variability in the instruments used (Aloe et al. 2014). In a review of measurement scales with classroom management items, O'Neill and Stephenson (2011) categorize CMSE into six categories: classroom organization; establishing and maintaining routines and expectations; gaining and maintaining student attention; facilitating student cooperation; maintaining order and control; and general classroom management. The authors find that the different CMSE scales include different CMSE categories in their measurement, which reflects the different approaches to conceptualization and measurement. Most scales we reviewed included items on maintaining order and control; the least common were those pertaining to resource allocation. We reviewed CMSE scales that align with the operationalization proposed by Aloe et al. (2014), which characterizes classroom management self-efficacy as "controlling disruptive behavior, calming and responding to defiant students, and establishing a routine and order to keep learning activities running smoothly" (105). Given the issues of discipline and disruptive behavior associated with a gang-related presence in the school or community that might be expected in Salvadoran classrooms, we focused on measures that align with this conceptualization. We found only one scale designed to measure CMSE as a single domain and five CMSE subscales from broader self-efficacy scales. The scales and subscales ranged from 3 to 14 items and all had reliability coefficients above .8.

APPENDIX B

PSYCHOMETRIC RESULTS

Table B1: Item Correlation by Scale (polychoric correlation)

Emotion regulation reappraisal		(1)	(2)	(3)	(4)	(5)	(6)			
(1)	When I want to feel more positive emotion (such as joy or amusement)...	1.00								
(2)	When I want to feel less negative emotion (such as sadness or anger)...	0.42	1.00							
(3)	When I'm faced with a stressful situation, I make myself think about it...	0.29	0.40	1.00						
(4)	When I want to feel more positive emotion, I change the way...	0.38	0.36	0.41	1.00					
(5)	I control my emotions by changing the way I think about the situation I'm in.	0.28	0.35	0.41	0.57	1.00				
(6)	When I want to feel less negative emotion, I change the way I'm thinking...	0.29	0.42	0.39	0.49	0.55	1.00			
Emotion regulation suppression		(1)	(2)	(3)	(4)					
(1)	I keep my emotions to myself.	1.00								
(2)	When I am feeling positive emotions, I am careful not to express them.	0.59	1.00							
(3)	I control my emotions by not expressing them.	0.51	0.58	1.00						
(4)	When I am feeling negative emotions, I make sure not to express them.	0.33	0.31	0.49	1.00					

DEVELOPING THE WELLBEING HOLISTIC ASSESSMENT FOR TEACHERS

Emotion regulation reappraisal		(1)	(2)	(3)	(4)	(5)	(6)			
Emotion exhaustion		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	Drained	1.00								
(2)	Used up	0.74	1.00							
(3)	Fatigued	0.63	0.63	1.00						
(4)	End of my rope	0.57	0.57	0.59	1.00					
(5)	Burned out	0.66	0.63	0.68	0.70	1.00				
(6)	Frustrated	0.47	0.43	0.58	0.47	0.64	1.00			
(7)	Working too hard	0.46	0.43	0.46	0.48	0.60	0.42	1.00		
(8)	Stress	0.44	0.47	0.45	0.61	0.58	0.52	0.45	1.00	
(9)	Strain	0.50	0.44	0.55	0.54	0.69	0.63	0.53	0.59	1.00
Perceived stress I (In the past two weeks, how often have you...)		(1)	(2)	(3)	(4)	(5)	(6)			
(1)	been upset because of something that happened unexpectedly?	1.00								
(2)	felt "stressed"?	0.49	1.00							
(3)	felt nervous?	0.47	0.63	1.00						
(4)	found that you could not cope with all the things that you had to do?	0.28	0.34	0.32	1.00					
(5)	been angered because of things that were outside of your control?	0.44	0.45	0.44	0.36	1.00				
(6)	felt difficulties were piling up so high that you could not overcome them?	0.43	0.53	0.48	0.48	0.58	1.00			

Emotion regulation reappraisal		(1)	(2)	(3)	(4)	(5)	(6)			
Perceived stress II (In the past two weeks, how often have you...)		(1)	(2)	(3)	(4)					
(1)	felt confident about your ability to handle your personal problems?	1.00								
(2)	felt that things were going your way?	0.57	1.00							
(3)	been able to control irritations in your life?	0.61	0.50	1.00						
(4)	felt that you were on top of things?	0.61	0.53	0.79	1.00					
Classroom management		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
(1)	How much can you do to control disruptive behavior in the classroom?	1.00								
(2)	How much can you do to get children to follow classroom rules?	0.63	1.00							
(3)	How much can you do to calm a student who is disruptive or noisy?	0.60	0.70	1.00						
(4)	To what extent can you establish a classroom management system...	0.51	0.61	0.65	1.00					
(5)	To what extent can you keep a few problem students from ruining an entire lesson?	0.56	0.61	0.70	0.61	1.00				
(6)	To what extent can you respond to defiant students?	0.48	0.50	0.56	0.50	0.64	1.00			
(7)	To what extent can you make your expectation clear about student behavior?	0.57	0.58	0.56	0.58	0.55	0.62	1.00		
(8)	To what extent can you establish routines to keep activities running smoothly?	0.49	0.58	0.51	0.59	0.55	0.50	0.61	1.00	

Table B2: EFA All Scales–Statistics

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor 1	9.24	4.69	0.25	0.25
Factor 2	4.55	1.53	0.12	0.37
Factor 3	3.02	0.42	0.08	0.45
Factor 4	2.59	0.88	0.07	0.52
Factor 5	1.72	0.37	0.05	0.57
Factor 6	1.35	0.28	0.04	0.61
Factor 7	1.06	0.09	0.03	0.64

Figure B1: EFA All Scales–Scree Plot

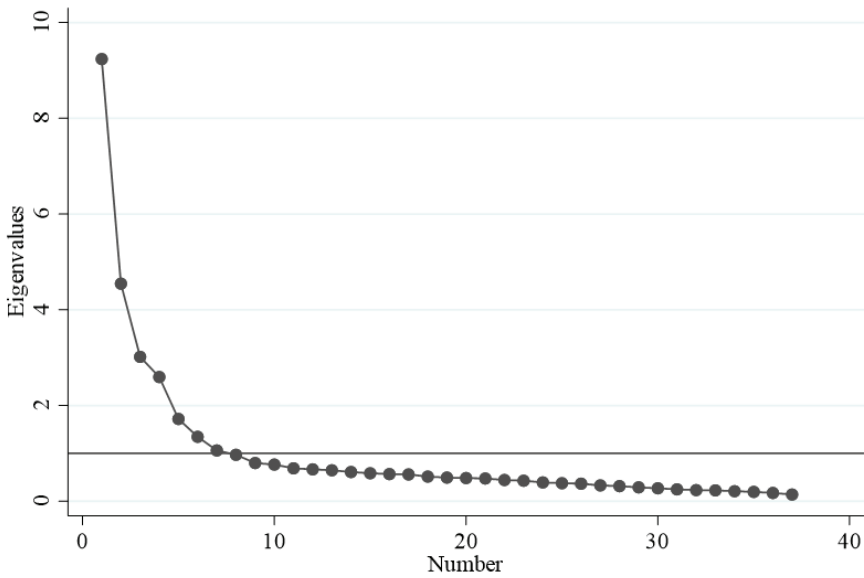


Table B3: Parallel Analysis

Real Data % of Variance	Mean of Random % of Variance	95 th Percentile of Random % of Variance
26.2106**	5.4262	5.8937
12.8515**	5.1485	5.5164
8.5394**	4.9475	5.265
7.3024**	4.7705	5.0868
4.8575*	4.6119	4.9044

Note: Parallel analysis implemented using FACTOR, following Lorenzo-Seva and Ferrando (2006), based on minimum rank factor analysis, as recommended by Timmerman and Lorenzo-Seva (2011). Polychoric correlation matrices used. Permutation of the raw data was performed to obtain random correlation matrices, as suggested by Buja and Eyuboglu (1992). **Advised number of dimensions when 95th percentile is considered. *Advised number of dimensions when mean is considered.

APPENDIX C

WHAT TOOL: SPANISH VERSION

Table C1: Spanish Version of the Scales

Item
Emotion regulation reappraisal
Si quiero tener una emoción más positiva a la que estoy sintiendo trato de pensar en algo mas/en otra cosa/de cambiar mi pensamiento
Si quiero tener una emoción menos negativa a la que estoy sintiendo trato de pensar en algo mas /en otra cosa/de cambiar mi pensamiento
Cuando me enfrento a una situación estresante, trato de pensar de tal forma que me ayude a estar en calma.
Cuando quiero sentir una emoción más positiva, modifico mi forma de pensar acerca de la situación actual.
Controlo mis emociones por medio de cambiar la forma en que pienso sobre la situación en la que me encuentro.
Cuando quiero sentir una emoción menos negativa, modifico mi forma de pensar sobre la situación en la que me encuentro.

Item
Emotion regulation suppression
Me guardo mis emociones para mí misma / mismo
Cuando siento emociones positivas me cuido de no expresarlas
No expresar mis emociones es una forma de controlarlas
Cuando estoy sintiendo emociones negativas, me aseguro de no expresarlas.
Perceived stress I (Durante las últimas (2) dos semanas, ¿Con cuánta frecuencia ha sentido estas emociones?)
Con molestia a causa de algo que ocurrió de forma repentina
Con estrés
Con nerviosismo
Llegar a la conclusión que no puede hacer frente con todas las cosas que tiene que hacer
Con enfado a causa de cosas que están fuera de su control
Sentir que las dificultades se acumulan de tal manera que no puede superarlas
Perceived stress II (Durante las últimas (2) dos semanas, ¿Con cuánta frecuencia ha sentido estas emociones?)
Con confianza acerca de su habilidad para lidiar / manejar sus problemas personales
La sensación que las cosas salen a su manera
En capacidad de controlar las irritaciones en su vida
En control de las cosas
Classroom management (¿En qué medida puede enfrentar de forma positiva las siguientes situaciones?)
Controlar el comportamiento que genera desorden dentro del aula
Lograr que las y los estudiantes sigan las reglas dentro del aula
Calmar estudiantes que generan ruido o alboroto
Establecer un sistema de manejo del aula con cada grupo de estudiantes
Evitar que algunas / algunos estudiantes problemáticos estropeen una clase completa
Lidiar con estudiantes desafiantes
Establecer de forma clara sus expectativas acerca del comportamiento que se espera de las y los estudiantes
Establecer rutinas que permitan que las actividades se desarrollen de forma ágil y continua

Note: Negative items were reverse scored.

EVALUATING THE 3CS PROGRAM FOR CAREGIVERS OF YOUNG CHILDREN AFFECTED BY THE ARMED CONFLICT IN COLOMBIA

LINA MARÍA GONZÁLEZ BALLESTEROS, JOSÉ M. FLORES,
ANA MARÍA ORTIZ HOYOS, AMALIA LONDOÑO TOBÓN, SASCHA HEIN,
FELIPE BOLÍVAR RINCON, OSCAR GÓMEZ,
AND LILIANA ANGÉLICA PONGUTA

ABSTRACT

Colombia has endured one of the world's longest internal displacement crises in recent history. Programs that address the practices and psychosocial wellbeing of the community of caregivers of young children in protracted crises are urgently needed. We developed and implemented a program aimed at strengthening the resilience and wellbeing of caregivers (parents, grandparents, and educators) of children enrolled in home-based and institutional centers for early childhood development in Colombia. The program, Conmigo, Contigo, Con Todos, or 3Cs, used purposive sampling across 14 municipalities disproportionately impacted by the armed conflict in Colombia. It consisted of two modules, a skills-building program (SBP) module and a psychotherapy intervention (PTI). The program content drew from cognitive behavioral therapy and mindfulness, and from inputs from local stakeholders. By applying a pragmatic evaluation strategy, we explored the pre-post intervention changes in parental resilience (the primary outcome of interest) through self-reports on the Connor-Davidson Resilience Scale (CD-RISC). The analysis of the pre-post intervention outcomes showed statistically significant improvements in CD-RISC in both intervention arms (SBP and PTI). Caregivers in the PTI group started with lower CD-RISC scores than caregivers who did not receive the PTI, and they showed the most improvement over time. Caregivers who had lower than average participation in the SBP (M=1-3 sessions out of a total of 6) did not

Received May 20, 2020; revised June 7, 2021, and July 15, 2021; accepted September 23, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

show significant changes in CD-RISC. Additionally, caregivers who had higher than average participation in the SBP showed significantly more improvement in CD-RISC scores than caregivers who did not attend any sessions. We discuss the implications of these findings for future applications of the program and substantiate the measurable impact of interventions for caregivers in conflict settings.

INTRODUCTION

The interplay between bioecological risks and protective factors during early childhood critically influences children's learning and developmental trajectories (Hein, Reich, and Grigorenko 2015; Wachs and Rahman 2013). Macro-level risks (e.g., disasters, conflict, and extreme poverty) are often juxtaposed with protective factors (e.g., peacebuilding strategies and education policies). As such, humanitarian crises and conflict can disrupt the ecology of human development at the macro, meso, and micro levels (Bronfenbrenner 2009). Here we describe the context and developmental underpinnings of the *Conmigo, Contigo, Con Todos* (With Me, With You, With All) or 3Cs program. The program targeted the community of caregivers (parents, grandparents, and teachers) of children enrolled in early childhood development (ECD) centers located in municipalities that have been disproportionately impacted by the armed conflict and internal displacement in Colombia. We introduce the intervention's context, with an emphasis on key macro-, meso-, and micro-level risks and protective factors in the target communities.

MACRO-LEVEL RISK AND PROTECTIVE FACTORS

Colombia experienced a 60-year civil conflict, which resulted in one of the largest internal displacement crises in recent history (UN High Commissioner for Refugees 2018). Extreme intergenerational poverty in the country has been concomitant and persistent. In 2020, the national monetary poverty level (a baseline acquisition power for food and goods) was 42.5 percent and the extreme monetary poverty level (a baseline acquisition power for basic foods) was 15.1 percent, with the incidence of the latter generally higher in peri-urban, rural, and disperse rural areas, which are farthest from urban centers, often nondelimited, and usually without access to basic public services (DANE 2021). Colombia also has had historically higher social inequality indices than other countries, both regionally and globally (Reliefweb 2020). The Colombian government has responded to these challenges by enacting several strategic geopolitical and social policies. From a peacebuilding perspective, the peace accords that ended the

government's armed conflict with the biggest guerilla group in the nation were a landmark effort (Gobierno de la República de Colombia 2016). The accords were built on several pillars, such as targeted investment in education, including early childhood education, health, and job opportunities, particularly in rural areas and localities severely affected by the war. Concomitantly, the government has been committed to the promotion of quality ECD through the implementation of the National ECD Law, *De Cero a Siempre*, or DCAS. Investments in quality ECD have been shown to be the most cost-effective social policies, and they are linked with pathways to social equality, inclusion, and the fomentation of a culture of peace (Rolnick and Grunewald 2003). Against this backdrop, DCAS aims to provide quality and equitable ECD services to all children while prioritizing those living in extreme poverty, and to provide holistic services through different contextualized modalities, including improving access to and the quality of early and primary education (Consejería Presidencial para la Niñez y la Adolescencia 2014).

MESO-LEVEL RISK AND PROTECTIVE FACTORS

Many families and children in Colombia were impacted by the war through massacres, attacks using explosive devices, forced recruitment into the armed forces, and community violence due to organized crime (Reliefweb 2020). Moreover, despite great strides forward in the implementation of DCAS, challenges to quality ECD access persist, in particular challenges to providing socioemotional and psychosocial support for young children and their caregivers (Gómez Cardona 2017). From a sociocultural perspective, community support and networks are crucial buffers against meso-level risks for families and children. Strong networks and community-based strategies can repair the social environment and renew trust within communities and toward institutions, which was disrupted by the conflict (Lozano Montilla, Parra Giraldo, and Uribe Ortiz 2019). For example, targeted programs to promote the wellbeing of children and their primary caregivers through ECD settings that stem from the peace accords have emerged as a significant social investment strategy (ICBF 2020).

Ecological approaches to children's development involve interactions among individuals, families, peers, and communities, which may increase or decrease the risk of negative outcomes in the face of adversity (Bronfenbrenner 2009). For example, the incidence and prevalence of psychopathologies resulting from exposure to war are associated with the degree of trauma experienced and the physical and emotional support available to a community (Murthy and Lakshminarayana 2006). Social support, broadly defined as material and interpersonal resources provided through social relationships, can deliver valuable resources in adverse contexts

(e.g., counseling, skills-building, information, or access to services), and it may act as a buffer to stress or provide direct benefits, despite the contextual stressors experienced by individuals (Thompson, Flood, and Goodvin 2006). In the context of child development, parental and community support are crucial during early childhood, primarily by fostering self-regulation, problem-solving, and other skills linked with positive developmental outcomes (Luthar, Crossman, and Small 2015). One potential byproduct of increased social support is increased social cohesion, which can manifest vertically (i.e., between individuals and groups and government institutions) and/or horizontally (i.e., in relationships between individuals, between individuals and groups, and between groups) (Pham and Vinck 2017). By fostering trust and improved relationships between individuals, institutions, and groups, strategic investment in effective ECD services has the potential to enhance social support and promote social cohesion (Leckman et al. 2019).

MICRO-LEVEL RISK AND PROTECTIVE FACTORS

PRACTICES OF THE COMMUNITY OF CAREGIVERS

Poverty, war, community violence, and barriers to early childhood services have directly affected the physical safety and security of many children in Colombia, as well as their psychosocial, emotional, and cognitive development (DCAS 2013). Nurturing care for young children is provided by an interconnected system of individuals inside and outside of the home, primarily mothers and fathers but also service providers, including early childhood educators (Britto et al. 2017). Supporting this notion are recent conceptual models that highlight the crucial role parents play as a buffer to the effects war can have on their children (Murphy et al. 2017). Recent literature reviews suggest that parenting programs in low- and middle-income countries have a measurable positive impact on children's cognitive and language development (Rao et al. 2014). A systematic review of 35 studies showed that young children exposed to war were at higher risk of developing posttraumatic stress disorder (PTSD) and posttraumatic stress symptoms, behavioral and emotional symptoms, sleep problems, and psychosomatic symptoms; however, these adverse effects were lower among young children who had higher-functioning parents and families (Slone and Mann 2016). A second review that explored the effects of war on children around the world found that the mental health effects appear to depend on the duration and acuity of the children's exposure to war. The worst outcomes have been observed among children who were the victim of or witnessed violent acts, had experienced threats to and the loss of loved ones, had experienced prolonged parental absence, and

were confronting forced displacement. Protective factors that mitigate the impact of war-related adversity on children include a strong bond between the primary caregiver and the child, and social support from teachers and peers (Werner 2012).

Grandparents are another critical element of family dynamics and structures around the globe (Sadrudin et al. 2019). Recent studies of interventions aimed at improving wellbeing in custodial families have called for strengthening programs by including grandparents, particularly those providing primary care for young children (Smith et al. 2018). Although data on the impact grandparental care has on the outcomes of young children are scarce, recent conceptual models call for research and practice agendas that consider the role grandparental care plays in children's physical health, and in their social-emotional, behavioral, cognitive, and educational development, particularly in contexts of high vulnerability (Sadrudin et al. 2019).

Current models of quality early childhood development and education go beyond family relations and are grounded in socioecological attachment and learning theories, which include process characteristics such as the interactions between educators and young children. Recent studies in Colombia have shown—for the first time at a national level—associations between positive and responsive interactions between teachers and children and the children's development outcomes (Maldonado-Carreño et al. 2018). The evidence points to the importance of considering parents, grandparents, and early childhood educators to be critical targets of programs that address the impact of adversity on children's learning and development. In early childhood education settings, the role teachers' mental health plays in their ability to support children's social-emotional learning has been well acknowledged, including prior studies showing associations between teacher depression and their negative relationships with children (Whitaker, Dearth-Wesley, and Gooze 2015). As such, programs should consider teacher wellbeing by lowering workplace stress and providing workplace support and training and targeted strategies that promote positive teacher-child interactions.

WELLBEING AND RESILIENCE IN THE COMMUNITY OF CARE

Conflict and poverty have multiple adverse effects on the wellbeing of caregivers that put at risk their ability to provide nurturing care for their children. This in turn poses a threat to children's positive cognitive, behavioral, and emotional adjustment (McEwen and McEwen 2017). Several studies addressing families living in adverse conditions have demonstrated that parenting programs can have a positive effect on a range of caregiver and/or child development outcomes (Annan et al. 2017; Dybdahl 2001; Ponguta et al. 2019). Recent studies also have

documented a broad range of positive wellbeing outcomes among caregivers in vulnerable contexts that result from targeted interventions (Hein et al. 2020). Recent global paradigms of nurturing care emphasize the importance of targeting caregivers' physical and mental health and overall wellbeing, while also enhancing their caregiving skills and strategies for helping their young children (Britto et al. 2017). More generally, global guidelines for providing mental health and psychosocial support in emergency settings provide frameworks that include multiple layers of support: basic services and security, community and family supports, focused nonspecialized supports, and specialized services (IASC 2007). There has been a particular increase in interventions to improve parenting practices, family relationships, and mental wellbeing for caregivers and children in low- and middle-income countries (Pedersen et al. 2019), which has led to a need for models that illustrate the operationalization of approaches in humanitarian settings and across children's community of care.

Key dimensions of caregiver wellbeing are resilience and resilience skills (Panter-Brick and Leckman 2013). Definitions of resilience vary across contexts and disciplines and are based on its characterization as a trait, a process, or an outcome (Ungar, Ghazinour, and Richter 2013). Resilience can constitute dynamic coping mechanisms, capacities, or resources that facilitate the successful endurance, recovery, and adaptability of individuals or groups of people who experience adversity that threatens their viability, ability to function, or development (Aburn, Gott, and Hoare 2016; Masten 2018). According to a multisystem resilience framework for disasters, resilience factors can be present simultaneously at the individual (child), family, and community (school or wider community) levels (Masten and Motti-Stefanidi 2020). Based on these observations, the resilience outcomes can be multifactoral. For example, multiple studies have shown a bidirectional relationship between cognitive and socioemotional development in the context of early life adversity (Osher et al. 2018). Resilience is specifically linked with later-life identity formation, which in turn impacts mental health and other individual developmental outcomes in adulthood (Smith and Pollak 2020). Furthermore, interventions that promote individual resilience have been shown to have a meaningful impact on limiting psychopathologies, such as depression, anxiety, and risk of suicide (Smith-Osborne, Maleku, and Morgan 2017; Zolkoski and Bullock 2012). Parental resilience can be defined as "the capacity of parents to deliver competent, quality parenting to children despite adverse personal, family, and social circumstances" (Gavidia-Payne et al. 2015, 111). A recent analysis of the socioecological factors that influence parenting behaviors suggests that parenting programs offer a promising approach to improving caregiving practices that help to promote children's resilience in the context of war (Murphy et al. 2017). However, systematic exploration of the evidence base suggests a deficit in reporting

on the design, implementation, and evaluation of resilience-focused interventions for caregivers in crisis contexts (Jordans et al. 2009). In that caregivers provide a critical buffer from the impact of conflict and other risks, it is necessary to explore resilience models that expand the focus from children's developmental trajectories to include caregiver resilience and the broader community context (Sim, Bowes, and Gardner 2019).

Global interventions in violence-affected settings have focused on working directly with children or promoting parenting skills. However, more recent interventions have also focused on improving the mental health and wellbeing of caregivers as a vehicle for improving child outcomes. Although improving the mental health and wellbeing of both caregivers and children can be seen as building resilience, few programs have examined resilience, and specifically caregiver resilience, as a main intervention outcome. Teachers are important people who provide care (i.e., attend to the personal needs of children from age 0 to 18) for several hours per day. Therefore, teachers are pivotal caregivers for young children right along with parents, grandparents, and other members of the child's family system. For example, a teacher-delivered protocol focused on enhancing personal resilience achieved significant improvement to stress, mood, and posttrauma symptoms among Israeli children exposed to the 2006 Lebanon War (Wolmer et al. 2011). However, there also are examples whereby psychosocial interventions aimed at increasing resilience among children exposed to war have shown null effects (Diab et al. 2015). Studies emerging from Colombia and the Latin American region that focus on resilience-building interventions in conflict settings are notably sparse. One example is a pilot of a school-based intervention aimed at fostering resilience among teachers and children. This model has shown positive effects on children's and teachers' self-esteem, humor, perseverance, assertiveness, and empathy (Acevedo and Restrepo 2012; Auyeung et al. 2012). Overall, a review of the literature suggests that there is a need to understand specifically how psychosocial interventions in violence-affected settings can affect caregiver resilience, and whether targeting caregiver resilience ultimately results in positive outcomes for children (Tol, Song, and Jordans 2013).

THE 3CS INTERVENTION AND THE PROCESS OF THE PRESENT STUDY

In 2015, a partnership between academia, the private sector, and the Colombian government led to the development and implementation of the 3Cs program. The program was designed to provide psychosocial support and resilience-building skills to caregivers (parents, grandparents, and teachers) of children enrolled

in ECD settings located in areas disproportionately impacted by the armed conflict, extreme poverty, and community violence. The program's theory of change drew from a peacebuilding-through-ECD paradigm, key social policy priorities in Colombia, the evidence base on resilience as a key protective factor in crisis contexts, and evidence from multiple psychotherapeutic interventions in conflict-affected areas (see Table A1 in the Appendix) (Yale University and AÇEV 2012). The aim of the present study was to develop, implement, and pragmatically evaluate the 3Cs program as a resilience-promotion intervention for caregivers of young children enrolled in ECD centers in Colombia. Since caregivers were the primary focus of the intervention, we hypothesized that participation in the 3Cs would be associated with improved parental resilience (the primary outcome of interest) when controlling for levels of parental psychopathology (i.e., symptoms of anxiety, depression, and PTSD). Below we discuss the implications of the study for the future application of psychosocial support and caregiver education in contexts affected by conflict and other risk factors.

METHODS

TARGET POPULATION

The 3Cs program was developed and implemented by an interdisciplinary team from *Fundación Saldarriaga Concha*, or FSC, a nongovernmental organization in Colombia. The program was funded by a leading child and family support government institute, the *Instituto Colombiano de Bienestar Familiar*, or ICBF. The ICBF is the leading publicly funded institution responsible for the provision of protection and ECD services for the most vulnerable children under the age of 18 in Colombia. Researchers from Yale University provided support for the formulation of the evaluation framework and execution of the data analysis. The program was implemented in 14 municipalities in Colombia.¹ These municipalities were selected because of their acute exposure to the armed conflict (e.g., direct presence of armed groups, geographic association with drug-trafficking routes) or because they were areas that hosted displaced rural communities. Participants were selected through a purposive sampling strategy, first from a list of ECD centers and community-based family homes provided by ICBF, and second, based on whether a person was a victim of armed conflict in accordance with the 1448 law, according to the ICBF register. If both applied, that person was invited to participate in the program.

1 The 14 municipalities were Medellín, Sincelejo, Pasto, Turbo, Soledad, Maicao, Buenaventura, Guapi, San Vicente del Caguán, Tame, Necoclí, Tumaco, Istmina, and El Tambo.

THEORY OF CHANGE

The program design was anchored in the intent to ameliorate macro- (effects of the armed conflict and extreme poverty), meso- (community violence and barriers to socioemotional and psychosocial support and education), and micro-level risk factors (maladaptive caregiver practices and poor caregiver wellbeing). Concomitantly, the program aimed to leverage macro- (peacebuilding and ECD policy landscape), meso- (targeted ECD services, community support and cohesion, interpersonal network), and micro-level (caregiver resilience) protective factors. The program content included several cognitive behavioral therapy and third-generation psychotherapeutic techniques shown to be effective among children and families in early childhood settings (Foa et al. 2009; Toth et al. 2002; Toth et al. 2006). The intervention design also included mindfulness techniques to address behavioral impulse control, impulsive regulation of stress, and emotional regulation, and to enhance the resilience of children recovering from traumatic events (Bethell et al. 2016). The content and approaches were selected by conducting a literature review and expert consultations. The program consisted of two overarching components. First, the skills-building program (SBP) module was offered to all caregivers (parents, grandparents, and teachers/educators working with children in the selected ECD settings). Second, the psychotherapy intervention (PTI) was offered to program participants who (1) self-reported to be “direct victims of the armed conflict” when enrolling children in the target ECD center, and (2) fulfilled the screening criteria described in the Group Assignment section below. Table A1 describes (1) the process applied in designing the program components, including the theoretical, conceptual, and stakeholder inputs and the process in which these inputs were integrated into the model; and (2) the overview of the content and implementation details of the SBP and the PTI.

The sessions were held once a week in community spaces (e.g., schoolrooms, community centers). The content and the approach of the PTI consisted of third-generation cognitive behavioral therapy techniques, namely, activation control therapy, behavioral activation technique, metacognitive therapy, mindfulness, schema-based therapies, and dialectical behavioral therapy. The key objective of the PTI was to bolster the learning of enduring bonding and to provide strategies to promote resilience, social skills, emotional processing, presentation techniques, activation control techniques, and self-control. The PTI included group-based discussions of the relevant concepts, as well as assignments and strategies to be implemented at home. Sessions were intended to be implemented only with parents and grandparents. However, due to other commitments or a lack

of childcare, some participants brought children to the sessions. When children were present, the facilitators were encouraged to normalize their presence and/or to ensure that one of the two facilitators provided focused support to the children brought to the sessions. If thematically relevant, the facilitators were encouraged to demonstrate activities by engaging with the children who were present.

GROUP ASSIGNMENT

We chose to conduct a pragmatic evaluation by combining qualitative and quantitative methods to explore the program's impact on parental resilience (Crane et al. 2019). A total of 2,448 consenting primary caregivers, including parents of children from 0 to 5 years old who were enrolled in ECD centers and homes, were invited to participate. They were screened for depression (Whooley depression screen), general anxiety (Hamilton-A), and PTSD (PTSD checklist-civilian version). This study pertains to the 331 of those 2,097 caregivers who were eligible to participate in the SBP and the PTI (see Figure A1 in the Appendix for a summary of group assignment). The PTI was delivered only to caregivers whose screening for PTSD, anxiety, and depression was negative. Those whose screening was positive were referred to the health services available in their municipality. While the SBP was intended for every caregiver, 40 caregivers (12.08%) did not participate in the SBP, and among those 40 caregivers, a small subset also did not participate in the PTI. Despite the small amount of cross-contamination due to the difficulties in access for real-time follow-up to the program's implementation, the majority of caregivers did in fact participate in the SBP, regardless of their assignment to the PTI ($N=291$ or 87.92%). The subset of caregivers ultimately selected for the PTI (whether or not they participated in the SBP) was $N=92$ of 331 caregivers (27.80%).

FACILITATOR AND DATA-COLLECTION TRAINING

The program facilitators were one psychologist and one social worker from each municipality. The facilitators were trained by members of the FSC in Bogotá. The trainers were a multidisciplinary team consisting of two psychiatrists, one psychologist, one general medicine practitioner, and one early childhood education expert. The training for program facilitators (or implementers) was held in Bogotá for five days in June 2015. The trainings included a combination of lectures and interactive and practice-oriented activities. Adjustments were made to the program content based on feedback from the facilitators and supervisors during the training. Teachers and educators were also trained in the SBP module for future implementation in ECD centers and community homes.

The data-collection teams were trained in Bogotá, and they used electronic tablets to collect demographic information, as well as primary caregiver outcomes. To obtain demographic data from the beneficiary primary caregivers and children, ICBF routinely collects a comprehensive set of variables. The data collectors were trained to transfer relevant data from the ICBF sociodemographic questionnaire to the tablets. Missing data from the ICBF's demographic questionnaire were imputed via direct interviews with participants. Data were uploaded from the tablets into a centralized RedCap data-management system. A sample of 10 percent of all data per municipality was verified by the lead project team. If there were errors or missing data, the data-collection teams in the municipalities were notified to review and amend accordingly. Data for the CD-RISC scores (the resilience measure used in this study) were entered on hard copies, then digitized by the lead project team at baseline and at the follow-up cross-sections of the program. Representatives of the project's lead team made site visits to all municipalities to oversee the onset of the program implementation and data collection. The supervisors continued to oversee the procedures throughout the implementation of the project.

MEASURES

DEMOGRAPHIC CHARACTERISTICS AND COVARIATES

Demographic variables included the age of caregivers (in years) and the caregivers' gender (male or female). Covariates of the program implementation included whether or not caregivers participated in the PTI, whether they participated in the SBP, and the average attendance at the SBP (0%, 17%, 33%, 50%, 67%, 83%, or 100%). The program facilitators tracked and entered attendance.

CD-RISC

This scale is comprised of 25 items designed to explore 5 factors: personal competence, tolerance and strength, positiveness, control, and spiritual influences. The original CD-RISC studies showed a high correlation between the scale and the measures of hardiness, perceived stress and stress vulnerability, disability, and social supports, which supports the convergent validity of the scale (Connor and Davidson 2003). A number of studies have focused on Spanish-speaking populations and validated different versions of the CD-RISC, and show that it is a reliable measure of resilience traits in Hispanic populations (Crespo, Fernández-Lansac, and Soberón 2014). In this study, we computed the total score as the sum of the 25-item and 10-item scales (Campbell-Sills and Stein 2007), respectively.

Several coauthors in this group are in the process of evaluating the validity and reliability of the CD-RISC scale for the population included in this study, which has not yet been presented in the literature and will be submitted for future publication. However, in this ongoing analysis, the internal consistencies of the 25-item scale and the 10-item scale were acceptable ($\alpha_{\text{CD-RISC } 25} = 88.35$; $\alpha_{\text{CD-RISC } 10} = 74.65$). In addition, the 10-item CD-RISC version had moderate to good validity indices, based on our initial assessments.

WHOOLEY DEPRESSION SCREEN

This is a two-question case-finding instrument for depression that asks about depressed moods and anhedonia. It has a sensitivity of 96 percent (95% CI=90-99%) and specificity of 57 percent (95% CI=53-62%) when a positive answer to any of the two items is given (Whooley et al. 1997). The Whooley questions are a recommended screening tool in the Colombian clinical practice guideline for depression, based on the operative characteristics stated above and a diagnostic odds ratio of 36.25 percent (95% CI=14.98-88.24%) (Ministerio de Salud y Protección Social 2013).

HAMILTON ANXIETY SCALE (HAM-A)

The HAM-A scale is a 14-item self-report measure developed as a scoring system for anxiety that has a good fit with clinical evaluation ($z=0.89$) (Hamilton 1959). Factor analysis showed a general factor clearly related to anxiety and a bipolar factor that grouped symptoms in psychic (i.e., mental agitation and psychological distress) and somatic (i.e., physical complaints related to anxiety) anxiety. Anxiety severity is rated as mild if scores are less than or equal to 17, mild to moderate if scores are between 18 and 24, moderate to severe if scores are between 25 and 30, and very severe for scores greater than 30 in a 0-56 score range (Hamilton 1959). The HAM-A has been validated in Spanish, with results showing psychometric properties similar to those of the original version (Cronbach's $\alpha=0.89$; intraclass correlation coefficient=0.92; effect size [sensitivity to change]=1.36) (Lobo et al. 2002). In this analysis, the internal consistency for HAM-A was 0.84.

PTSD CHECKLIST-CIVILIAN VERSION (PCL-C)

The PCL-C is a 17-item self-report measure of civilians' response to traumatic experiences (Wilkins, Lang, and Norman 2011). Total scores range from 17 to 85 and are based on the amount and severity of PTSD-related symptoms (symptoms severity range from 1=not at all to 5=extremely). Cutoff score for possible PTSD

is greater than or equal to 30 (sensitivity=82%; specificity=76%) (Walker et al. 2002). The PCL-C has been used to measure the response to behavioral cognitive interventions in Afro-descendant populations that are the victims of the armed conflict in Colombia (Bonilla-Escobar et al. 2018), to evaluate diagnostic criteria in mental health in victims of armed conflict in Colombia and Cambodia (Stammel et al. 2015), and to screen for PTSD in Colombia's 2015 National Mental Health Survey (Tamayo Martinez et al. 2016). The internal consistency for PCL-C was 0.87 in this analysis sample.

DATA ANALYSIS

Data management and statistical analysis were performed using STATA/IC v16 (Stata Corp). Continuous variables are presented as mean (standard deviation) or median (interquartile range). Categorical variables are presented as the number (proportion or percentage) of participants. Due to the nested nature of the CD-RISC scores measured before and after the intervention, random effects models were used to account for the covariance of scores among caregivers. Univariate linear regression with random intercepts estimated the association between CD-RISC scores and (1) the PTI and (2) the SBP (including SBP average attendance). While the primary outcome of interest was caregivers' CD-RISC 25-item scores, we also estimated associations with the 10-item version to evaluate whether magnitude and statistical significance differed from the 25-item scale. Multivariate linear mixed models estimated the independent effect of the PTI, the interaction of the PTI \times time of follow-up (follow-up vs. baseline scores), and the SBP average attendance. Multivariable models were adjusted for screening tools if they were statistically associated with the subgroups in Table 1 (Hamilton-A total score, PCL-C total score, and positive screening on the Whooley depression screener). All significant associations are reported at a threshold of $\alpha=0.05$.

RESEARCH ETHICS

All the program beneficiaries and study participants who enrolled signed an informed consent form that was administered by study personnel, per the regulations established by Colombia's ethics oversight committee and approved by the ICBF. The informed consent (and the application of all study instruments) was delivered by the psychologists trained by the FSC in Bogotá and subsequently deployed to the municipalities. To ensure application of the principle of do no harm, all of the 3Cs program facilitators were trained in the activation of a referral health system to provide specialized support as needed (e.g., in the presence of depression, PTSD, and/or anxiety). Furthermore, in partnership with ICBF,

families with special needs were given referrals to other family supports as needed. If personnel from the 3Cs program suspected child abuse or neglect, the study psychologists activated a referral to ICBF and the pertinent local entities.

RESULTS

Following our pragmatic evaluation strategy, we assessed the change in the CD-RISC score after the intervention. Relevant aspects of the program, such as the group that the caregivers were assigned to and the number of sessions attended, were assessed relative to the CD-RISC score. The results of the anxiety, depression, and PTSD screening tests are described and compared according to the group allocation. Table 1 shows caregiver characteristics at their baseline visit according to the 3Cs component (SBP vs. PTI; henceforth referred to as intervention groups) in which they participated. A total of 331 caregivers completed the CD-RISC before and after the intervention. Of those, 14 caregivers (4.23%) did not participate in the SBP or the PTI, 26 (7.85%) participated in the PTI but not the SBP, 225 (67.98%) participated in the SBP but not the PTI, while the remaining 66 caregivers (19.94%) participated in both the SBP and the PTI. Neither the caregivers' age nor the proportion of each gender differed significantly by intervention group. In terms of screening for anxiety using HAM-A scores, while there were significant baseline differences in the total HAM-A scores (mean differences, $p < 0.001$; median differences, $p < 0.001$; see Table 1), there were no differences across intervention groups when the scores were tabulated into severity categories using the cutoff scores ($p = 0.33$). We likewise observed significant differences in the total PCL-C scores at baseline (mean differences $p < 0.05$; median differences, $p < 0.008$), but there were no differences when the scores were tabulated into positive versus negative screening for PTSD, regardless of whether the 30- or 35-point cutoff score was used ($p = 0.21$ for the 30-point cutoff, $p = 0.68$ for the 35-point cutoff). In contrast, positive depression screening (+DS) at baseline using the Whooley two-item scale was significantly different across groups (+DS_{No SBT, No PTI} = 28.6%, +DS_{No SBT, PTI} = 15.4%, +DS_{SBT, No PTI} = 41.3%, +DS_{SBT & PTI} = 42.4%; $p < 0.007$). Average attendance (i.e., dose) at the SBP was 52.19 percent of the sessions ($SD = 35.19\%$). Average attendance at the SBP was not significantly different between the two participating subgroups. The average attendance at the SBP among those who participated only in the SBP (without the PTI) was 66.93 percent ($SD = 23.03\%$), compared to 71.35 percent ($SD = 23.84\%$) SBP attendance among caregivers who attended both the SBP and the PTI ($p = 0.19$). There were no significant differences in the distribution of attendance at the PTI between the group who only participated in the SBP and the group who participated in both the SBP and the PTI ($p = 0.30$).

Table 1: Baseline Characteristics of Colombian Caregivers Residing in Areas Affected by Armed Conflict

Characteristic	No SBP, No PTI	No SBP, PTI	SBP, No PTI	SBP & PTI	p-value
Intervention group sample size (<i>n</i>)	14	26	225	66	
Age, mean (SD)	29.1 (8.1)	30.0 (9.8)	30.7 (8.5)	30.0 (9.8)	0.88
Sex					
Male	1 (7.1%)	1 (3.8%)	7 (3.1%)	5 (7.6%)	0.55
Female	13 (92.9%)	24 (92.3%)	172 (76.4%)	54 (81.8%)	
HAM-A, total score, mean (SD)	7.4 (7.4)	6.2 (5.6)	11.1 (6.4)	9.2 (6.3)	<0.001
HAM-A, total score, median (IQR)	5.5 (2.0, 11.0)	4.5 (3.0, 7.0)	12.0 (6.0, 15.0)	7.0 (4.0, 14.5)	<0.001
Anxiety severity (based on HAM-A scores)					
Mild (scores<17)	13 (92.9%)	24 (92.3%)	160 (71.1%)	51 (77.3%)	0.33
Mild to moderate (scores 18-24)	0 (0.0%)	1 (3.8%)	18 (8.0%)	9 (13.6%)	
Moderate to severe (scores 25-30)	1 (7.1%)	1 (3.8%)	6 (2.7%)	0 (0.0%)	
PCL-civilian, total score, mean (SD)	23.4 (4.8)	21.1 (6.4)	26.4 (7.5)	24.4 (6.7)	0.012
PCL-civilian, total score, median (IQR)	22.0 (19.0, 27.0)	18.5 (17.5, 20.5)	26.0 (20.0, 31.0)	24.0 (17.0, 29.0)	0.007
PCL-civilian missing data	1 (7.1%)	10 (38.5%)	43 (19.1%)	5 (7.6%)	
PCL-civilian screening results based on cutoff value of 30					
Negative screening (scores<30)	11 (78.6%)	14 (53.8%)	129 (57.3%)	49 (74.2%)	0.21
Positive screening (scores≥30)	2 (14.3%)	2 (7.7%)	53 (23.6%)	12 (18.2%)	
PCL-civilian screening results based on cutoff value of 35					
Negative screening (scores<35)	13 (92.9%)	15 (57.7%)	165 (73.3%)	56 (84.8%)	0.68
Positive screening (scores≥35)	0 (0.0%)	1 (3.8%)	17 (7.6%)	5 (7.6%)	

EVALUATING THE 3CS PROGRAM FOR CAREGIVERS IN COLOMBIA

Characteristic	No SBP, No PTI	No SBP, PTI	SBP, No PTI	SBP & PTI	<i>p</i> -value
Whooley depression screen					
Negative screening	10 (71.4%)	22 (84.6%)	93 (41.3%)	34 (51.5%)	0.006
Positive screening*	4 (28.6%)	4 (15.4%)	93 (41.3%)	28 (42.4%)	
Missing data	0 (0.0%)	0 (0.0%)	39 (17.3%)	4 (6.1%)	
Skills-building program (SBP)					
No SBP	14 (100.0%)	26 (100.0%)	0 (0.0%)	0 (0.0%)	<0.001
SBP	0 (0.0%)	0 (0.0%)	225 (100.0%)	66 (100.0%)	

Note: IQR= interquartile range.

*Screening for depression was considered positive when both items reported in the Whooley Depression Screen were depression items (low mood/affect and anhedonia).

Bold values indicate p values below the alpha level of 0.05.

Table 2 shows unadjusted changes in CD-RISC scores, based on PTI/SBP grouping using paired t-tests. We tested differences using the 25-item and 10-item versions of the CD-RISC to evaluate the changes in magnitude and the statistical significance between both the longer and shorter versions of this survey. In terms of magnitude, Table 2 shows larger pre-post differences for the CD-RISC 25-item questionnaire than for the 10-item questionnaire. Consequently, the statistical significance of the t-test statistic is also considerably lower for the 10-item version.

Table 2: Descriptive Statistics and Unadjusted Paired T-Tests Showing Differences in CD-RISC Scores before and after Program Intervention

	No SBP, No PTI	No SBP, PTI	SBP, No PTI	SBP and PTI
10-item CD-RISC survey				
Baseline mean (SD)	28.8 (6.6)	22.1 (8.1)	28.2 (6.7)	26.1 (7.4)
Follow-up mean (SD)	27.4 (4.4)	26.8 (3.4)	29.7 (6.8)	28.7 (6.7)
Paired t-test <i>p</i> value:	0.4294	0.0150 *	0.0066**	0.0222*
25-item CD-RISC survey				
Baseline mean (SD)	73.4 (11.8)	56.2 (18.5)	70.8 (16.1)	66.8 (17.6)
Follow-up mean (SD)	70.6 (9.5)	71.9 (8.0)	75.7 (15.8)	73.3 (16.7)
Paired t-test <i>p</i> value:	0.3272	0.0008***	0.0002***	0.0125*

Note: This table compares the change in magnitude and statistical significance if measurements are conducted with 10-item versus 25-item versions of the CD-RISC.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The linear mixed effects model in Table 3 shows a main effect of time for caregivers who did not participate in the PTI (most of these caregivers did participate in the SBP). In the non-PTI group, the pre-post change shows an increase in resilience scores of +4.70 units, 95 percent CI=1.82 to 7.58, $p < 0.001$ (Figure 1).

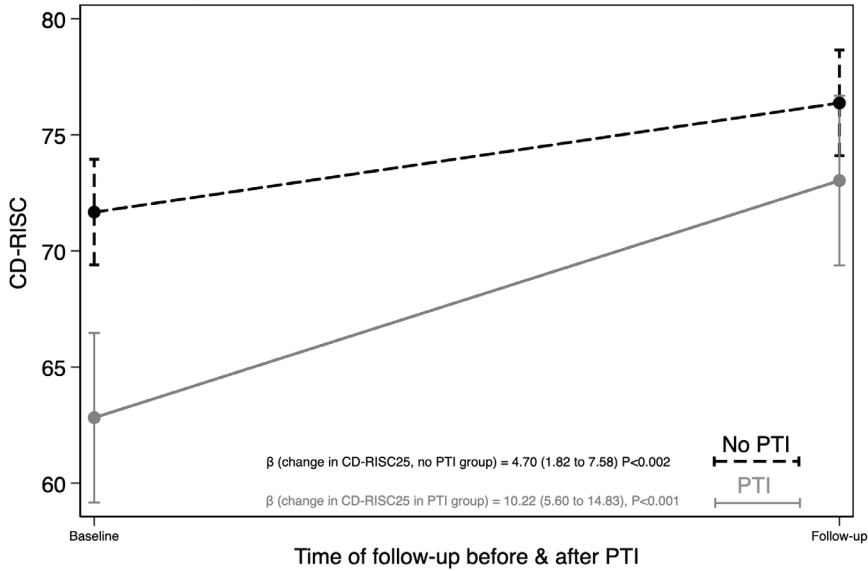
Table 3: Linear Random Effects Model Showing Significant Effects of Time of Follow-Up, PTI, the Time × PTI Interaction, as well as SBP Average Attendance

Change in CD-RISC 25 Total Score	β Coef.	[95% Confidence Interval]		p value
Follow-up vs. baseline in no PTI group	4.70	1.82	7.58	0.001 ***
PTI vs. no PTI at baseline	-8.05	-12.50	-3.61	0.001 ***
Time × psychotherapy interaction	5.51	0.07	10.95	0.047 *
Attendance at SBP				
17%	-1.73	-10.00	6.53	0.68
33%	-4.74	-12.16	2.67	0.21
50%	1.04	-5.31	7.40	0.74
67%	3.95	-1.72	9.62	0.17
83%	7.28	1.83	12.71	0.009 **
100%	2.32	-4.05	8.68	0.47
HAM-A total score (1-unit change)	0.25	-0.034	0.53	0.09
PCL-civilian total score (1-unit change)	-0.07	-0.33	0.19	0.60
Whooley positive screening vs. negative	2.95	-0.14	6.04	0.06

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The model also shows significant differences in CD-RISC at baseline, which suggests that the PTI group's caregivers started with significantly lower resilience scores at baseline than the SBP group (-8.05 units; 95% CI=-12 to -3). However, the interaction term of the PTI with time was statistically significant, which shows that, compared to the SBP group, the PTI group on average increased its CD-RISC scores significantly, despite have the lowest resilience scores at baseline.

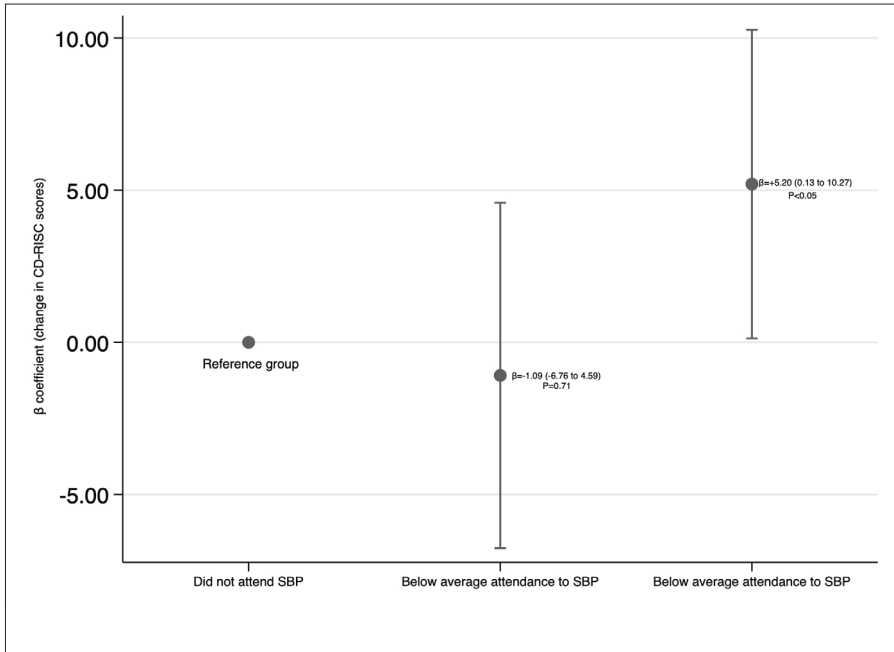
Figure 1: Changes in Pre-Post Intervention CD-RISC-25 Scores



Note: Figure 1 compares caregivers who (1) participated in the PTI and (2) caregivers who did not participate in the PTI (reference group). Adjusted for the independent effect of the PTI, the interaction of the PTI by time of follow-up (follow-up vs. baseline scores), attendance at the SBP, and for statistically significant screening tools (HAM-A total score, PCL-C total score, and Whooley positive depression screening).

We estimated whether attendance at the SBP was associated with significant differences in CD-RISC scores. As shown in Figure 2, we divided SBP attendance into three groups (no attendance, below average attendance, and above average attendance). Figure 2 shows that SBP attendance was associated with higher CD-RISC scores, but only if the participating caregivers had higher than average attendance (change=+5.20, $p<0.05$). Caregivers who participated in the SBP but did so with less than average attendance did not have significantly different scores than caregivers who did not attend the SBP at all.

Figure 2: Changes in CD-RISC-25 Scores Based on Participation



Note: Figure 2 indicates whether caregivers (1) did not attend the SBP program (reference group), (2) participated with below average attendance, or (3) participated with higher-than-average attendance. Adjusted for the independent effect of the PTI, the interaction of PTI \times time of follow-up (follow-up vs. baseline scores), and for statistically significant screening tools (HAM-A total score, PCL-C total score, and Whooley positive depression screening).

DISCUSSION

Our study describes the development and evaluation of the 3Cs program, a resilience and wellbeing promotion intervention for caregivers of young children enrolled in ECD centers in Colombia. The program targeted municipalities acutely affected by the country's armed conflict and by forced displacement. To our knowledge, this is one of the first studies to assess the impact of a program on parental resilience in crisis contexts in a Latin American country. This intervention also combined multiple inputs in its design, such as several psychosocial intervention approaches, the application of community-participatory research principles, and the utilization of ECD settings as an entry point for implementation. Supporting our hypothesis, the results of this study show statistically significant improvements in parental resilience (CD-RISC scores) as a result of participating in both program modules (the SBP and the PTI). Importantly, while participants in the PTI group started with lower resilience scores than the group who did not participate in the PTI, they

showed the most improvement after the intervention. One possible interpretation of these findings is that caregivers who rated their initial levels of resilience as relatively low compared to other caregivers are also the most likely to benefit from the PTI. Another interpretation of this finding could be social desirability and regression, in that caregivers with low baseline levels of resilience reported higher levels post-intervention because they felt that improved resilience was expected of them. More research is needed to determine the association and conclusions reported here. Future studies are needed in particular to determine the relationship between attendance, participant characteristics, and their impact on wellbeing outcomes and resilience. More research is also needed to determine and mitigate the reasons for not attending and/or dropping out.

While participation in the SBP was designed for all caregivers, 40 caregivers did not participate in those sessions. While finding a group of caregivers who did not participate in the SBP or the PTI was not the original intention of this intervention, the scale and complicated nature of emergency and fragile settings resulted in a small number of people enrolling in the study but not attending either the SBP or the PTI. We made use of this natural experimental (pseudo-control) group to evaluate the effects of average attendance at the SBP. Caregivers who had lower than average participation (mean attendance at the SBP=52.19% of sessions) had resilience scores similar to the group who did not attend any sessions. In contrast, caregivers who had higher than average participation in the SBP showed significantly more improvement in their resilience scores than the group who did not attend any sessions. Therefore, the SBP program's benefits seem to have a threshold beyond which caregivers living in fragile contexts benefit, and below which caregivers have resilience scores comparable to the general caregiver population in similar circumstances. The findings of this pragmatic evaluation are important in informing the design and implementation of controlled randomized interventions. They have also been used to design program models in communities with a similar background and in the context of the education system in Colombia, directed toward vulnerable communities who possibly benefit from interventions that aim to promote the development, strengthening, and maintenance of resilience. The findings of this study also demonstrate that programmatic approaches that target meso-level risk and protective factors (e.g., targeted ECD services, community networks, and psychosocial support) have the potential to promote caregivers' micro-level outcomes (e.g., resilience), which presumably has spillover effects on other caregiver outcomes, such as wellbeing and psychopathology.

We conducted focus group discussions and in-depth interviews with caregivers who were randomly selected. Our preliminary analyses following these discussions indicated that the resilience promotion program may have led to a reduction in the physical and verbal punishment of children, increased recognition of children's emotions, enhanced parental empathy toward their children, increased compassion for others, and a recognition of self-resilience among caregivers.² These possible effects need to be explored further, including a comprehensive qualitative analysis of the data to validate the findings. Nevertheless, there is evidence that the practice of mindfulness can enhance neuroplasticity and functional changes in the brain regions involved in the regulation of attention, emotions, and self-awareness (Tang, Hölzel, and Posner 2015). An exploration of the program's impact on neurobiological markers of stress and other markers of wellbeing could inform the intervention's mechanisms of action.

STRENGTHS AND INNOVATION

The program has several innovative attributes that contribute to the current evidence base. First, the program combines multiple approaches to mental health support, including a behavioral-cognitive model that emphasizes emotional regulation techniques (e.g., breathing), problem-solving, self-control, and social abilities (e.g., assertiveness and empathy). Second, the development of the 3Cs program included a qualitative exploration of the perspectives of parents, teachers, local health secretariats, and ICBF officials on resilience-building topics (e.g., strategies to facilitate conflict resolution, spirituality, and the role of grandparents), which were incorporated into the program sessions. Although spirituality was not an explicit component of the program, caregivers highlighted it as a key tool for strengthening resilience. This is consistent with the existing literature on the importance of spirituality in other contexts (e.g., among a sample of executives in the United States) when used as a mechanism to confront difficult situations, solve problems, and recover the meaning and purpose of life (Shelton, Hein, and Phipps 2019; Smith et al. 2012).

Third, given that the community of care for young children in many of the sites was diverse and often intergenerational, grandparents were eligible to participate. Despite the fact that grandparents around the globe often provide care to children, they are generally overlooked in the design of caregiver programs (Sadruddin et al. 2019). Fourth, the program is one of few that, to our knowledge, has been implemented in ECD settings in war-affected and other highly vulnerable

² Possible effects are shown in unpublished ICBF and FSC data.

contexts, such as extreme poverty and insecurity. Interestingly, recent studies have shown that developing positive parenting skills is related to greater childhood resilience and family resources when facing displacement (Domínguez de la Ossa 2018). In the case of Colombia, working through ECD settings enabled the rapid identification of highly at-risk communities (e.g., a high poverty index, victims of the armed conflict, family violence and abuse) because these populations are prioritized for enrollment in publicly funded ECD centers. Working in ECD settings also afforded the possibility of including parents and other primary caregivers (e.g., grandparents, extended family) so that the 3Cs program reached the wider community of care. The program welcomed the participation of male and female caregivers, which is particularly relevant, as recent studies have shown positive (indirect) associations between paternal engagement and maternal distress, harsh parenting, and parenting stress (Hein et al. 2020).

CHALLENGES AND ENABLERS OF IMPLEMENTATION

The program's implementation and evaluation were enabled through the Colombian government's prioritization of ECD, mental health, and social strategies for peacebuilding as mechanisms for national development. The program content was aligned with the vision of multiple legislative frameworks, including DCAS and the implementation of the peace accords. A key justification for developing the program was the evidence brought forth by the peacebuilding through ECD paradigm (Yale University and AÇEV 2012) and its alignment with Colombia's policy priority to invest in strategies to bolster socioemotional skills and conflict resolution that included young children, caregivers, and parents. The government buy-in facilitated the program implementers' ability to engage with the communities and local leadership (e.g., community leaders, *juntas*, churches, cultural centers, early childhood education center directors and teachers). These relationships were key to building trust in the communities and aligning the program's content with a culturally diverse group of municipalities. The mechanisms of community engagement featured several communication strategies (e.g., community radio, flyers) that were used effectively to raise awareness of the program and encourage participation.

To ensure that the fidelity and quality of the program implementation was sustained, an intensive and structured training program for the facilitators was delivered by the FSC. Crucial on-site support and mentorship were also provided regularly throughout the process for facilitators at all sites. In some locations where internet access was available, the use of online social networks was an effective way to streamline referrals to other services and to offer additional

support for program beneficiaries. It is important to note that the program involved training psychologists and mental health professionals who lived in the target municipalities. As a result, the skills and capacities that the participants acquired through the program may have been introduced in a sustainable way in the communities, partly overcoming the shortage of skilled people, a difficulty prevalent in disperse rural parts of Colombia.

One main challenge in implementing the program was to sustain enrollment and attendance. Parents' program attendance in humanitarian crisis settings has been reported elsewhere to be one of the main barriers to implementation and is associated with program effects (Ponguta et al. 2020; Ponguta et al. 2019). To encourage attendance at both program modules, participants were offered snacks as an added incentive and the session scheduling was conducted in close consultation with the beneficiaries. These incentive strategies were aligned with what was thought to be acceptable to the context and the local partners.

Conducting controlled evaluations of psychosocial interventions in conflict-affected contexts is known to be challenging (Hein and Weeland 2019), and this case was no exception. Training and deploying reliable data collectors required funding earmarked for the program evaluation, and the partnership between the academic, public, and private sectors was a key enabler to the data collection and analysis. However, because of a lack of internet connectivity in several of the municipalities, real-time data entry was not possible. This delayed the analyses and challenged the quality control of data management. Furthermore, from an evaluation design perspective, it was necessary to conduct an observational evaluation because of ethical concerns in the assignment of control arms to ensure that the members of the communities involved in the 3Cs program had access to mental health promotion strategies likely to be beneficial for them. Conducting randomized controlled trials would strengthen the evidentiary base for this and similar programs.

STUDY LIMITATIONS AND FUTURE DIRECTIONS

This study focused on the impact of a program to promote parental resilience among a subgroup of caregivers who participated in the SBP, a combination of the SBP and the PTI, or the PTI alone. More research is needed to determine the impact of the 3Cs program components on caregiver outcomes, namely, their practices and wellbeing, and on all outcomes for teachers and grandparents. Furthermore, evaluations are needed to establish the program's impact on

vertical (e.g., trust in institutions and institutional capacities) and horizontal (e.g., trust across families and communities) social cohesion as a result of program participation. In fact, a key issue ECD programs face is assessing their potential to build cohesion and establish a pathway to intergenerational peace (Connolly, Hayden, and Levin 2007; Leckman, Panter-Brick, and Salah 2014). Importantly, future evaluations should explore the effects of enhanced caregiver resilience, wellbeing, and social cohesion on children's short-term (e.g., socioemotional and cognitive development and strengthened resilience) and long-term (e.g., reduced risk of psychopathologies and improved developmental outcomes later in life) outcomes.

Limited funding for the evaluation of programs of this nature is a persistent challenge. We were only able to apply a self-reported measure of parental resilience to explore the impact of the 3Cs program and apply a pragmatic evaluation. To advance the field, future studies should apply observational and behavioral measures beyond self-reports that include all caregivers and children, and that follow the effects over time. Our study was not able to empirically determine the elements of the program modules that contributed to its impact. We also were unable to assess the program's impact on children's outcomes, primarily due to resource constraints. However, focus group discussions with facilitators after the program implementation suggest that sessions that addressed problem-solving skills and offered concrete techniques to develop self-control and self-regulation were well received, and they seemed to be integrated more easily into participants' behavior changes. More research is needed to validate these preliminary observations and make determinations about which programmatic elements and attendance could be linked to positive outcomes. Importantly, in part due to limited funding, we were not able to follow up with participants after the post-intervention assessment. Ideally, determining the sustainability of the program's impact should include a one-year follow-up assessment. Our study was challenged by various issues, such as higher attrition rates than those reported in similar studies conducted in more controlled settings. The challenges our team encountered in following up with participants are common for studies performed in real-world settings, especially conflict and postconflict settings. We argue nonetheless that the pragmatic nature of our research design counters this weakness and increases the external validity of our findings.

From a systems perspective, it is crucial to identify national- and municipal-level entry points to ensure that programs addressing young children's and caregivers' wellbeing are included in development and investment plans. By aligning this program to Colombia's National Law for Early Childhood, for example, the

initiative was anchored in strategic actions to provide holistic support for young children and their families. The program to promote resilience was one of the strategic actions included in Colombia's National Policy for Mental Health 2019 and is one of the programs adopted to promote mental health in the country. Integrating these programs into the country's public policy vision are key to their scalability and sustainability.

REFERENCES

- Aburn, Gemma, Merryn Gott, and Karen Hoare. 2016. "What Is Resilience? An Integrative Review of the Empirical Literature." *Journal of Advanced Nursing* 72 (5): 980-1000. <https://doi.org/10.1111/jan.12888>.
- Acevedo, Victoria Eugenia, and Lucía Restrepo. 2012. "De Profesores, Familias y Estudiantes: Fortalecimiento de la Resiliencia en la Escuela." *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud* 10: 301-319. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1692-715X2012000100019&nrm=iso.
- Annan, Jeannie, Amanda Sim, Eve S. Puffer, Carmel Salhi, and Theresa S. Betancourt. 2017. "Improving Mental Health Outcomes of Burmese Migrant and Displaced Children in Thailand: A Community-Based Randomized Controlled Trial of a Parenting and Family Skills Intervention." *Prevention Science* 18 (7): 793-803. <https://doi.org/10.1007/s11121-016-0728-2>.
- Auyeung, Bonnie, Carrie Allison, Sally Wheelwright, and Simos Baron-Cohen. 2012. "Brief Report: Development of the Adolescent Empathy and Systemizing Quotients." *Journal of Autism and Developmental Disorders* 42 (10): 2225-35. <https://doi.org/10.1007/s10803-012-1454-7>.
- Bethell, Christina, Narangerel Gombojav, Michele Solloway, and Lawrence Wissow. 2016. "Adverse Childhood Experiences, Resilience and Mindfulness-Based Approaches: Common Denominator Issues for Children with Emotional, Mental, or Behavioral Problems." *Child and Adolescent Psychiatric Clinics of North America* 25 (2): 139-56. <https://doi.org/10.1016/j.chc.2015.12.001>.

- Bonilla-Escobar, Francisco J., Andrés Fandino-Losada, Diana M. Martínez-Buitrago, Julián Santaella-Tenorio, Daniel Tobon-García, Edgar J. Muñoz-Morales, Iván D. Escobar-Roldán et al. 2018. "A Randomized Controlled Trial of a Transdiagnostic Cognitive-Behavioral Intervention for Afro-Descendants' Survivors of Systemic Violence in Colombia." *PLoS ONE* 13 (12): e0208483. <https://doi.org/10.1371/journal.pone.0208483>.
- Britto, Pia R., Stephen J. Lye, Kerrie Proulx, Aisha K. Yousafzai, Stephen G. Matthews, Tyler Vaivada, Rafael Pérez-Escamilla et al. 2017. "Nurturing Care: Promoting Early Childhood Development." *Lancet* 389 (10064): 91-102. [https://doi.org/10.1016/S0140-6736\(16\)31390-3](https://doi.org/10.1016/S0140-6736(16)31390-3).
- Bronfenbrenner, Urie. 2009. *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.
- Campbell-Sills, Laura, and Murray B. Stein. 2007. "Psychometric Analysis and Refinement of the Connor-Davidson Resilience Scale (CD-RISC): Validation of a 10-Item Measure of Resilience." *Journal of Traumatic Stress* 20 (6): 1019-28. <https://doi.org/10.1002/jts.20271>.
- Connolly, Paul, Jacqueline Hayden, and Diane Levin. 2007. *From Conflict to Peace Building: The Power of Early Childhood Initiatives - Lessons from Around the World*. Lincoln, NE: Exchange Press.
- Connor, Kathryn M., and Jonathan R. T. Davidson. 2003. "Development of a New Resilience Scale: The Connor-Davidson Resilience Scale (CD-RISC)." *Depression and Anxiety* 18 (2): 76-82. <https://doi.org/10.1002/da.10113>.
- Consejería Presidencial para la Niñez y la Adolescencia. 2014. *de Cero a Siempre*. Bogotá: Presidencia de la República de Colombia. <http://www.deceroasiempre.gov.co/QuienesSomos/Paginas/QuienesSomos.aspx>.
- Crane, Melanie, Adrian Bauman, Beverley Lloyd, Bronwyn McGill, Chris Rissel, and Anne Grunseit. 2019. "Applying Pragmatic Approaches to Complex Program Evaluation: A Case Study of Implementation of the New South Wales Get Healthy at Work Program." *Health Promotion Journal of Australia* 30 (3): 422-32. <https://doi.org/10.1002/hpja.239>.

- Crespo, Maria, Violeta Fernández-Lansac, and Carmen Soberón. 2014. "Adaptación Española de La 'Escala de Resiliencia de Connor-Davidson' (CD-RISC) En Situaciones de Estrés Crónico." *Behavioral Psychology* 22: 219-38. <https://doi.org/10.1037/t62957-000>.
- DANE (Departamento Administrativo Nacional de Estadística). 2021. *Pobreza y Desigualdad*. Bogotá: DANE. <https://www.dane.gov.co/index.php/estadisticas-por-tema/pobreza-y-condiciones-de-vida/pobreza-monetaria>.
- DCAS (De Cero a Siempre). 2013. *Estrategia de Atención Integral a la Primera Infancia: Fundamentos Políticos, Técnicos y de Gestión*. Bogotá: ICBF. <http://www.deceroasiempre.gov.co/QuienesSomos/Documents/Fundamentos-politicos-tecnicos-gestion-de-cero-a-siempre.pdf>.
- Diab, Marwan, Kirsi Peltonen, Samir R. Qouta, Esa Palosaari, and Raija-Leena Punamäki. 2015. "Effectiveness of Psychosocial Intervention Enhancing Resilience among War-Affected Children and the Moderating Role of Family Factors." *Child Abuse & Neglect* 40: 24-35. <https://doi.org/10.1016/j.chiabu.2014.12.002>.
- Domínguez de la Ossa, Elsy Mercedes. 2018. "Los Recursos Familiares para Afrontar la Adversidad del Desplazamiento Forzado." *Psicología desde el Caribe* 35: 145-55. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-417X2018000200145&nrm=iso.
- Dybdahl, Ragnhild. 2001. "Children and Mothers in War: An Outcome Study of a Psychosocial Intervention Program." *Child Development* 72 (4): 1214-30. <https://doi.org/10.1111/1467-8624.00343>.
- Foa, Edna B., Terence M. Keane, Matthew J. Friedman, and Judith A. Cohen. 2009. *Effective Treatments for PTSD: Practice Guidelines from the International Society for Traumatic Stress Studies* (2nd ed.). New York: The Guilford Press.
- Gavidia-Payne, Susana, Bianca Denny, Kate Davis, Andrew Francis, and Merv Jackson. 2020. "Parental Resilience: A Neglected Construct in Resilience Research." *Clinical Psychologist* 19 (3): 111-21. <https://doi.org/10.1111/cp.12053>.
- Gobierno de la República de Colombia. 2016. *ABC del Acuerdo Final*. Bogotá: Gobierno de la República de Colombia. <https://www.cancilleria.gov.co/sites/default/files/cartillaabcedelacuerdofinal2.pdf>.

- Gómez Cardona, Liliana María. 2017. "Primera Infancia y Educación Emocional." *Revista Virtual Universidad Católica del Norte* 52: 174-84. (52): 174-184.
- Hamilton, Max. 1959. "The Assessment of Anxiety States by Rating." *British Journal of Medical Psychology* 32 (1): 50-55. <https://doi.org/10.1111/j.2044-8341.1959.tb00467.x>.
- Hein, Sascha, Jodi Reich, and Elena Grigorenko. 2015. "Cultural Manifestation of Intelligence in Formal and Informal Learning Environments During Childhood." In *The Oxford Handbook of Human Development and Culture: An Interdisciplinary Perspective*, edited by Lene Arnett Jensen, 214-29. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199948550.013.14>.
- Hein, Sascha, Johanna Bick, Ghassan Issa, Lara Aoude, Cosette Maalouf, Abir Awar, Sawsan Nourallah et al. 2020. "Maternal Perceptions of Father Involvement among Refugee and Disadvantaged Families in Beirut, Lebanon." *PLoS ONE* 15 (3): e0229670. <https://doi.org/10.1371/journal.pone.0229670>.
- Hein, Sascha, and Joyce Weeland. 2019. "Randomized Controlled Trials (RCTs) in Clinical and Community Settings: Challenges, Alternatives, and Supplementary Designs." *New Directions for Child and Adolescent Development* (167): 7-15. <https://doi.org/10.1002/cad.20312>.
- IASC (Inter-Agency Standing Committee). 2007. *IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings*. Geneva: IASC. https://interagencystandingcommittee.org/system/files/legacy_files/guidelines_iasc_mental_health_psychosocial_june_2007.pdf.
- ICBF (Instituto Colombiano de Bienestar Familiar). 2020. *Acciones Desarrolladas por el ICBF que Aportan a la Implementación del Acuerdo Final para la Terminación del Conflicto y la Construcción de una Paz Estable y Duradera: Vigencia 2018 y 2019*. Bogotá: ICBF. https://www.icbf.gov.co/sites/default/files/acciones_desarrolladas_por_el_icbf_que_aportan_a_la_implementacion_del_acuerdo_final_para_la_terminacion_del_conflicto_y_la_construccion_de_una_paz_estable_y_duradera_-_vigencia_2018_y_2019.pdf.

- Jordans, Mark J. D., Wietse A. Tol, Ivan H. Komproe, and Joop V. T. M. de Jong. 2009. "Systematic Review of Evidence and Treatment Approaches: Psychosocial and Mental Health Care for Children in War." *Child and Adolescent Mental Health* 14 (1): 2-14. <https://doi.org/10.1111/j.1475-3588.2008.00515.x>.
- Leckman, James F., Catherine Panter-Brick, and Rima Salah. 2014. *Pathways to Peace: The Transformative Power of Children and Families*. Cambridge, MA: MIT Press.
- Leckman, James F., Chelsea K. Donaldson, Friedrich W. Affolter, and Liliana Angélica Ponguta. 2019. "Pathways to Wellbeing and a More Peaceful and Sustainable World: The Transformative Power of Children and Families." *Japanese Journal of Child and Adolescent Psychiatry* 60 (3): 278-98. https://doi.org/10.20615/jscap.60.3_278.
- Lobo, Antonio, Lorenzo Chamorro, Antonio Luque, Rafael Dal-Re, Xavier Badia, and Eva Baro. 2002. "Validación de las Versiones en Español de la Montgomery-Asberg Depression Rating Scale y la Hamilton Anxiety Rating Scale para la Evaluación de la Depresión y de la Ansiedad." *Medicina Clínica* 118 (13): 493-99. [https://doi.org/10.1016/s0025-7753\(02\)72429-9](https://doi.org/10.1016/s0025-7753(02)72429-9).
- Lozano Montilla, Dahyana Andrea, Mateo Parra Giraldo, and Daniela Victoria Uribe Ortiz. 2019. "Resiliencia y Constructos Personales de Victimización en Individuos Vulnerados por el Conflicto Armado." *Poiésis* (37): 65-84. <https://doi.org/10.21501/16920945.3339>.
- Luthar, Suniya S., Elizabeth J. Crossman, and Phillip J. Small. 2015. "Resilience and Adversity." In *Handbook of Child Psychology and Developmental Science, Vol. 3*, edited by Richard M. Lerner, 1-40. Hoboken, NJ: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy307>.
- Maldonado-Carreño, Carolina, Carolina Rey Guerra, Jorge Cuartas, Tatiana Plata-Caviedes, Jesús J. Rodríguez De Luque, Eduardo Escallón Largacha, and Hirokazu Yoshikawa. 2018. *Quality Measurement of the Institutional Modality of Early Childhood Education in Colombia*. Bogotá: Universidad de Los Andes. <https://educacion.uniandes.edu.co/sites/default/files/Notes-In-Education.pdf>.
- Masten, Ann. S. 2018. "Resilience Theory and Research on Children and Families: Past, Present, and Promise." *Journal of Family Theory & Review* 10 (1): 12-31. <https://doi.org/10.1111/jftr.12255>

- Masten, Ann S., and Frosso Motti-Stefanidi. 2020. "Multisystem Resilience for Children and Youth in Disaster: Reflections in the Context of COVID-19." *Adversity and Resilience Science* 1 (2): 1-12. <https://doi.org/10.1007/s42844-020-00010-w>.
- McEwen, Craig A., and Bruce S. McEwen. 2017. "Social Structure, Adversity, Toxic Stress, and Intergenerational Poverty: An Early Childhood Model." *Annual Review of Sociology* 43 (1): 445-72. <https://doi.org/10.1146/annurev-soc-060116-053252>.
- Ministerio de Salud y Protección Social. 2013. *Detección Temprana y Diagnóstico del Episodio Depresivo y Trastorno Depresivo Recurrente en Adultos: Atención Integral de los Adultos con Diagnóstico de Episodio Depresivo o Trastorno Depresivo Recurrente*. Bogotá: Ministerio de Salud y Protección Social. http://gpc.minsalud.gov.co/gpc_sites/Repositorio/Conv_500/GPC_td/GPC_Comple_Depre.pdf.
- Murphy, Katie Maeve, Katherine Rodrigues, Jaime Costigan, and Jeannie Annan. 2017. "Raising Children in Conflict: An Integrative Model of Parenting in War." *Peace and Conflict: Journal of Peace Psychology* 23 (1): 46-57. <https://doi.org/10.1037/pac0000195>.
- Murthy, R. Srinivasa, and Rashmi Lakshminarayana. 2006. "Mental Health Consequences of War: A Brief Review of Research Findings." *World Psychiatry* 5 (1): 25-30. <https://www.ncbi.nlm.nih.gov/pubmed/16757987>.
- Osher, David, Pamela Cantor, Juliette Berg, Lily Steyer, and Todd Rose. 2018. "Drivers of Human Development: How Relationships and Context Shape Learning and Development." *Applied Developmental Science* 24 (1): 6-36. <https://doi.org/10.1080/10888691.2017.1398650>.
- Panter-Brick, Catherine, and James F. Leckman. 2013. "Resilience in Child Development: Interconnected Pathways to Wellbeing." *Journal of Child Psychology and Psychiatry* 54 (4): 333-36. <https://doi.org/10.1111/jcpp.12057>.
- Pedersen, Gloria A., Eva Smallegange, April Coetzee, Kim Hartog, Jasmine Turner, Mark J. D. Jordans, and Felicity L. Brown. 2019. "A Systematic Review of the Evidence for Family and Parenting Interventions in Low- and Middle-Income Countries: Child and Youth Mental Health Outcomes." *Journal of Child and Family Studies* 28 (8): 2036-55. <https://doi.org/10.1007/s10826-019-01399-4>.

- Pham, Phuong, and Patrick Vinck. 2017. *Indicators Framework for Peacebuilding, Education and Social Cohesion*. Boston, MA: UNICEF and Harvard Humanitarian Initiative. https://inee.org/system/files/resources/Social_Cohesion_Assessmen_Framework_20170420_%28HHI%29.pdf.
- Ponguta, Liliana A., Ghassan Issa, Lara Aoudeh, Cosette Maalouf, Sasha D. Hein, Anna L. Zonderman, Liliya Katsovich et al. 2020. "Effects of the Mother-Child Education Program on Parenting Stress and Disciplinary Practices Among Refugee and Other Marginalized Communities in Lebanon: A Pilot Randomized Controlled Trial." *Journal of the American Academy of Child & Adolescent Psychiatry* 59 (6): 727-38. <https://doi.org/10.1016/j.jaac.2019.12.010>.
- Ponguta, Liliana Angélica, Ghassan Issa, Lara Aoudeh, Cosette Maalouf, Sawsan Nourallah, Kaveh Khoshnood, Anna Leslie Zonderman et al. 2019. "Implementation Evaluation of the Mother-Child Education Program Among Refugee and Other Vulnerable Communities in Lebanon." *New Directions for Child and Adolescent Development* 2019 (167): 91-116. <https://doi.org/10.1002/cad.20314>.
- Rao, Nirmala, Jin Sun, Jessie M. S. Wong, Brendan Weekes, Patrick Ip, Sheldon Shaeffer, Mary Young, Mark Bray, Eva Chen, and Diana Lee. 2014. *Early Childhood Development and Cognitive Development in Developing Countries: A Rigorous Literature Review*. London: Department for International Development. <https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/ECD%202014%20Rao%20report.pdf?ver=2014-10-02-145634-017>.
- Reliefweb. 2020. *Colombia Violence against Children and Youth Survey 2018*. Bogotá: Reliefweb. <https://reliefweb.int/report/colombia/colombia-violence-against-children-and-youth-survey-2018>.
- Rolnick, Arthur J., and Rob Grunewald. 2003. *Early Childhood Development: Economic Development with a High Public Return*. Minneapolis, MN: Federal Reserve Bank of Minneapolis. <https://www.minneapolisfed.org/article/2003/early-childhood-development-economic-development-with-a-high-public-return>.

- Sadrudin, Aalyia F. A., Liliana A. Ponguta, Anna L. Zonderman, Kyle S. Wiley, Alyssa Grimshaw, and Catherine Panter-Brick. 2019. "How Do Grandparents Influence Child Health and Development? A Systematic Review." *Social Science & Medicine* 239: 112476. <https://doi.org/10.1016/j.socscimed.2019.112476>.
- Shelton, Charlotte D., Sascha Hein, and Kelly A. Phipps. 2019. "Resilience and Spirituality: A Mixed Methods Exploration of Executive Stress." *International Journal of Organizational Analysis* 28 (2): 399-416. <https://doi.org/10.1108/ijoa-08-2019-1848>.
- Sim, Amanda, Lucy Bowes, and Frances Gardner. 2019. "The Promotive Effects of Social Support for Parental Resilience in a Refugee Context: A Cross-Sectional Study with Syrian Mothers in Lebanon." *Prevention Science* 20 (5): 674-83. <https://doi.org/10.1007/s11121-019-0983-0>.
- Slone, Michelle, and Shiri Mann. 2016. "Effects of War, Terrorism and Armed Conflict on Young Children: A Systematic Review." *Child Psychiatry & Human Development* 47 (6): 950-65. <https://doi.org/10.1007/s10578-016-0626-7>.
- Smith, Bruce W., J. Alexis Ortiz, Kathryn T. Wiggins, Jennifer F. Bernard, and Jeanne Dalen. 2012. "Spirituality, Resilience, and Positive Emotions." In *The Oxford Handbook of Psychology and Spirituality*, edited by Lisa J. Miller, 37-454. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199729920.013.0028>.
- Smith, Gregory C., Bert Hayslip, Gregory R. Hancock, Frederick H. Strieder, and Julian Montoro-Rodriguez. 2018. "A Randomized Clinical Trial of Interventions for Improving Well-Being in Custodial Grandfamilies." *Journal of Family Psychology* 32 (6): 816-27. <https://doi.org/10.1037/fam0000457>.
- Smith, Karen E., and Seth D. Pollak. 2020. "Early Life Stress and Development: Potential Mechanisms for Adverse Outcomes." *Journal of Neurodevelopmental Disorders* 12 (1): 1-15. <https://doi.org/10.1186/s11689-020-09337-y>.
- Smith-Osborne, Alexa, Arati Maleku, and Sarolyn Morgan. 2017. "Impact of Applied Suicide Intervention Skills Training on Resilience and Suicide Risk in Army Reserve Units." *Traumatology* 23 (1): 49-55. <https://doi.org/10.1037/trm0000092>.

- Stammel, Nadine, Eva M. Abbing, Carina Heeke, and Christine Knaevelsrud. 2015. "Applicability of the ICD-11 Proposal for PTSD: A Comparison of Prevalence and Comorbidity Rates with the DSM-IV PTSD Classification in Two Post-Conflict Samples." *European Journal of Psychotraumatology* 6 (1): 27070. <https://doi.org/10.3402/ejpt.v6.27070>.
- Tamayo Martinez, Nathalie, Carlos Javier Rincón Rodríguez, Cecilia de Santacruz, Nubia Bautista Bautista, Jaime Collazos, and Carlos Gomez-Restrepo. 2016. "Mental Problems, Mood and Anxiety Disorders in the Population Displaced by Violence in Colombia: Results of the National Mental Health Survey 2015." *Revista Colombiana de Psiquiatría* 45: 113-18. <https://doi.org/10.1016/j.rcp.2016.09.004>.
- Tang, Yi-Yuan, Britta K. Hölzel, and Michael I. Posner. 2015. "The Neuroscience of Mindfulness Meditation." *Nature Reviews Neuroscience* 16 (4): 213-25. <https://doi.org/10.1038/nrn3916>.
- Thompson, Ross A., Mary Fran Flood, and Rebecca Goodvin. 2006. "Social Support and Developmental Psychopathology." In *Developmental Psychopathology: Risk, Disorder, and Adaptation, Vol. 3*, edited by Dante Cicchetti and Donald J. Cohen, 1-37. Hoboken, NJ: John Wiley & Sons, Inc.
- Tol, Wietse A., Suzan Song, and Mark J. D. Jordans. 2013. "Annual Research Review: Resilience and Mental Health in Children and Adolescents Living in Areas of Armed Conflict: A Systematic Review of Findings in Low- and Middle-Income Countries." *Journal of Child Psychology and Psychiatry* 54 (4): 445-60. <https://doi.org/10.1111/jcpp.12053>.
- Toth, Sheree L., Angeline Maughan, Jody Todd Manly, Mary Spagnola, and Dante Cicchetti. 2002. "The Relative Efficacy of Two Interventions in Altering Maltreated Preschool Children's Representational Models: Implications for Attachment Theory." *Development and Psychopathology* 14 (4): 877-908. <https://doi.org/10.1017/s095457940200411x>.
- Toth, Sheree L., Fred A. Rogosch, Jody Todd Manly, and Dante Cicchetti. 2006. "The Efficacy of Toddler-Parent Psychotherapy to Reorganize Attachment in the Young Offspring of Mothers with Major Depressive Disorder: A Randomized Preventive Trial." *Journal of Consulting and Clinical Psychology* 74 (6): 1006-16. <https://doi.org/10.1037/0022-006X.74.6.1006>.

- Ungar, Michael, Mehdi Ghazinour, and Jörg Richter. 2013. "Annual Research Review: What Is Resilience within the Social Ecology of Human Development?" *Journal of Child Psychology and Psychiatry* 54 (4): 348-66. <https://doi.org/10.1111/jcpp.12025>.
- UN High Commissioner for Refugees. 2018. *Global Trends: Forced Displacement in 2018*. <https://www.unhcr.org/globaltrends2018/>.
- Wachs, Theodore D., and Atif Rahman. 2013. "The Nature and Impact of Risk and Protective Influences on Children's Development in Low-Income Countries." In *Handbook of Early Childhood Development Research and Its Impact on Global Policy*, edited by Pia Rebello Britto, Patrice L. Engle, and Charles M. Super, 85-122. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199922994.003.0005>.
- Walker, Edward A., Elana Newman, Dorcas J. Dobie, Paul Ciechanowski, and Wayne Katon. 2002. "Validation of the PTSD Checklist in an HMO Sample of Women." *General Hospital Psychiatry* 24 (6): 375-80. [https://doi.org/10.1016/s0163-8343\(02\)00203-7](https://doi.org/10.1016/s0163-8343(02)00203-7).
- Werner, Emmy E. 2012. "Children at War: Risk, Resilience, and Recovery." *Development and Psychopathology* 24 (2): 553-58. <https://doi.org/10.1017/S0954579412000156>.
- Whitaker, Robert C., Tracy Dearth-Wesley, and Rachel A. Gooze. 2015. "Workplace Stress and the Quality of Teacher-Children Relationships in Head Start." *Early Childhood Research Quarterly* 30: 57-69. <https://doi.org/10.1016/j.ecresq.2014.08.008>.
- Whooley, Mary A., Andrew L. Avins, Jeanne Miranda, and Warren S. Browner. 1997. "Case-Finding Instruments for Depression: Two Questions Are as Good as Many." *Journal of General Internal Medicine* 12 (7): 439-45. <https://doi.org/10.1046/j.1525-1497.1997.00076.x>.
- Wilkins, Kendall C., Ariel J. Lang, and Sonya B. Norman. 2011. "Synthesis of the Psychometric Properties of the PTSD Checklist (PCL) Military, Civilian, and Specific Versions." *Depression and Anxiety* 28 (7): 596-606. <https://doi.org/10.1002/da.20837>.

- Wolmer, Leo, Daniel Hamiel, Jack D. Barchas, Michelle Slone, and Nathaniel Laor. 2011. "Teacher-Delivered Resilience-Focused Intervention in Schools with Traumatized Children Following the Second Lebanon War." *Journal of Trauma Stress* 24 (3): 309-16. <https://doi.org/10.1002/jts.20638>.
- Yale University and AÇEV. 2012. *The Ecology of Peace: Formative Childhood and Peace Building. A Brief Note*. New Haven, CT and Istanbul, Turkey: Yale-AÇEV Partnership. <https://medicine.yale.edu/childstudy/fcpb/approach/ecologyofpeace/>.
- Zolkoski, Staci M., and Lyndal M. Bullock. 2012. "Resilience in Children and Youth: A Review." *Children and Youth Services Review* 34 (12): 2295-2303. <https://doi.org/10.1016/j.childyouth.2012.08.009>.

APPENDIX

Table A1: Design and Implementation Processes of the 3Cs Program Modules (SBP and PTI)

t		
Source of Data/Information	Overview of Data/Information	Process of Integration of Data/Information
Literature and theoretical frameworks by expert team	The lead Fundación Saldarriaga Concha team conducted the review of best practices in resilience skills-building, informed by the Ecology of Peace Framework (Yale University and AÇEV 2012). It also held technical meetings with national experts and technical work consultations across the organization.	<ul style="list-style-type: none"> • The technical proposal was designed and presented to the ICBF, which recommended introducing pedagogical materials on other resilience promotion strategies that were being implemented in the country.
Participatory approach and community inputs	Interviews and focus group discussions were conducted with early childhood educators and caregivers from some of the target municipalities to inform the program design. Key areas of consultation included ways to frame content to promote uptake and acceptability, and delivery methods to facilitate delivery of the content.	<ul style="list-style-type: none"> • Framed all the content in “first person” to increase the extent to which participants related to it. • Employed collaborative learning to promote the discussion of content and its relation to everyday life experiences. • Introduced culturally relevant activities and practices, such as singing and dancing. • Included commitments and tasks to apply the content of the sessions at home, in the community, and/or in the workplace. • Adjusted schedules, location, and frequency of the program delivery based on caregiver groups’ preferences and availability.

EVALUATING THE 3CS PROGRAM FOR CAREGIVERS IN COLOMBIA

SBP IMPLEMENTATION		
Target Group	Overview of Content	Process of Implementation
<p>All parents of children enrolled in ECD centers and grandparents of children enrolled in ECD centers who self-identified as primary caregivers</p>	<p>Resilience</p> <ul style="list-style-type: none"> • Assertive and interpersonal communication • Stress management and emotional regulation • Decision-making, problem-solving, creative thinking, critical thinking <p>Wellbeing</p> <ul style="list-style-type: none"> • Self-knowledge • Empathy • Assertiveness • Life skills <p>Parenting practices</p> <ul style="list-style-type: none"> • Protective factors and safe, effective bonding • How to promote assertive communication in early childhood as a life skill that generates peacebuilding in different environments <p>Additional topics for grandparents</p> <ul style="list-style-type: none"> • Self-knowledge, interpersonal relationships, reconciliation and resilience, and realities of aging • Abuse, risk factors, violence, and intergenerational protection • Life skills • Transference of knowledge to other family members 	<ul style="list-style-type: none"> • Four lecture sessions (60-90 minutes per session) • Held in groups ranging from 15 to 20 participants • Sessions also included group discussions about specific strategies to promote resilience capacities

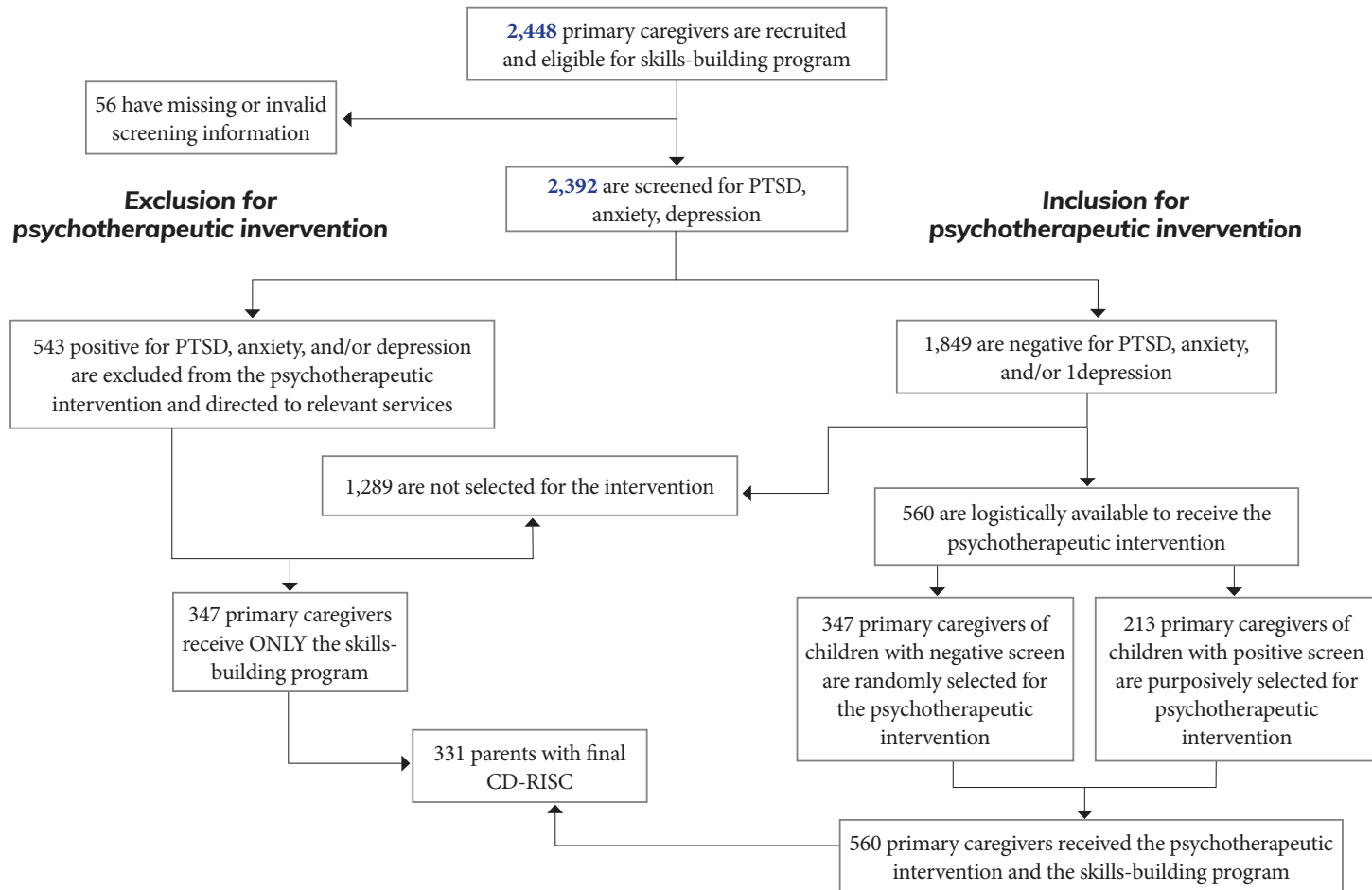
SBP IMPLEMENTATION (CONT.)		
Target Group	Overview of Content	Process of Implementation
Educators/Teachers	<p>Interactions and emotional support</p> <ul style="list-style-type: none"> • Safe and effective bonding • Socioemotional development in early childhood <p>Resilience</p> <ul style="list-style-type: none"> • Assertive communication • Relationships • Stress management—management of emotions 	Ten interactive workshops (60-90 minutes per session) Held in groups ranging from 15 to 20 participants. Sessions included activities to openly discuss and internalize the concepts addressed in the lecture sessions. Activities, such as games and role-play, were used to reinforce knowledge, elaborate on concepts through active questioning, and state personal commitments and good practices.

PTI DESIGN		
Source of Data/Information	Overview of Data/Information	Process of Integration of Data/Information
Theory/Literature	<p>A review of the literature was conducted to identify best practices for effectively promoting resilience capacities in conflict-affected settings, including those designed for parents of young children (with no clinical presentation of anxiety, depression, or PTSD). Based on the literature review, third-generation cognitive behavioral techniques were selected as part of the program content.</p> <p>Group psychotherapy was selected, based on literature review. With results similar to those of individual psychotherapy, group interventions have some advantages. They allow a larger number of people to be treated by each available therapist, and they reinforce positive beliefs not available in individual therapy, such as a sense of belonging, peer support, and feeling a connection to a group of people who value a shared environment.</p>	The technical proposal was designed and presented to the ICBF. Additional consultations were held with a group of psychiatrists with experience in community interventions, who made final adjustments to the proposed model.
Pilot	A pilot study was conducted to optimize the content and program modality in one municipality (Medellín).	The results of the pilot led to shortening the intervention and to summarizing the contents the subjects found similar.

EVALUATING THE 3CS PROGRAM FOR CAREGIVERS IN COLOMBIA

PTI IMPLEMENTATION		
Target Group	Overview of Content	Process of Implementation
<p>Parents who self-reported to be victims of the armed conflict and had a negative screen for depression, PTSD, and/or anxiety.</p>	<p>Parental psychosocial wellbeing and resilience</p> <ul style="list-style-type: none"> • Characterization and initial psychoeducation • Behavioral activation model • Rational behavioral emotional therapy • Management of emotions • Mindfulness • Acceptance and commitment techniques 	<p>The overarching aim of the PTI was to promote enduring bonding and provide strategies to promote resilience, social skills, emotional processing, presentation techniques, activation control techniques, and self-control. The PTI module consisted of eight in-person sessions, each lasting an average of two hours.</p>

Figure A1: Recruitment, Screening, and Group Assignment to the 3C Program Modules (SBP and PTI)



HOW FAMILY RELATIONSHIPS PREDICT THE EFFECTIVENESS OF A PSYCHOSOCIAL GROUP INTERVENTION AMONG WAR-AFFECTED CHILDREN

RAIJA-LEENA PUNAMÄKI, KIRSI PELTONEN, MARWAN DIAB,
AND SAMIR R. QOUTA

ABSTRACT

Family relationships habitually shape the way traumatic events affect children's mental health in a context of war and violence, but research is scarce on the role these relationships play in the success of psychosocial interventions. This study is a secondary analysis of previously identified family system types that are based on attachment, parenting, and siblingship, and of the influence they have on the effectiveness of a psychosocial group intervention (the Teaching Recovery Techniques, TRT). The TRT is aimed at reducing children's mental health problems and increasing their psychosocial resources. We tested three noncompeting hypotheses based on family system dynamics. First was the compensation hypothesis, which holds that children from families with negative relationships benefit a great deal from the TRT intervention. The second was the accumulative hypothesis, which maintains that children from families with negative relationships do not benefit from the intervention. The third, the buffering hypothesis, states that children from families with positive relationships benefit a great deal from the intervention. The family sample consists of 325 Palestinian mothers and fathers and one child from each family between the ages of 10 and 13. The children participated either in the TRT intervention or control groups. Their self-reported posttraumatic stress symptoms, emotional and conduct problems, positive resources, and prosocial behavior were assessed at baseline, three months post-intervention, and at a six-month follow-up. We found that family type was significantly associated with TRT effectiveness, which supports the compensation and buffering hypotheses. In the

Received January 18, 2021; revised September 27, 2021; accepted November 4, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

intervention group, children with insecure and negative family relationships showed a reduction in emotional problems across the three assessments, and an increase in positive resources from baseline to post-intervention. Children with secure and positive family relationships showed a reduction in emotional problems and increase in positive resources also in the control group. We argue that a family system approach can deepen understanding of the mechanisms of successful psychosocial interventions and, therefore, that family relations should be taken into account when tailoring such interventions for traumatized children.

INTRODUCTION

War as a developmental environment puts overwhelming demands on children and their families. Losses, horrors, and threats to life force them to seek a balance between their strengths and vulnerabilities. Researchers have documented both. The research shows associations between children's exposure to traumatic war events and increased posttraumatic stress disorder (PTSD), as well as symptoms of depression, anxiety, and dissociation (Charlson et al. 2019; Slone and Mann 2016). Some studies also show an increase in aggression and antisocial behavior (Keresteš 2006; Qouta et al. 2008). However, war-affected children also demonstrate psychosocial resources and can even blossom despite trauma, which is conceptualized as resilience (Masten and Narayan 2012) and evidenced by observations of intact cognitive and emotional skills (Punamäki, Qouta, and El-Sarraj 2001), an improved sense of social affiliation (Diab 2011; Punamäki et al. 2006), and phenomenal recovery when help is received (Bonanno 2004; Tol, Song, and Jordans 2013).

A number of community- and school-based interventions have been designed to prevent and heal children's mental health and developmental problems and to enhance their psychosocial resources and resilience (Purgato et al. 2020). Their healing elements include social and emotional learning (SEL) and psychosocial support (PSS), such as creative expression through dance, storytelling, music, and psychodrama. PSS/SEL also can reinforce a child's sense of safety, mastery, and belonging through cooperative games, reflective thinking, group cohesion, and emotion expression and regulation (Kangaslampi and Peltonen 2019; Schnyder et al. 2015). Ultimately, the aim is to reinforce children's individual and social resources and strengths and support their empowerment, effective coping strategies, and trust in self and others, which in turn can reduce psychological suffering (Bosqui and Marshoud 2018). However, children who participate in psychosocial interventions can differ greatly in terms of family resources, support, and security, which may affect their potential to benefit from the interventions.

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

Research is scarce on the impact family relations have on the effectiveness of interventions for war-affected children. Accordingly, the present study examines how the quality of family attachments, parenting, and siblingship helps children benefit from psychosocial intervention activities, thereby enhancing the success of the intervention.

FAMILY SYSTEM DYNAMICS OF RECOVERY FROM TRAUMA

According to family system theories, parents and children face traumatic events together and show endurance, manifest symptoms, and care for each other in different and unique ways (Crittenden and Dallos 2009; Montgomery 2004; Punamäki, Qouta, and Peltonen 2017; Riggs and Riggs 2011). These theories offer the possibility of understanding how trauma affects children's mental health and how and why successful interventions can help them. Family experiences can result in compensatory, accumulative, or buffering dynamics in members who manifest mental health problems, social support, or emotional sharing (Coyne, Downey, and Boergers 1992; Minuchin 1974; Punamäki et al. 2010). Compensation dynamics suggest that maternal problems, such as depression, do not constitute a risk for child development if the children enjoy a good relationship with their father (Vänskä et al. 2015) or warm and intimate siblingship (Peltonen et al. 2010). Accumulative dynamics in turn refer to a spillover of negative responses from marital and parenting systems into siblingship, thus multiplying family stressors (Lindblom et al. 2014; Lindblom et al. 2017; Minuchin 1974). Buffering dynamics indicate that traumatic war events do not pose a risk to children's mental health in families with supportive, wise, and sensitive parents (Montgomery 2004).

Based on family system dynamics, we may analogously propose similar dynamics concerning the role family relationships play in psychosocial interventions for war-affected children. According to compensation dynamics, children who have insecure and unsupportive parent and sibling relationships enjoy successful intervention effects (reduced symptoms and increased psychosocial resources), as the intervention experiences give them opportunities to satisfy their deep need for care, genuine listening, sharing, and attention. In contrast, according to accumulation dynamics, children with insecure and unsupportive family relationships experience unsuccessful intervention effects (stable or even increased symptoms and reduced resources), due to the stress caused by family burdens and a spillover of mistrust to other adults and peers, which prevents the children from engaging in and benefitting from intervention activities. Finally, according to buffering dynamics, children with secure and supportive family relationships enjoy successful intervention effects, due to family-based protection

from traumatic war events, readiness to trust adults and peers to get help, and adaptive expression and regulation of emotion. In this study, we test whether these family system dynamics also work in terms of the intervention's effectiveness. The corresponding compensation, additive, and buffering hypotheses are not competing and can exist simultaneously.

FAMILY RELATIONSHIPS IN PSYCHOSOCIAL INTERVENTIONS

Research on the effectiveness of interventions conceptualizes family relationships as an underlying mechanism for outcome change (statistical mediators) or as factors that affect the degree of outcome change (statistical moderators). As typical mediators, psychosocial interventions can enhance family mental health and security and improve parents' capacity to deal with traumatized children, often by providing psychoeducation and support (Betancourt et al. 2013; Jordans, Pigott, and Tol 2016). Research has shown that moderating factors such as family structure (e.g., single-parent families, family separation, foster parenting), socioeconomic status, and family connectedness affect the success of psychosocial interventions (Betancourt et al. 2012; Bryant 2016; Panter-Brick, Grimon, and Eggerman 2014). However, few studies have empirically tested the effects of family relationship quality on intervention-induced improvements, even though supportive parenting and family security are considered essential to supporting recovery among traumatized children, both in general (Cohen, Mannarino, and Murray 2011) and in conditions of war and political violence (Bosqui and Marshoud 2018; Jordans et al. 2016).

We found one study that examined the impact parent-child attachment relationships have on the effectiveness of a school-based psychosocial intervention for war-affected Palestinian children (Eloranta et al. 2017). The results showed that the intervention was effective in reducing PTSD symptoms among secure and preoccupied children, but not among those who were avoidant. Importantly, secure children's symptoms also decreased in control conditions, whereas those of the avoidant children in the control group increased. We were unable to find research on the role siblingship plays in the success of psychosocial interventions. One Palestinian study (Diab et al. 2014) found that a psychosocial intervention could reduce sibling rivalry among girls, although it was not effective enough to produce optimal sibling relations that provided warmth and intimacy. Sibling conflict increased in the control conditions. Two studies confirmed that warm and intimate sibling relationships can prevent traumatic war events from negatively affecting the mental health of Palestinian children (Diab, Guillaume, and Punamäki 2018; Peltonen et al. 2010).

Research shows that treatments are effective in providing parental support, competence, and optimal parent-child communication in families exposed to interpersonal and community violence (Barber, Stoltz, and Olsen 2005; Johnson et al. 2018). However, we could not find studies that examined the influence of parenting quality on intervention outcomes among war-traumatized children, despite abundant research showing that high-quality parenting enhances recovery from trauma (Eltanamy et al. 2021; Feldman et al. 2013; Thabet et al. 2009).

System theories conceptualize family relationships comprehensively, which involves multiple subsystems, the most important being parent-child bonds and attachments, parenting, and siblingship (Minuchin 1974). According to attachment theories, experiencing danger and threats intensifies protection-motivated family dynamics, which partly legitimizes the examination of the effects different family types have on recovery from war trauma (Bowlby 1969; Masten and Monn 2015; Mikulincer et al. 1999). In this study, we utilize the person-oriented approach inherent in family system theories, which enables us to capture the complexity of multiple dyadic and triadic relationships by identifying homogenous groups (Bergman, Magnusson, and El Khouri 2003). Rather than focusing on separate dimensions that describe attachment relationships, parenting styles, and siblingship, the person-oriented approach summarizes dynamic information to identify unique family types in the context of war. Qualitatively different family types have distinct and meaningful effects on children's development, as the young ones learn to adjust their social, emotional, and cognitive responses to match their specific family environment (Coyne et al. 1992; Crittenden and Dallos 2009).

CONTEXT OF THE STUDY

Palestinians, who currently live under foreign military occupation, have a long history of accumulated war trauma and deportations. Families in Gaza are severely affected by the current military siege and repeated Israeli military operations, and by an international economic boycott imposed in 2007 as a response to the Islamic Hamas party winning national elections. As a result, Palestinians in Gaza are denied freedom of movement, access to clean water, and regular electricity (World Bank 2015). The current study was conducted in the aftermath of the war on Gaza in 2008-2009 (Operation Cast Lead in Israeli military terms). The 23-day war resulted in 1,417 Palestinians dead, including 313 children, and 5,303 injured, among them 1,606 children. Approximately 100,000 people were displaced due to their homes being destroyed and had to seek refuge in the besieged Gaza Strip (UNHRC 2009; UN OCHA 2009).

The participating families belong to Islamic Arab culture, where parents' main tasks are to protect their children and secure their physical, spiritual, and social development. Children, who are dependent on their parents, show them great respect (Al-Krenawi and Graham 2005). These cultural norms emphasize social harmony and connectedness rather than individual thriving, ambition, and autonomy (Kagitcibasi 2005). Parenting goals among these families include teaching appropriate behavior (politeness, respecting elders, obedience, family loyalty), decency (honesty, charity, and responsibility), and lovingness (sharing, loving family, and compassion) (Al-Hassan and Takash 2011).

Some research shows that the Palestinian national struggle for liberation has affected marriage and parenting traditions. Traumatic events with political, military, and ideological connotations have been found to empower rather than harm women, whereas everyday stress constitutes a risk for their mental distress (Diab 2011; Khamis 1998; Punamäki 1986). For instance, as almost one-third of Palestinian men were detained, imprisoned, or deported during the First Intifada (1987-1991), a national uprising, women took responsibility for political and community actions (Khamis 1998). In terms of parenting, one study found that Palestinian mothers exposed to severe traumatic war events endorsed more relatedness and autonomy as desirable characteristics of their children (Kankaanpää et al. 2020). The Palestinian national struggle demands that mothers socialize their children for traditional relatedness while also instilling in them the autonomy they need in the face of military and war violence. It thus is possible that individual and collectivistic values are not opposing but complementary and dynamic (Kagitcibasi 2005).

RESEARCH AIMS

The present study involves a secondary analysis of the role played by four previously identified family types before the children participated in the Teaching Recovery Techniques (TRT) psychosocial intervention (Punamäki et al. 2017). The first aim was to examine how these family types (highly secure and positive relationships, moderately secure and neutral relationships, discrepant experiences, and highly insecure and negative relationships) influence the effectiveness of the TRT intervention among war-affected Palestinian children. The criteria for intervention effectiveness were the following: a decrease in PTSD symptoms; a decrease in emotional and conduct problems; and an increase in positive resources and prosocial behavior. According to the compensation hypothesis, children from families with insecure and negative relationships and from families with discrepant experiences would show decreased symptoms and

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

increased psychosocial resources at baseline, post-intervention, and follow-up assessments. According to the accumulation hypothesis, children from families with insecure and negative relationships and discrepant experiences would not show reduced mental health problems and increased psychosocial resources across the assessments. Finally, according to the buffering hypothesis, children from families with secure and positive relationships and those with moderately secure and neutral relationships would show reduced mental health problems and increased psychosocial resources across the assessments. The second aim was to analyze how family types are generally associated with children's mental health problems and positive resources, and the third aim was to determine whether traumatic war events and child gender influence the impact family type has on the intervention's effectiveness.

METHOD

PARTICIPANTS AND PROCEDURE

The current study uses data from a subsample of 325 families from a larger randomized intervention study (N=482, Punamäki et al. 2017) that examined the effectiveness of a psychosocial intervention program three months after the Gaza War in 2008-2009. The subsample consisted of families in which both parents and one child between age 10 and 13 (M=11.35, SD=0.57; 49% girls) from each family responded to the questionnaires. The children participated at baseline, pre-intervention (T1), three months after, at post-intervention (T2), and six months later in a follow-up (T3).

The children and their parents were given information sheets about the intervention, which explained the study procedure and asked about their willingness to participate in the questionnaire. Six research assistants collected the children's data in the classrooms, and the children took the parents' questionnaires home to be completed and returned them to the assistants in closed envelopes. The fourth author of this paper (Samir R. Qouta) supervised the data collection through weekly consultations with research assistants and school visits.

In the data on the families, the participation rates were 68.2 percent (n=377) for mothers and 68.0 percent (n=328) for fathers. The subsample of 325 families did not differ from the families that did not participate (n=157) in terms of fathers' or mothers' work status (respectively, $\chi^2(1)=0.38$, $p=ns$; and $\chi^2(1)=0.01$, $p=ns$); place of residence ($\chi^2(1)=0.28$, $p=ns$); family structure ($\chi^2(2)=2.3$, $p=ns$); or family size

($\chi^2(2)=0.11$, $p=ns$). However, the family sample was more biased toward having girls as the target child (56.1%) than the nonparticipating families (35.2%) ($\chi^2(1)=17.72$, $p<.0001$).

The sampling of the randomized intervention study participants involved four phases. First, we selected two regions from the Gaza Strip that were heavily bombed during the war on Gaza in 2008-2009, North Gaza and Gaza City. Second, we selected the participant schools using a simple random-sampling algorithm (8 schools from 160 potential schools located in the two areas). Third, one sixth-grade and one seventh-grade class were randomly chosen in the eight schools, resulting in 16 classes being used in the study. Finally, of the 16 total school classes, every second class was randomly selected and placed in either the intervention or the control group ($n=242$ and $n=240$).

The ethics boards of the Palestinian Ministry of Education and the Gaza Community Mental Health Program reviewed and accepted the study's protocol and measurements, and the school authorities granted permission to conduct the study.

THE INTERVENTION

The TRT is a manualized intervention procedure with clear session procedures developed by Smith, Dyregrov, and Yule (2000) for the Children and War Foundation. The intervention, which consisted of two weekly two-hour sessions, lasted for four weeks. The counsellors followed an Arab-language manual, and fidelity was ensured by holding weekly preparatory and supervisory meetings. The TRT involves evidence-based tools based on elements of cognitive behavioral therapy, and it aims to help children develop effective coping skills, empowerment, and emotion regulation through narrative, imagery, and body- and mind-related and psycho-educational techniques. The intervention techniques were aimed at reducing intrusion, avoidance, and arousal symptoms by enhancing children's symbolic, verbal, and kinaesthetic processing of traumatic experiences.

All sessions started with a warm-up, an introduction to the topic, and a review of the homework. It was crucial to create a sense of safety and to provide meaningful tools that helped the children frame and regulate their overwhelming negative emotions and painful experiences, to recognize their own and others' stress reactions, and to revitalize their numbed feelings, as well as to enhance peer relations in the group. Problem-solving, storytelling, drawing pictures of their frightening experiences and dreams, and role-play techniques were also applied.

Learning and practicing how to regulate their fears and horrors were introduced in a relaxed and playful manner, which taught the children to link their bodily sensations with their traumatic experience, feelings, and emotions. The parents were informed about and encouraged to join in the children's TRT homework, which included sleep hygiene, and talking about dreams and nightmares.

MEASURES

FAMILY TYPES

In a previous study (Punamäki et al. 2017), four family types based on parents' and children's attachment relationships, parenting practices, and sibling relationships were identified by cluster analysis. Parental attachment security was measured by a ten-item scale of "parents' acceptance and willingness to serve as an attachment figure" (Kerns et al. 2000). Mothers and fathers used a five-point scale to indicate how well the descriptions corresponded to their attitudes and behavior toward the target child. The averaged sum variables that were formed showed moderate reliability (Cronbach's α -values .69 for mothers and .68 for fathers). The scale had not previously been used with Palestinian adults. Children's attachment style was measured by the shortened coping strategies questionnaire (CSQ; Finnegan, Hodges, and Perry 1996) and the security scale (Kerns, Klepac, and Cole 1996), which involved 28 everyday situations that reflected avoidant, preoccupied, and secure attachment to mothers. Children answered using two-stage methods (Harter 1982) to evaluate their typical responses to threatening situations. The sum variables formed showed low to moderate reliabilities (Cronbach's α -values .66 for felt security, .63 for avoidant, and .54 for preoccupied attachment styles). The scales had not been used previously with Palestinian children. Parenting practices were assessed by the 20-item child psychological maltreatment questionnaire (Khamis 2000), on which the child participants estimated how well the descriptions fit their parents, and mothers and fathers responded about their own behavior toward the target child using a five-point scale. Three averaged sum variables were constructed for emotional abuse, emotional neglect, and harsh parenting, both separately for the child and combined with the parents (Cronbach's α -values ranging between .77 and .89). The questionnaire had been used previously with Palestinian children and showed good reliability and validity (Khamis 2000). Siblingship quality was assessed by the scale developed by Dunn et al. (1994), which reflected positive (warmth and intimacy) and negative (conflict and rivalry) relationships. Children estimated their interactions with an older (11 items) and a younger (11 items) sibling using a five-point scale. Averaged sum variables

were constructed for warmth, intimacy, conflict, and rivalry in siblingship with reasonable reliability (Cronbach's α -values .72, .68, .75, and .79, respectively). The siblingship quality scale had been used previously among Palestinian children, and it showed sufficient reliability and good validity (Diab et al. 2018; Diab et al. 2014; Peltonen et al. 2010).

We used cluster analyses to identify distinct family types, based on 15 sum variables of mother, father, and child attachment, parenting practices, and sibling relationships. Our analysis involved a hierarchical cluster analysis to define the number of initial clusters and the dendrogram for their visual inspection, as well as a K-means cluster analysis to confirm the cluster membership (Tabachnick and Fidell 2013). The analyses identified four family types. In families with security and positive relationships (36.2%, $n=102$), both parents had high secure attachment and the child showed low avoidant attachment, high warmth and intimacy, and low conflict and rivalry in siblingship; all family members reported low levels of abuse and harsh parenting. In families experiencing insecurity and negative relationships (15.6%; $n=44$), both parents and the child reported low secure and high avoidant attachments and high abusive and harsh parenting practices, and the siblingships were characterized by high conflict and rivalry and low warmth. In the families with moderately secure and neutral relationships (25.2%; $n=71$), both parents showed low secure attachment, while the child showed mixed secure and insecure attachments, combined with positive sibling relationships; all family members reported moderate levels of abusive and harsh parenting practices. Finally, in the families with discrepant experiences (23.0%; $n=65$), both parents showed high secure attachment, while the child showed low secure and high avoidant attachments. The parents reported very low abuse and harsh parenting practices, while the children reported that both were high. For a more detailed description of these measures and methods for forming the family system types, see Punamäki et al. (2017).

CHILDREN'S MENTAL HEALTH: SYMPTOMS AND RESOURCES

CHILDREN'S PTSD

We evaluated the symptoms using the 13-item Children's Revised Impact Event Scale (CRIES-13; Dyregrov, Gjestaad, and Raundalen 2002). The scale covers the three core dimensions: re-experiencing (4 items), avoidance (4 items), and hyperarousal (5 items) symptoms. Children indicated on a four-point scale how often they had experienced each symptom during the previous two weeks (from 0, "not at all," to 4, "often"). A total score was constructed showing moderate

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL GROUP INTERVENTION EFFECTIVENESS

reliabilities (Cronbach's α -values .62 at T1, .72 at T2, and .63 at T3). The CRIES-13 has been shown to be reliable and is validated for use with Palestinian children (Kolltveit et al. 2012; Veronese et al. 2019), although use of the PTSD concept in ongoing life-threatening conditions has been criticized (Altawil, El Asam, and Khadaroo 2018).

CHILDREN'S EMOTIONAL AND CONDUCT PROBLEMS

We used the Strengths and Difficulties Questionnaire (SDQ; Goodman, Meltzer, and Bailey 1998) to measure emotional and conduct problems. Emotional problems include symptoms of depression and anxiety, and conduct problems include hyperactive and aggressive behaviors. All four scales contain five items that describe behaviors, thoughts, and moods. The children used a three-point scale to estimate how well the description fit them (0, "not at all," 1, "somewhat," 2, "yes, fits well"). We constructed sum scores for emotional and conduct problems, both of which showed good reliability (Cronbach's α -values for emotional problems .71 at T1, .73 at T2, and .72 at T3, and for conduct problems .79 at T1, .76 at T2, and .75 at T3). The SDQ had been used previously with Palestinian children, and showed good or moderate reliability and good validity (El-Khodary and Samara 2020; Peltonen et al. 2010; Thabet, Stretch, and Vostanis 2000).

CHILDREN'S POSITIVE RESOURCES

We used the 14-item Mental Health Continuum-Short Form (MHC-SF) for youth (Keyes et al. 2008) to measure the degree of emotional and social resources available to children. The scale includes dimensions of positive emotional affects ("I have warm and trusting relationships with others"), psychological autonomy and self-acceptance ("I feel happy"), and social contribution and coherence ("The way our society works makes sense to me"). Children used a five-point scale to evaluate how often they had felt or thought in the described ways during the previous month (0, "never," 1, "rarely," 2, "sometimes," 3, "often," 4, "every day"). We calculated total sum variables (Cronbach's α -values .83 at T1, .82 at T2, and .85 at T3).

PROSOCIAL BEHAVIOR

We used the five-item prosocial behavior scale of the SDQ (Goodman et al. 1998) that covers willingness to share with and help others. Examples of the prosocial behavior items are "I usually share toys and school tools with other children" and "I help other people if something bad happens to them or if I see them upset."

The children used the three-point scale to evaluate how well the descriptions fit them (0, “not at all,” 1, “somewhat,” 2, “yes, fits well”). We constructed a sum variable of prosocial behavior, but the internal consistency was weak (Cronbach’s α -values .63 at T1, .62 at T2, and .65 at T3).

CHILDREN’S TRAUMATIC WAR EXPERIENCES

We measured traumatic war experiences by a scale of 14 traumatic events corresponding to Criterion A of the PTSD diagnosis in *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association 2013), which include experiencing and witnessing actual or threatened serious injury or death. The children reported whether they had had such an experience during the war (0, “no,” 1, “yes”). We constructed a sum variable by counting the “yes” answers.

DEMOGRAPHIC VARIABLES

Mothers and fathers reported family income, parental education, work situation, family size, nature of the family (extended or core), and their children’s ages and genders.

TRANSLATIONS

The research instruments for sibling relationships, parenting practices, CRIES-13, SDQ, and traumatic war experiences scales were available in Arabic. For the present study, a bilingual psychologist first translated the children’s and parents’ attachment scales from English into Arabic and the research group then translated it back.

STATISTICAL ANALYSES

We conducted all analyses using IBM SPSS Statistics 25. We present the distributions of demographic variables as percentages, and we determined correlations between study variables using the Pearson Cross Product method. To analyze how family type is associated with the effectiveness of the TRT psychosocial intervention, we applied multiple repeated-measure analyses of variance with covariates (MANCOVA) to a three-wave (baseline T1, post-intervention T2, and follow-up T3) assessment of PTSD symptoms, emotional and conduct problems, and positive resources and prosocial behavior. We calculated the main effects for time (T1, T2, and T3, indicating dependent variables) and independent between-subject variables of the intervention versus the control group and four family types (security and positive relationships, insecurity and negative relationships, discrepant

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

experiences, and moderately secure and neutral relationships). We further tested the interaction term between the intervention and family types, and included children's traumatic war experiences and gender as the covariates. We analyzed the tests of within-subject contrasts (three-wave assessment) as either linear or quadratic.

The sample included 16 school classes, with 30 pupils per class on average. The nonindependence of their responses could cause bias, due to the reduced sample variation in cluster sampling (Ukoumunne et al. 1999). The within-school class biases were checked by estimating the intraclass correlations, the average being .036 (CI=.018-.059), and the design effects (DEFFs, all <2.00). The design effects were close to one for the attachment, parenting practices, and siblingship variables, and ranged between 1.5 and 2.78 for emotional and conduct problems in some of the assessment points.

There may be floor and ceiling effects, as children from different family types may report either maximum or minimum scores of mental health problems and psychosocial resources. The criterion for the flooring effect is that more than 15 percent of the children from a certain family type would score the lowest values of mental health problems or resources. The corresponding ceiling effect is that more than 15 percent of children from a certain family type would score the highest values of mental health problems or resources. We found a flooring effect for conduct problems (22.5%) in children from families with secure and positive relationships, and a ceiling effect for prosocial behavior in children from families with secure and positive relationships (27.4%) and from families with moderately secure and neutral relationships (22.5%). However, as the values were not outliers, we did not replace them.

RESULTS

DESCRIPTIVE STATISTICS

Table 1 shows the demographic characteristics reported by the parents and children. About a quarter (24%) of the fathers had a university education, while fewer than 10 percent of the mothers had one. Despite their education, about half (49%) of the fathers were unemployed, and almost all (93%) of the mothers worked at home. The statistics correspond with the problematic economic and social situation in the Gaza Strip that is largely caused by the Israeli military siege and international economic boycott (UN OCHA 2009). As for family size, about

a quarter (26%) of families had more than eight children, and almost one-third (29%) of the participants lived in extended families.

Table 1: Percentage and Frequencies of Demographic Factors

	%	N
Place of living ^a		
City	84.3	284
Refugee camp	3.3	11
Village	12.5	42
Status ^a		
Refugee	11.3	38
Citizen	88.7	299
Father's education ^b		
Elementary	21.1	69
Preparatory	28.3	93
Secondary	26.2	86
University	24.4	80
Mother's education ^a		
Elementary	19.6	66
Preparatory	32.4	109
Secondary	39.9	134
University	8.0	27
Father's work situation ^b		
Unemployed	49.3	161
Worker	12.8	42
Public employee	24.9	82
Entrepreneur/ self-employed	13.1	43
Mother's work situation ^a		
Works at home	93.2	314
Worker	3.0	10
Public employee	3.9	13
Family type ^a		
Immediate	61.9	209
Extended	28.9	97
Tribe	9.1	31

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

	%	N
Family size ^a		
Small (1-4)	23.8	80
Medium (5-7)	50.0	168
Large (8 or more)	26.2	88

Note: ^a N=336-337; mother-reported, difference due to missing information; ^b N=328; father-reported

The supplemental table (see Appendix) presents the means, standard deviations, and bivariate correlations between children’s mental health problems and psychosocial resources and the covariates at baseline, post-intervention, and follow-up. The results show that emotional and conduct problems were significantly correlated within a single assessment and between the three assessments, whereas PTSD symptoms were not significantly correlated with emotional and conduct problems within or between assessment times. However, PTSD symptoms were correlated with each other between the three assessment times. Traumatic war experiences were positively correlated with children’s mental health problems at the three assessment times, except for follow-up PTSD symptoms. Correlation analysis further revealed that boys reported more conduct disorders and girls more positive resources at all assessment times. Boys reported more than girls traumatic war experiences.

FAMILY TYPE, INTERVENTION, AND CHILDREN’S MENTAL HEALTH

Table 2 summarizes the MANCOVA’s main and interaction effects between intervention and family type on children’s mental health problems and psychosocial resources at baseline, post-intervention, and follow-up (F and η^2 values). In terms of the impact family type has on the effectiveness of the intervention, the results showed significant interaction effects on emotional problems and positive resources between the intervention and family types, whereas the intervention and family type had separate main effects on PTSD symptoms and prosocial behavior, and family type alone had a main effect on conduct problems. Figures 1a and 1b illustrate the impact family type had on changes in emotional problems in the intervention (1a) and control (1b) groups, and on changes in positive resources in the intervention (Figure 2a) and control (Figure 2b) groups.

Table 2: Summary of Repeated-Measure MANCOVA Main and Interaction Effects of Psychosocial Intervention and Family Types on Change in PTSD^a Symptoms, Emotional and Conduct Problems, and Positive Resources across Three-Wave Follow-Up: F-values and Partial Eta Squared η^2

	PTSD Symptoms		Emotional Problems		Conduct Problems		Positive Resources		Prosocial Behavior	
Within-subject ANCOVA ^b	F-values ^c	η^2	F-values ^c	η^2	F-values ^c	η^2	F-values ^c	η^2	F-values ^c	η^2
<i>Main effects</i>										
Time (pre-, postfollow-up)	3.28* 4.42* _{Linear}	0.01	5.87** 8.22** _{Linear}	0.02	6.47** 10.80*** _{Quadratic}	0.03	2.41 3.91* _{Quadratic}	0.01	2.66 3.79* _{Linear}	0.01
Intervention groups	5.77** 9.77** _{Linear}	0.02	1.04	0.01	0.66	0.01	0.31	0.01	2.88* 4.48* _{Linear}	0.01
Family type	2.10	0.03	1.95	0.02	2.55* 2.61* _{Quadratic}	0.03	3.35** 5.34*** _{Linear}	0.04	2.99** 3.06* _{Linear}	0.04
<i>Interaction effects</i>										
Intervention * Family type	0.91	0.01	2.06* 3.84** _{Quadratic}	0.03	0.64	0.01	2.79** 4.34** _{Linear}	0.03	0.70	0.01
<i>Covariates</i>										
Traumatic war events	1.67	0.01	0.97	0.01	0.92	0.01	1.83	.01	0.76	0.01
Gender	1.83	0.01	0.89	0.01	3.30* 4.88* _{Linear}	0.02	0.68	.01	1.02	

Note: ^a PTSD symptoms; ^b Mauchly's test did not violate the sphericity assumption for PTSD, positive resources, and prosocial behavior, but violated the sphericity assumption for emotional problems, $F(2,305)=11.42$, $p<.003$, and conduct problems, $F(2,305)=9.18$, $p<.01$. For these variables, Huyn-Felt corrected results are reported (for others Greenhouse-Geisser). ^c The upper F-values are for within-subject effects and the lower F-values are for within-subject contrasts (either linear or quadratic; not reported, if within-subject effect is nonsignificant)

* $p<.05$, ** $p<.01$, *** $p<.001$

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

Figure 1a: Emotional Problems According to Family Types—
Intervention Group

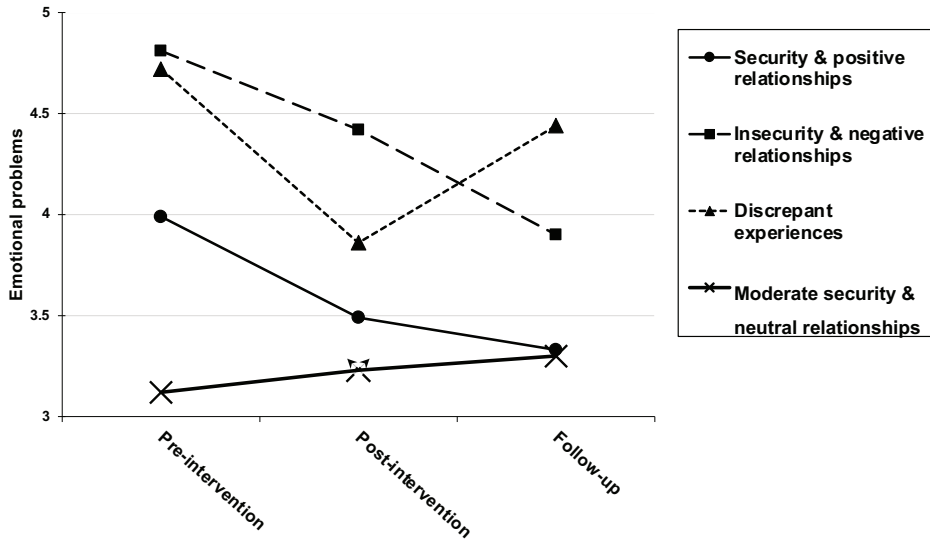


Figure 1b: Emotional Problems According to Family Types—Control Group

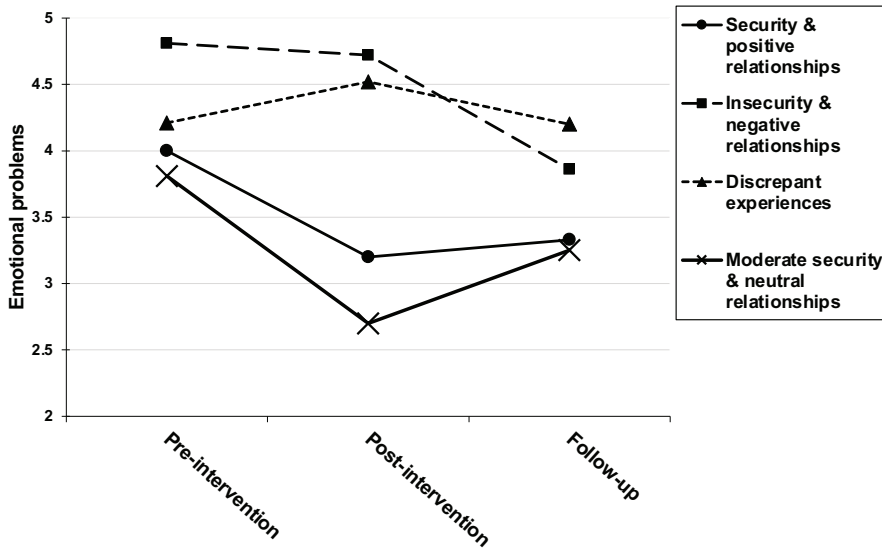


Figure 2a: Positive Resources According to Family Types—Intervention Group

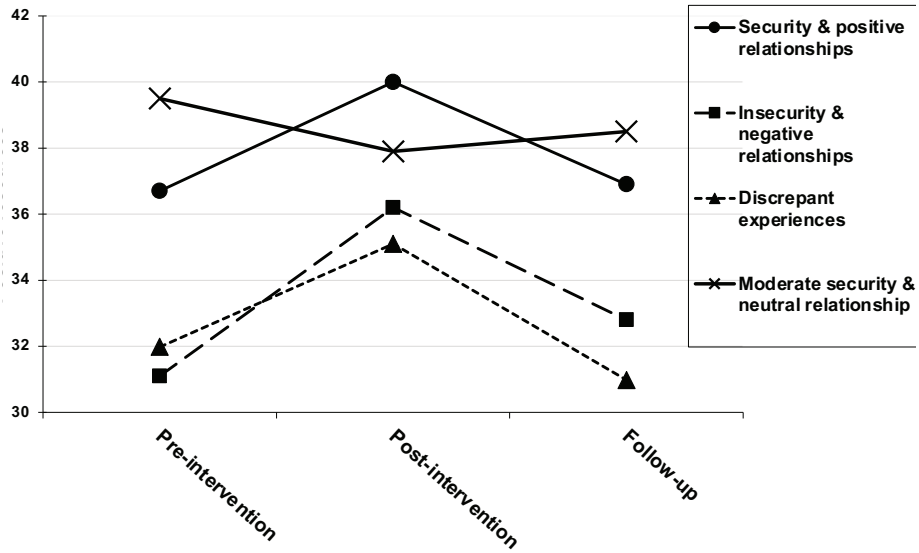
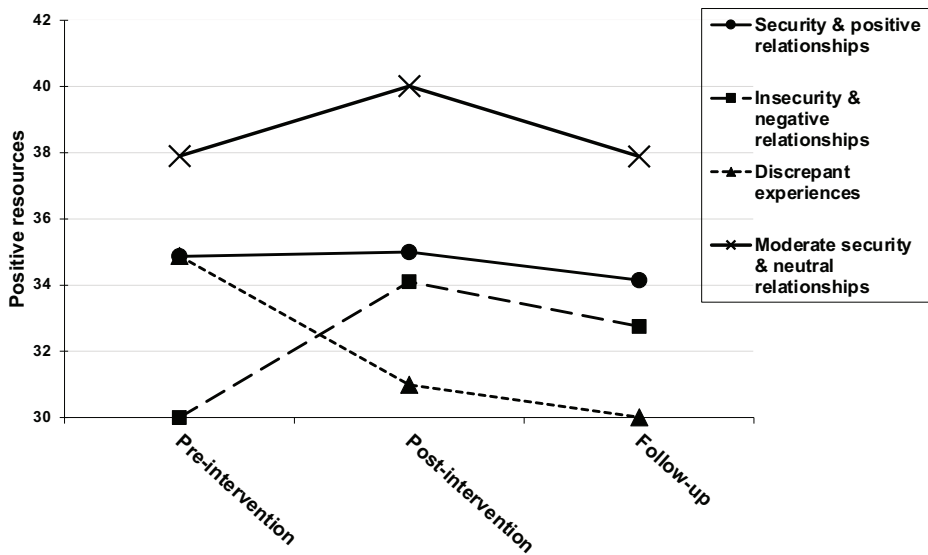


Figure 2b: Positive Resources According to Family Types—Control Group



In accordance with the compensation hypothesis, children from families with insecure and negative relationships and discrepant experiences benefitted from the TRT intervention. As illustrated in Figure 1a, among children from families with insecure and negative relationships, emotional problems decreased significantly and linearly from baseline through post-intervention to follow-up in the

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

intervention group. Among children from families with discrepant experiences, emotional problems decreased from baseline to post-intervention, but they then increased by follow-up, which indicates a quadratic change.

Further, in accordance with compensation hypothesis, children from families with insecure and negative relationships and discrepant experiences reported an increase in positive resources from baseline to post-intervention in the intervention group, as illustrated in Figure 2a. However, positive resources then decreased by follow-up, which indicates quadratic change.

Buffering hypothesis was substantiated concerning emotional problems, but not positive resources. Among children from families with secure and positive relationships emotional problems decreased linearly in the intervention group. Further, emotional problems decreased also in control group from baseline to post-intervention among children from families with secure and positive relationships and with moderately secure and neutral relationships.

Table 2 also shows repeated-measure main effects of time and intervention on mental health problems and psychosocial resources. Significant time main effects indicate that PTSD symptoms and emotional problems decreased linearly, whereas conduct problems decreased quadratically from baseline through post-intervention into follow-up. The general change in terms of positive resources and prosocial behavior was not significant, although the within-subject contrasts were significant. Significant intervention main effects indicated that PTSD symptoms decreased linearly and prosocial behavior increased linearly, especially in the intervention group.

The family types differed significantly (between-subject effects) in children's general levels of emotional problems ($F(3,244)=9.23, p<.0001$) and conduct problems ($F(3,244)=11.52, p<.0001$), as well as in positive resources ($F(3,244)=16.23, p<.0001$), and prosocial behavior ($F(3,244)=8.07, p<.0001$). Children from the families with secure and positive relationships and moderately secure and neutral relationships showed the lowest levels of emotional and conduct problems, while those in the families with insecure and negative relationships and discrepant experiences showed the highest levels. Meanwhile, children in the families with secure and positive relations and moderately secure and neutral relations showed the highest levels of positive resources and prosocial behavior, while those with insecure and negative relations and discrepant experiences showed the lowest.

Child gender was a significant covariate, but the effect of interaction between gender and family type, when added in the within-measure MANCOVA analysis, was nonsignificant. Traumatic war events were a nonsignificant covariate, which indicates that they did not affect the family type and intervention-related mental health changes.

DISCUSSION

The present study analyzed how dynamic family relationships influence the effectiveness of psychosocial interventions in improving children's mental health after experiencing war in the Palestinian context. Research on traumatized children emphasizes the importance of families as helping agents (Betancourt et al. 2013; Panter-Brick et al. 2014), and secure and supportive relationships are known to contribute to good mental health and even to function as buffers against the negative effects of war (Barber et al. 2005; Eltanamly et al. 2021; Montgomery 2011). However, few studies have examined the moderating impact of attachment, parenting, or sibling relationships on the effectiveness of help provided to war-traumatized children. In accordance with family systems theories, we employed a person-oriented approach to depict the complex dynamics of these relationships (Bergman et al. 2003). Family system dynamics also provided hypotheses about how and why children living in unique family types may benefit differently from psychosocial interventions, here the TRT, dedicated to improving their mental health. Our results supported the compensation hypothesis of psychosocial intervention, as emotional problems decreased and positive resources increased, especially among children from families with insecure and negative relationships and discrepant experiences. The buffering hypothesis was also substantiated, as children from families with secure and positive relationships showed decreased emotional problems both in the intervention and control groups. The hypothesis of accumulation dynamics among children with insecure and negative family relationships was not supported, as the intervention effects were not significant among them.

WHY COMPENSATION IN WAR?

Children from different family types have highly unequal access to social and emotional resources. In families with insecure and negative relationships, children are deprived of familial safety, trust, and warmth, which is especially harmful when experiencing traumatic war events. These children often lack parental support, stabilization, and consoling when afraid and worried. They also have fewer opportunities to learn the optimal cognitive and emotional responses

needed to manage hardships (Meiser-Stedman 2002), or to share and vent about painful and shameful experiences (Punamäki 2014). Children living in families with discrepant experiences may be more likely to experience a sense of isolation and mistrust, as well as a lack of affiliation and sharing with other family members (Lindblom et al. 2014). Therefore, the core elements of psychosocial interventions, including the TRT, must meet and satisfy the security and survival needs of children whose family relationships are insecure, negative, and discrepant. In other words, interventions can compensate by providing children with experiences that are not possible in their families.

In the TRT, children learn to create a safe place and imaginary helpers, and to enjoy playful ways of learning emotion recognition, regulation, and expression. They gradually learn to manage their overwhelming trauma-related fears and anxieties, and to employ effective coping strategies in multimodal ways through visual, auditory, kinaesthetic, and symbolic activities (Smith et al. 2000). Children can learn to trust their peers through group processes and enjoy the emotional availability of the adult facilitators. These compensation dynamics may also be attributed to feelings of surprise, joy, and empowerment, which explains the effective mental health recovery of children who are habituated to insecurity and disappointing social experiences. In interviews with the participating children, some said they were amazed at how much strength, sense of belonging, and power they could find in themselves. Similar dynamics were documented among torture survivors with insecure attachment styles, who benefitted from therapy treatments once they were able to trust in the benevolence of others (Kanninen, Salo, and Punamäki 2000).

Research has found analogous compensation dynamics around early interventions, biological vulnerability, and adolescents' search for safety. There is evidence that small children from socially deprived environments (e.g., low socioeconomic standing, overcrowded residences, and substance use) especially benefit from good-quality day care (Carneiro and Ginja 2014). Infant-related stressors, such as low birth weight and developmental deficits, can intensify a mother's efforts to provide compensatory and sensitive care, thereby supporting optimal child development (Korja, Latva, and Lehtonen 2012). Compensation dynamics also emerge in conditions where adolescents who have insecure attachment to their parents seek to create secure attachments to their peers and friends, which can result in good mental health (Helsen, Vollebergh, and Meeus 2000; Kobak et al. 2007).

We could not find empirical evidence of psychosocial interventions or therapies that work especially well among participants with low resources, which indicates compensation dynamics. The available studies on the preconditions (moderators) of psychosocial interventions among war-traumatized children emphasize their vulnerabilities when explaining the ineffectiveness (Brown et al. 2017), thus implicitly supporting the accumulation hypothesis, which suggests that children who lack social resources are less able to benefit from enriching experiences. Yet, many interventions and therapies for traumatized people (especially those based on attachment theories) explicitly aim at compensating for insecure, neglected, and abusive experiences by providing a safe haven, empowerment, and trust in therapeutic relationships (Kinniburgh et al. 2005), and by improving parental sensitivity and competence and family relations (Johnson et al. 2018).

The buffering hypothesis was substantiated concerning children from families with highly secure and positive relationships. Their emotional symptoms decreased not only in the intervention, but also in the control group, thus suggesting that children in secure families can improve their good mental health without the help of the psychosocial intervention. In other words, these children were not in urgent need of psychosocial help, as time and life itself were healing them. This result emphasizes the importance of tailoring interventions to children's comprehensive needs, including a sense of security in human relationships. War-affected children should be screened for mental health problems and, according to the triaged model of services (Murray and Jordans 2016), they can be assigned to different interventions based on the severity of their problems and diagnoses. The present study suggests that the screening and treatment assignment could be based on both children's psychosocial vulnerabilities and resources.

The participants in our study have had long-term and multiple exposure to severe war events, such as human and material losses, horrors, and threats to life. Personal exposure was not, however, a significant covariate in the role of family type in the effectiveness analysis. It is possible that high-quality family relationships can decisively enhance children's recovery from traumatic war events, as these families invest greatly in protecting their children during wartime (Panter-Brick et al. 2014; Thabet et al. 2009). However, correlational analysis revealed that a high level of children's traumatic war experiences was significantly associated with a high degree of mental health problems and low psychosocial resources across all assessment points (except PTSD symptoms at follow-up). Traumatic experiences thus constitute a severe threat to children's wellbeing and psychosocial adjustment in war conditions; fortunately, they did not influence the effectiveness of the intervention.

STUDY LIMITATIONS AND STRENGTHS

The study deserves criticism for its reliance on self-reports, the use of a cluster analysis method, and the negatively oriented parenting measure. We had to use self-reports of attachment styles due to our limited budget and relatively large sample. To study parents' attachment, a high standard could theoretically be the adult attachment interview, which provides dynamic dimensions of coherence or an unresolved traumatic past (Bakermans-Kranenburg and van IJzendoorn 2009). The criteria for intervention effectiveness (children's PTSD, emotional and conduct problems, and positive resources and prosocial behavior) were also based on self-reports. Clinical interviews on mental health problems would have been more accurate and insightful than their own responses, and parent and teacher observations about children's behavior are considered reliable. We used cluster analyses to identify the family types, which is based on the similarities of family members' responses. Latent profile modeling provides goodness-of-fit indices for latent class membership, which might strengthen the determination of cluster numbers and sizes (McCutcheon 2002; Tabachnick and Fidell 2013). Finally, the parenting quality measure was focused on highly negative aspects of parenting, including harsh practices and neglect. This choice was based on earlier studies on the transgenerational transmission of trauma (Yehuda, Halligan, and Grossman 2001); including supportive and loving parenting styles would have provided more information and validity.

The study contributes to family systems and trauma research by identifying dynamic family groups based on relevant relationships in attachment, parenting, and siblingship subsystems. The person-oriented family approach enabled us to obtain insightful information about the relationships between civilian mothers, fathers, children, and siblings in the highly stressful and traumatic sociopolitical context of war, military violence, and the struggle for national independence. A family systems approach is considered important but is seldom studied empirically. The dynamic approach is also appropriate, as a large number of families have been forced to flee from their homes due to war, persecution, and human rights abuses and to seek safety and refuge in more peaceful countries.

CONCLUSIONS

The political-military situation involving siege, occupation, and international boycott is depriving Palestinian children of governmental protection, safety, and basic human rights. Families carry the heavy burden of shielding their children from insecurity, and from the detrimental effects of war and violence. Psychosocial interventions, including the TRT, can help to maintain children's positive mental health; this effect may be even more pronounced when families struggle with overwhelming insecurity.

REFERENCES

- Al-Hassan, Suha, and Hanan Takash. 2011. "Attributions and Attitudes of Mothers and Fathers in Jordan." *Parenting* 11 (2-3): 142-51. <https://doi.org/10.1080/15295192.2011.585559>.
- Al-Krenawi, Alean, and John R. Graham. 2005. "Marital Therapy for Arab Muslim Palestinian Couples in the Context of Reacculturation." *The Family Journal* 13 (3): 300-10. <https://doi.org/10.1177/1066480704273640>.
- Altawil, Mohamad A. S., Aiman El Asam, and Ameerah Khadaroo. 2018. "The Effectiveness of Therapeutic and Psychological Intervention Programs in PTC-GAZA." *Journal of Child & Adolescence Trauma* 11: 473-86. <https://doi.org/10.1007/s40653-018-0213-0>.
- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Arlington, VA: American Psychiatric Association.
- Bakermans-Kranenburg, Marian J., and Marinus H. van IJzendoorn. 2009. "The First 10,000 Adult Attachment Interviews: Distributions of Adult Attachment Representations in Clinical and Non-Clinical Groups." *Attachment & Human Development* 11 (3): 223-63. <http://doi.org/10.1080/14616730902814762>.
- Barber, Brian K., Heidi E. Stoltz, and Joseph A. Olsen. 2005. "Parental Support, Psychological Control, and Behavioral Control: Assessing Relevance across Time, Culture, and Method." *Monographs of the Society for Research in Child Development* 70 (4): i-147. <https://doi.org/10.1111/j.1540-5834.2005.00365.x>.

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

- Bergman, Lars R., David Magnusson, and Bassam M. El Khouri. 2003. *Studying Individual Development in an Interindividual Context: A Person-Oriented Approach*, Vol. 4. London: Psychology Press.
- Betancourt, Theresa S., Carmel Salhi, Stephen Buka, Jennifer Leaning, Gillian Dunn, and Felton Earls. 2012. "Connectedness, Social Support and Internalising Emotional and Behavioural Problems in Adolescents Displaced by the Chechen Conflict." *Disasters* 36 (4): 635-55. <https://doi.org/10.1111/j.1467-7717.2012.01280.x>.
- Betancourt, Theresa S., Sarah E. Meyers-Ohki, Alexandra P. Charrow, and Wietse A. Tol. 2013. "Interventions for Children Affected by War: An Ecological Perspective on Psychosocial Support and Mental Health Care." *Harvard Review of Psychiatry* 21 (2): 70-91. <https://doi.org/10.1097/HRP.0b013e318283bf8f>.
- Bonanno, George A. 2004. "Loss, Trauma, and Human Resilience: Have We Underestimated the Human Capacity to Thrive after Extremely Aversive Events?" *American Psychologist* 59 (1): 20-28. <https://doi.org/10.1037/0003-066x.59.1.20>.
- Bosqui, Tania Josiane, and Bassam Marshoud. 2018. "Mechanisms of Change for Interventions Aimed at Improving the Wellbeing, Mental Health and Resilience of Children and Adolescents Affected by War and Armed Conflict: A Systematic Review of Reviews." *Conflict and Health* 12 (1): 1-17. <https://doi.org/10.1186/s13031-018-0153-1>.
- Bowlby, John. 1969. *Attachment and Loss, Vol. 1: Attachment*. London: The Tavistock Institute of Human Relations.
- Brown, R. C., A. Witt, J. M. Fegert, F. Keller, M. Rassenhofer, and P. L. Plener. 2017. "Psychosocial Interventions for Children and Adolescents after Man-Made and Natural Disasters: A Meta-Analysis and Systematic Review." *Psychological Medicine* 47 (11): 1893-1905. <https://doi.org/10.1017/S0033291717000496>.
- Bryant, Richard A. 2016. "Social Attachments and Traumatic Stress." *European Journal of Psychotraumatology* 7 (1): 29065. <https://doi.org/10.3402/ejpt.v7.29065>.

- Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6 (4): 135-73. <https://doi.org/10.1257/pol.6.4.135>.
- Charlson, Fiona, Mark van Ommeren, Abraham Flaxman, Joseph Cornett, Harvey Whiteford, and Shekhar Saxena. 2019. "New WHO Prevalence Estimates of Mental Disorders in Conflict Settings: A Systematic Review and Meta-Analysis." *The Lancet* 394 (10194): 240-48. [https://doi.org/10.1016/S0140-6736\(19\)30934-1](https://doi.org/10.1016/S0140-6736(19)30934-1).
- Cohen, Judith A., Anthony P. Mannarino, and Laura K. Murray. 2011. "Trauma-Focused CBT for Youth Who Experience Ongoing Traumas." *Child Abuse & Neglect* 35 (8): 637-46. <https://doi.org/10.1016/j.chiabu.2011.05.002>.
- Coyne, James C., Geraldine Downey, and Julia Boergers. 1992. "Depression in Families: A Systems Perspective." In *Rochester Symposium on Developmental Psychopathology: Developmental Perspectives on Depression*, edited by Dante Cicchetti and Sheree L. Toth, 211-49. Rochester, NY: University of Rochester Press.
- Crittenden, Patricia McKinsey, and Rudi Dallos. 2009. "All in the Family: Integrating Attachment and Family Systems Theories." *Clinical Child Psychology and Psychiatry* 14 (3): 389-409. <https://doi.org/10.1177/1359104509104048>.
- Diab, Marwan, Raija-Leena Punamäki, Esa Palosaari, and Samir R. Qouta. 2014. "Can Psychosocial Intervention Improve Peer and Sibling Relations among War-Affected Children? Impact and Mediating Analyses in a Randomized Controlled Trial." *Social Development* 23 (2): 215-31. <https://doi.org/10.1111/sode.12052>.
- Diab, Safwat Y. M. 2011. *The Psychosocial Factors That Undermine Children's Academic Potentials: A Multi-Level Study of the Main Psychosocial Factors That Contribute to Palestinian Children's Academic Underachievement*. Berlin: Lambert Academic Publishing.
- Diab, Safwat Y. M., Marja Guillaume, and Raija-Leena Punamäki. 2018. "Ecological Approach to Academic Achievement in War and Military Violence: Mediating and Moderating Role of School, Family, and Child Characteristics." *The Elementary School Journal* 119 (1) :1-28. <https://doi.org/10.1086/698730>.

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

- Dunn, Judy, Cheryl Slomkowski, Lynn Beardsall, and Richard Rende. 1994. "Adjustment in Middle Childhood and Early Adolescence: Links with Earlier and Contemporary Sibling Relationships." *Journal of Child Psychology and Psychiatry* 35 (3): 491-504. <https://doi.org/10.1111/j.1469-7610.1994.tb01736.x>.
- Dyregrov, Atle, Rolf Gjestad, and Magne Raundalen. 2002. "Children Exposed to Warfare: A Longitudinal Study." *Journal of Traumatic Stress* 15 (1): 59-68. <https://doi.org/10.1023/A:1014335312219>.
- El-Khodary, Basel, and Muthanna Samara. 2020. "The Relationship between Multiple Exposures to Violence and War Trauma, and Mental Health and Behavioural Problems among Palestinian Children and Adolescents." *European Child & Adolescent Psychiatry* 29 (5): 719-31. <https://doi.org/10.1007/s00787-019-01376-8>.
- Eloranta, Sami J., Kirsi Peltonen, Esa Palosaari, Samir R. Qouta, and Raija-Leena Punamäki. 2017. "The Role of Attachment and Emotion Regulation in the Psychosocial Intervention among War-Affected Children." *Journal of Child & Adolescent Trauma* 10 (4): 301-14. <https://doi.org/10.1007/s40653-016-0115-y>.
- Eltanamy, Hend, Patty Leijten, Suzanne Jak, and Geertjan Overbeek. 2021. "Parenting in Times of War: A Meta-Analysis and Qualitative Synthesis of War Exposure, Parenting, and Child Adjustment." *Trauma, Violence, & Abuse* 22 (1): 147-60. <https://doi.org/10.1177/1524838019833001>.
- Feldman, Ruth, Adva Vengrober, Moranne Eidelman-Rothman, and Orna Zagoory-Sharon. 2013. "Stress Reactivity in War-Exposed Young Children with and without Posttraumatic Stress Disorder: Relations to Maternal Stress Hormones, Parenting, and Child Emotionality and Regulation." *Development and Psychopathology* 25 (4): 943-55. <https://doi.org/10.1017/S0954579413000291>.
- Goodman, Robert, Howard Meltzer, and Veira Bailey. 1998. "The Strengths and Difficulties Questionnaire: A Pilot Study on the Validity of the Self-Report Version." *European Child Adolescent Psychiatry* 7 (3): 125-30. <https://doi.org/10.1007/s007870050057>.
- Finnegan, Regina A., Ernest V. E. Hodges, and David G. Perry. 1996. "Preoccupied and Avoidant Coping during Middle Childhood." *Child Development* 67 (4): 1318-28. <https://doi.org/10.2307/1131702>.

- Harter, Susan. 1982. "The Perceived Competence Scale for Children." *Child Development* 53 (1): 87-97. <https://doi.org/10.2307/1129640>.
- Helsen, Marianne, Wilma Vollebergh, and Wim Meeus. 2000. "Social Support from Parents and Friends and Emotional Problems in Adolescence." *Journal of Youth and Adolescence* 29 (3): 319-35. <https://doi.org/10.1023/A:1005147708827>.
- Johnson, Sarah Lindstrom, Kit Elam, Adam A. Rogers, and Chanler Hilley. 2018. "A Meta-Analysis of Parenting Practices and Child Psychosocial Outcomes in Trauma-Informed Parenting Interventions after Violence Exposure." *Prevention Science* 19 (7): 927-38. <https://doi.org/10.1007/s11121-018-0943-0>.
- Jordans, Mark J. D., Hugo Pigott, and Wietse A. Tol. 2016. "Interventions for Children Affected by Armed Conflict: A Systematic Review of Mental Health and Psychosocial Support in Low- and Middle-Income Countries." *Current Psychiatry Reports* 18 (9): 1-15. <https://doi.org/10.1007/s11920-015-0648-z>.
- Kagitcibasi, Cigdem. 2005. "Autonomy and Relatedness in Cultural Context: Implications for Self and Family." *Journal of Cross-Cultural Psychology* 36 (4): 403-22. <https://doi.org/10.1177/0022022105275959>.
- Kangaslampi, Samuli, and Kirsi Peltonen. 2019. "Mechanisms of Change in Psychological Interventions for Posttraumatic Stress Symptoms: A Systematic Review with Recommendations." *Current Psychology* (2019): 1-18. <https://doi.org/10.1007/s12144-019-00478-5>.
- Kankaanpää, Saija L., Sanna Isosävi, Safwat Y. M. Diab, Samir R. Quota, and Raija-Leena Punamäki. 2020. "Trauma and Parenting Beliefs: Exploring the Ethnotheories and Socialization Goals of Palestinian Mothers." *Child and Family Studies* 29 (8): 2130-45. <https://doi.org/10.1007/s10826-020-01746-w>.
- Kanninen, Katri, Jari Salo, and Raija-Leena Punamäki. 2000. "Attachment Patterns and Working Alliance in Trauma Therapy for Victims of Political Violence." *Psychotherapy Research* 10 (4): 435-49. <https://doi.org/10.1093/ptr/10.4.435>.
- Keresteš, Gordana. 2006. "Children's Aggressive and Prosocial Behavior in Relation to War Exposure: Testing the Role of Perceived Parenting and Child's Gender." *International Journal of Behavioral Development* 30 (3): 227-39. <https://doi.org/10.1177/0165025406066756>.

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

- Kerns, Kathryn A., Lisa Klepac, and AmyKay Cole. 1996. "Peer Relationships and Preadolescents' Perceptions of Security in the Child-Mother Relationship." *Developmental Psychology* 32 (3): 457-66. <https://doi.org/10.1037/0012-1649.32.3.457>.
- Kerns, Kathryn A., Patricia L. Tomich, Jeffery E. Aspelmeier, and Josefina M. Contreras. 2000. "Attachment-Based Assessments of Parent-Child Relationships in Middle Childhood." *Developmental Psychology* 36 (5): 614-26. <https://doi.org/10.1037/0012-1649.36.5.614>.
- Keyes, Corey L. M., Marié Wissing, Johan P. Potgieter, Michael Temane, Annamari Kruger, and Sinette van Rooy. 2008. "Evaluation of the Mental Health Continuum-Short Form (MHC-SF) in Setswana-Speaking South Africans." *Clinical Psychology & Psychotherapy* 15 (3): 181-92. <https://doi.org/10.1002/Cpp.572>.
- Khamis, Vivian. 1998. "Psychological Distress and Well-Being among Traumatized Palestinian Women during the Intifada." *Social Science & Medicine* 46 (8): 1033-41. [https://doi.org/10.1016/S0277-9536\(97\)10032-6](https://doi.org/10.1016/S0277-9536(97)10032-6).
- Khamis, Vivian. 2000. "Child Psychological Maltreatment in Palestinian Families." *Child Abuse & Neglect* 24 (8): 1047-59. [https://doi.org/10.1016/S0145-2134\(00\)00157-5](https://doi.org/10.1016/S0145-2134(00)00157-5).
- Kinniburgh, Kristine Jentoft, Margaret Blaustein, Joseph Spinazzola, and Bessel A. van der Kolk. 2005. "Attachment, Self-Regulation, and Competency: A Comprehensive Intervention Framework for Children with Complex Trauma." *Psychiatric Annals* 35 (5): 424-30. <https://doi.org/10.3928/00485713-20050501-08>.
- Kobak, Roger, Natalie L. Rosenthal, Kristyn Zajac, and Stefanie D. Madsen. 2007. "Adolescent Attachment Hierarchies and the Search for an Adult Pair-Bond." *New Directions for Child and Adolescent Development* 2007 (117): 57-72. <https://doi.org/10.1002/cd.194>.
- Kolltveit, Silje, Ida Ingridsdatter Lange-Nielsen, Abdel Aziz Mousa Thabet, Atle Dyregrov, Ståle Pallesen, Tom Backer Johnsen, and Jon Christian Laberg. 2012. "Risk Factors for PTSD, Anxiety, and Depression among Adolescents in Gaza." *Journal of Traumatic Stress* 25 (2): 164-70. <https://doi.org/10.1002/jts.21680>.

- Korja, Riikka, Reija Latva, and Liisa Lehtonen. 2012. "The Effects of Preterm Birth on Mother-Infant Interaction and Attachment during the Infant's First Two Years." *Acta Obstetrica et Gynecologica Scandinavica* 91 (2): 164-73. <https://doi.org/10.1111/j.1600-0412.2011.01304.x>.
- Lindblom, Jallu, Marjo Flykt, Asko Tolvanen, Mervi Vänskä, Aila Tiitinen, Maija Tulppala, and Raija-Leena Punamäki. 2014. "Dynamic Family System Trajectories from Pregnancy to Child's First Year." *Journal of Marriage and Family* 76 (4): 796-807. <https://doi.org/10.1111/jomf.12128>.
- Lindblom, Jallu, Mervi Vänskä, Marjo Flykt, Asko Tolvanen, Aila Tiitinen, Maija Tulppala, and Raija-Leena Punamäki. 2017. "From Early Family Systems to Internalizing Symptoms: The Role of Emotion Regulation and Peer Relations." *Journal of Family Psychology* 31 (3): 316-26. <https://doi.org/10.1037/fam0000260>.
- Masten, Ann S., and Amy R. Monn. 2015. "Child and Family Resilience: A Call for Integrated Science, Practice, and Professional Training." *Family Relations* 64 (1): 5-21. <https://doi.org/10.1111/fare.12103>.
- Masten, Ann S., and Angela J. Narayan. 2012. "Child Development in the Context of Disaster, War, and Terrorism: Pathways of Risk and Resilience." *Annual Review of Psychology* 63: 227-57. <https://doi.org/10.1146/annurev-psych-120710-100356>.
- McCutcheon, Allan L. 2002. "Basic Concepts and Procedures in Single- and Multiple-Group Latent Class Analysis." In *Applied Latent Class Analysis*, edited by Jacques A. Hagenaars and Allan L. McCutcheon, 56-86. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511499531.003>.
- Meiser-Stedman, Richard. 2002. "Towards a Cognitive-Behavioral Model of PTSD in Children and Adolescents." *Clinical Child and Family Psychology Review* 5 (4): 217-32. <https://doi.org/10.1023/A:1020982122107>.
- Mikulincer, Mario, Netta Horesh, Ilana Eilati, and Moshe Kotler. 1999. "The Association between Adult Attachment Style and Mental Health in Extreme Life-Endangering Conditions." *Personality and Individual Differences* 27 (5): 831-42. [https://doi.org/10.1016/S0191-8869\(99\)00032-X](https://doi.org/10.1016/S0191-8869(99)00032-X).

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

- Minuchin, Salvador. 1974. *Families and Family Therapy*. Cambridge, MA: Harvard University Press.
- Montgomery, Edith. 2004. "Tortured Families: A Coordinated Management of Meaning Analysis." *Family Process* 43 (3): 349-71. <https://doi.org/10.1111/j.1545-5300.2004.00027.x>.
- Montgomery, Edith. 2011. "Trauma, Exile and Mental Health in Young Refugees." *Acta Psychiatrica Scandinavica* 124: 1-46. <https://doi.org/10.1111/j.1600-0447.2011.01740.x>.
- Murray, Laura K., and Mark J. D. Jordans. 2016. "Rethinking the Service Delivery System of Psychological Interventions in Low and Middle Income Countries." *BMC Psychiatry* 16 (1): 1-6. <https://doi.org/10.1186/s12888-016-0938-y>.
- Panter-Brick, Catherine, Marie-Pascale Grimon, and Mark Eggerman. 2014. "Caregiver-Child Mental Health: A Prospective Study in Conflict and Refugee Settings." *Journal of Child Psychology and Psychiatry* 55 (4): 313-27. <https://doi.org/10.1111/jcpp.12167>.
- Peltonen, Kirsi, Samir R. Qouta, Eyad El Sarraj, and Raija-Leena Punamäki. 2010. "Military Trauma and Social Development: The Moderating and Mediating Roles of Peer and Sibling Relations in Mental Health." *International Journal of Behavioral Development* 34 (6): 554-63. <https://doi.org/10.1177/0165025410368943>.
- Punamäki, Raija-Leena. 1986. "Stress among Palestinian Women under Military Occupation; Women's Appraisal of Stressors, Their Coping Modes, and Their Mental Health." *International Journal of Psychology* 21 (1-4): 445-62. <https://doi.org/10.1080/00207598608247601>.
- Punamäki, Raija-Leena. 2014. "Mental Health and Development among Children Living in Violent Conditions: Underlying Mechanisms for Promoting Peace." In *Pathways to Peace: The Transformative Power of Children and Families*, edited by James F. Leckman, Catherine Panter-Brick, and Rima Salah, 213-32. Cambridge, MA: MIT Press.

- Punamäki, Raija-Leena, Samir R. Qouta, and Eyad El-Sarraj. 2001. "Resiliency Factors Predicting Psychological Adjustment after Political Violence among Palestinian Children." *International Journal of Behavioral Development* 25 (3): 256-67. <https://doi.org/10.1080/01650250042000294>.
- Punamäki, Raija-Leena, Samir R. Qouta, Eyad El-Sarraj, and Edith Montgomery. 2006. "Psychological Distress and Resources among Siblings and Parents Exposed to Traumatic Events." *International Journal of Behavioral Development* 30 (5): 385-97. <https://doi.org/10.1177/0165025406066743>.
- Punamäki, Raija-Leena, Samir R. Qouta, and Kirsi Peltonen. 2017. "Family Systems Approach to Attachment Relations, War Trauma, and Mental Health among Palestinian Children and Parents." *European Journal of Psychotraumatology* 8 (sup7): 1-16. <https://doi.org/10.1080/20008198.2018.1439649>.
- Purgato, Marianna, Federico Tedeschi, Theresa S. Betancourt, Paul Bolton, Chiara Bonetto, Chiara Gastaldon, James Gordon et al. 2020. "Mediators of Focused Psychosocial Support Interventions for Children in Low-Resource Humanitarian Settings: Analysis from an Individual Participant Dataset with 3,143 Participants." *Journal of Child Psychology and Psychiatry* 61 (5): 584-93. <https://doi.org/10.1111/jcpp.13151>.
- Qouta, Samir R., Raija-Leena Punamäki, Thomas Miller, and Eyad El-Sarraj. 2008. "Does War Beget Child Aggression? Military Violence, Gender, Age and Aggressive Behavior in Two Palestinian Samples." *Aggressive Behavior: Official Journal of the International Society for Research on Aggression* 34 (3): 231-44. <https://doi.org/10.1002/ab.20236>.
- Riggs, Shelley A., and David S. Riggs. 2011. "Risk and Resilience in Military Families Experiencing Deployment: The Role of the Family Attachment Network." *Journal of Family Psychology* 25 (5): 675-87. <https://doi.org/10.1037/a0025286>.
- Schnyder, Ulrich, Anke Ehlers, Thomas Elbert, Edna B. Foa, Berthold P. R. Gersons, Patricia A. Resick, Francine Shapiro, and Marylène Cloitre. 2015. "Psychotherapies for PTSD: What Do They Have in Common?" *European Journal of Psychotraumatology* 6 (1): 1-10. <https://doi.org/10.3402/ejpt.v6.28186>.

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

- Slone, Michelle, and Shiri Mann. 2016. "Effects of War, Terrorism and Armed Conflict on Young Children: A Systematic Review." *Child Psychiatry & Human Development* 47 (6): 950-65. <https://doi.org/10.1007/s10578-016-0626-7>.
- Smith, Peter, Atle Dyregrov, and William Yule. 2000. *Children and War: Teaching Recovery Techniques*. Bergen, Norway: Foundation for Children and War.
- Tabachnick, Barbara G., and Linda S. Fidell. 2013. *Using Multivariate Statistics*, 6th ed. Northridge: California State University.
- Thabet, Abdel Azis, A. N. Ibraheem, R. Shivram, E. A. Winter, and Panos Vostanis. 2009. "Parenting Support and PTSD in Children of a War Zone." *International Journal of Social Psychiatry* 55 (3): 226-37. <https://doi.org/10.1177/0020764008096100>.
- Thabet, Abdel Azis, David Stretch, and Panos Vostanis. 2000. "Child Mental Health Problems in Arab Children: Application of the Strengths and Difficulties Questionnaire." *International Journal of Social Psychiatry* 46 (4): 266-80. <https://doi.org/10.1177/002076400004600404>.
- Tol, Wietse A., Suzan Song, and Mark J. D. Jordans. 2013. "Annual Research Review: Resilience and Mental Health in Children and Adolescents Living in Areas of Armed Conflict—A Systematic Review of Findings in Low- and Middle-Income Countries." *Journal of Child Psychology and Psychiatry* 54 (4): 445-60. <https://doi.org/10.1111/jcpp.12053>.
- Ukoumunne, Obioha C., Martin C. Gulliford, Susan Chinn, Jonathan A. C. Sterne, Peter G. J. Burney, and Allan Donner. 1999. "Evaluation of Health Interventions at Area and Organization Level." *British Medical Journal* 319: 376-79. <https://doi.org/10.1136/bmj.319.7206.376>.
- UNHRC (UN Human Rights Council). 2009. *Human Rights in Palestine and Other Occupied Arab Territories: Report of the United Nations Fact Finding Mission on the Gaza Conflict*. New York: UN General Assembly. https://www.un.org/ruleoflaw/files/UNFFMGC_Report.pdf.

- UN OCHA (UN Office for the Coordination of Humanitarian Affairs). 2009. *Locked In: The Humanitarian Impact of Two Years of Blockade on the Gaza Strip*. East Jerusalem: UN OCHA, Occupied Palestinian Territory. https://www.ochaopt.org/sites/default/files/Ocha_opt_Gaza_impact_of_two_years_of_blockade_August_2009_english.pdf.
- Vänskä, Mervi, Raija-Leena Punamäki, Jallu Lindblom, Asko Tolvanen, Marjo Flykt, Leila Unkila-Kallio, Maija Tulppala, and Aila Tiitinen. 2015. "Timing of Early Maternal Mental Health and Child Cortisol Regulation." *Infant and Child Development* 25 (6): 461-83. <https://doi.org/10.1002/icd.1948>.
- Veronese, Guido, Alessandro Pepe, Federica Cavazzoni, Hania Obaid, and Jesus Perez. 2019. "Agency via Life Satisfaction as a Protective Factor from Cumulative Trauma and Emotional Distress among Bedouin Children in Palestine." *Frontiers in Psychology* 10: 1674. <https://doi.org/10.3389/fpsyg.2019.01674>.
- World Bank. 2015. *Economic Monitoring Report to the Ad Hoc Liaison Committee*. Washington, DC: World Bank. <http://documents.worldbank.org/curated/en/563181468182960504/pdf/96601-REVISED-WP-Box391464B-AHLC-May-21-Book-fix-footnotes.pdf>.
- Yehuda, Rachel, Sarah L. Halligan, and Robert Grossman. 2001. "Childhood Trauma and Risk for PTSD: Relationship to Intergenerational Effects of Trauma, Parental PTSD, and Cortisol Excretion." *Development and Psychopathology* 13 (3): 733-53. <https://doi.org/10.1017/s0954579401003170>.

FAMILY RELATIONSHIPS AND PSYCHOSOCIAL
GROUP INTERVENTION EFFECTIVENESS

APPENDIX

Table A1: Means, Standard Deviations, and Bivariate Correlations between Children's Mental Health Problems and Psychosocial Resources and Covariates

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
<i>Pre-intervention^a</i>									
1. PTSD symptoms	29.53	10.09							
2. Emotional problems	4.08	2.11	.15**						
3. Conduct problems	2.53	2.05	.09*	.37***					
4. Positive resources	34.86	9.56	-.07	-.13*	-.19**				
5. Prosocial behavior	7.60	2.04	-.05	-.22**	-.48***	.27***			
<i>Post-intervention^a</i>									
6. PTSD symptoms	26.30	11.26	.21**	.27***	.07	-.04	-.09*		
7. Emotional problems	3.55	2.02	.12*	.31***	.19**	-.12*	-.17**	.44***	
8. Conduct problems	1.93	1.80	.08	.17**	.19**	-.08	-.16**	.11*	.31***
9. Positive resources	36.04	9.12	-.03	-.25***	-.26***	.34***	.21**	-.08	-.24**
10. Prosocial behavior	7.85	1.89	-.05	-.06	-.16**	.18**	.20**	-.02	-.15**
<i>Follow-up^b</i>									
11. PTSD symptoms	24.91	9.43	.14*	.18**	.12*	.02	-.09	.42***	.25**
12. Emotional problems	3.51	2.21	.20**	.30***	.29***	-.12*	-.17**	.38***	.47***
13. Conduct problems	2.35	1.87	.04	.23***	.25***	-.13*	-.13*	.09*	.21**
14. Positive resources	34.05	9.76	-.12*	-.15*	-.18**	.32***	.12*	-.11*	-.18**
15. Prosocial behavior	6.69	2.21	-.07	-.06	-.21***	.21***	.19**	-.04	-.16**
16. Traumatic war events	8.06	4.83	.09*	.10*	.13*	-.09*	-.09*	.09*	.18**
17. Gender			.01	-.02	.13*	-.13*	-.14*	-.01	-.02

Note: ^a N=377, ^b N=321.

* $p < .05$, ** $p < .01$, *** $p < .001$ or $p < .0001$

Variable	8	9	10	11	12	13	14	15	16
<i>Pre-intervention^a</i>									
1. PTSD symptoms									
2. Emotional problems									
3. Conduct problems									
4. Positive resources									
5. Prosocial behavior									
<i>Post-intervention^a</i>									
6. PTSD symptoms									
7. Emotional problems									
8. Conduct problems									
9. Positive resources	-.25**								
10. Prosocial behavior	-.41***	.27**							
<i>Follow-up^b</i>									
11. PTSD symptoms	.08	.01	-.02						
12. Emotional problems	.18**	-.22**	-.08	.35***					
13. Conduct problems	.29***	-.26***	-.26***	.08	.44***				
14. Positive resources	-.17**	.43***	.19**	.02	-.33***	-.39***			
15. Prosocial behavior	-.24***	.27***	.18**	-.03	-.28***	-.52***	.39***		
16. Traumatic war events	.18**	-.20**	-.10*	.06	.25***	-.19**	-.21**	-.17**	
17. Gender	.16**	-.13*	-.18**	-.22**	.04	.30***	-.21**	-.27***	.28***

USING A PARTICIPATORY APPROACH TO CREATE SEL PROGRAMMING: THE CASE OF *AHLAN SIMSIM*

SHANNA KOHN, KIM FOULDS, CHARLOTTE COLE,
MACKENZIE MATTHEWS, AND LAILA HUSSEIN

ABSTRACT

This paper highlights the use of a participatory, trauma-informed approach in the creation of Ahlan Simsim, a Sesame Street television program for the Middle East, and asserts the importance of using a participatory approach to designing culturally relevant social and emotional learning (SEL) content. Ahlan Simsim is a component of a larger initiative of the same name, which was created by Sesame Workshop and the International Rescue Committee and funded by the MacArthur and LEGO foundations. This program brings early learning and nurturing care to children and families affected by the Syrian crisis through a combination of mass media and direct service programming. In this article, we present a review of the research and consultations Sesame Workshop conducted with local communities and local child-development experts in Iraq, Jordan, Syria, and Lebanon from August to November 2018. Sesame Workshop's aim was to identify and refine the television program's focus area and to create locally relevant, trauma-informed content that draws from SEL strategies that resonate most and have the greatest impact with audiences in the Syrian response region. We argue that, for SEL programming to achieve maximum impact, it is critical that program designers develop social-emotional frameworks for children from the ground up by working with local caregivers and practitioners.

Received October 17, 2019; revised March 17, 2020, June 29, 2020, and August 3, 2020; accepted November 2, 2020; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2
Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).
ISSN 2518-6833

INTRODUCTION

The value of social and emotional learning (SEL) programming cannot be overstated. Research shows that SEL programs help children build skills that are critical to resilience, healing, and coping with long-term exposure to trauma (Umiltà et al. 2013) and can enhance children's academic outcomes by providing them with tools that help them focus, regulate their emotions, and cope with stress (Durlak et al. 2011). The importance of SEL interventions for children who have experienced violent conflict is particularly pronounced. Research on children in war zones has shown that exposure to traumatic events is strongly associated with physical and mental health outcomes, and that experiencing five or more traumatic events triples the risk of having psychiatric disorders and posttraumatic stress (Panter-Brick et al. 2009; Shonkoff et al. 2012). These effects are particularly acute in the first years of a child's life, when the brain is undergoing its most rapid development, making it extremely sensitive to environmental influence.

In order for SEL programming to resonate most fully with its beneficiaries and have the greatest impact, it is critical to take a participatory approach to the development of such programs (Cornwall 2002; Tufte and Mefalopulos 2009). The INEE Minimum Standards for Education dictate that community members should participate "actively, transparently and without discrimination" (INEE 2010, 22) in the design and analysis of education responses. The Nurturing Care Framework, an accepted approach to ECD in the humanitarian community, advises that, "when families feel valued, and when they are involved in [a] programme's design and delivery, they are likely to be more successful and to sustain their efforts" (World Health Organization, UNICEF, and World Bank 2018, 38).

Equally important to program design is acknowledging the effects of trauma. Studies show that the persistent stress common in humanitarian settings can affect children's outcomes later in life (Jabbar and Zaza 2014; Britto et al. 2016; Bouchane et al. 2018). Sesame Workshop employs what the US Department of Health and Human Services Substance Abuse and Mental Health Services Administration (2014) refers to as a trauma-informed approach that is aware of the widespread effects trauma has on children, recognizes the signs and symptoms, and responds by integrating knowledge of trauma into its program content and practices and actively resists retraumatization.

Despite evidence of the importance of SEL programming for young children affected by conflict and displacement, as well as recommendations for using a participatory, trauma-informed approach to program design and delivery, few programs for early childhood development (ECD) in emergencies focus on SEL, and even fewer engage local communities meaningfully in their content development (Measham et al. 2014). Short project timeframes (typically six months to one year), insufficient funds, and the emergency nature of the work contribute to this lack. However, while project timelines may be short, the impact of displacement is far from short term; refugees today are displaced on average for 10 to 26 years (World Bank 2016). Therefore, it is critical that funders of education in emergencies programs and the organizations delivering such programs recognize the detrimental effects conflict and displacement have on children's mental health and direct their resources toward culturally relevant SEL programming that resonates with its intended audience.

Using research conducted with local communities and experts to inform curriculum design and the production of educational television programs is key to Sesame Workshop's international coproduction model, as is seeking advice on program content from local child-development experts (Cole 2016). The founders of Sesame Workshop developed and led this approach, now known as the Sesame Workshop Model (Cooney 1966; Lesser 1974). In this article, we outline Sesame Workshop's process for adapting a participatory approach to creating television and digital programming in the context of the Syrian response. We illustrate how taking a participatory approach helped us identify SEL as a critical focus for *Ahlan Simsim* and helped us to create a culturally relevant, trauma-informed television program. We also highlight the importance of program developers checking their assumptions and developing SEL programs from the ground up in order to have the maximum impact.

OVERVIEW OF THE AHLAN SIMSIM PROGRAM

Nearly 71 million people are currently displaced worldwide, and nearly half of them are children. Some four million Syrian children in the Syrian response region—one-half of all Syrian children—have been born since the war began (UNICEF 2018). The healthy cognitive, social, and emotional development of displaced children living in Jordan, Lebanon, Iraq, and Syria is particularly at risk, due to the lack of access to early childhood care and education. These effects are not limited to children who have been displaced; children living in host communities, particularly in low-income areas and those with limited access

to ECD services, are also missing important opportunities. Children who have adverse experiences are at severe risk for impairments that will follow them throughout their lives, including poor physical and mental health, cognitive deficits, and reduced earnings (Felitti et al. 1998; Lahiri, van Ommeren, and Roberts 2017; Measham et al. 2014; Forman-Hoffman et al. 2013; Cook et al. 2017).

In response to this crisis, in December 2017, the MacArthur Foundation awarded Sesame Workshop and the International Rescue Committee (IRC) a 100&Change grant to bring early childhood care and education to millions of children living in the Syrian response region. This initiative, Ahlan Simsim, offers direct ECD services through classrooms, community centers, social-protection programs, health clinics, and other spaces frequented by children and families. These services are complemented by culturally relevant video content offered on broadcast and digital platforms. The program, which uses the Nurturing Care Framework, highlights two key pillars: offering children opportunities for early learning, and providing responsive caregiving (World Health Organization, UNICEF, and World Bank 2018). Ahlan Simsim also integrates the three other aspects of the Nurturing Care Framework—good health, adequate nutrition, and security and safety—into the IRC’s existing early childhood health and child-protection programs.

At the moment, less than 3 percent of the global humanitarian budget goes to education, and only a fraction of that goes to early childhood programming (UNESCO 2017). This means that there is insufficient support for the youngest children living in emergency contexts. Television programming, when carefully crafted to ensure its cultural relevance, has the ability to make up for this lack by presenting characters and scenarios that reflect these children’s lived experiences. This in turn helps to shift the knowledge, attitudes, and behaviors of children and families at a fraction of the cost of many humanitarian response models (Mares and Pan 2013). The IRC’s direct services programming, which takes a research-based approach to providing children and caregivers with facilitated instruction, has the ability to provide children and the adults who care for them with engaging, play-based programming that meets their specific needs. This combined approach to providing ECD services—using the reach and resonance of culturally relevant and trauma-informed mass media and the deep impact of direct service programming—holds great promise. Through the production of evidence-based, participatory models for ECD programming that combine the power of television and direct services and are operationally successful at scale, Ahlan Simsim offers a new model for ECD programming in humanitarian response settings.

SUPPORTING POSITIVE CHILDHOOD DEVELOPMENT THROUGH TELEVISION

A foundational element of Ahlan Simsim's work is the creation of *Ahlan Simsim*, an Arabic-language version of *Sesame Street* for children ages three to eight who are living in Syria, Iraq, Jordan, Lebanon, and the broader Middle East. Broadcast and digital platforms are formidable ways to deliver quality content at scale in the region, as household television access exceeds 90 percent in all four countries (Melki et al. 2012; Sweis and Baslan 2013; MICT 2014; Wagner, Glioti, and Hussein 2021).¹ One survey of Syrians living in their home country and Syrian refugees living in Jordan, Lebanon, and Turkey found that satellite television is almost ubiquitous—even in the refugee camps (MICT 2014). Mobile phone access is also high, exceeding 70 percent in Lebanon, 50 percent in Syria, 60 percent in Iraq, and nearly 70 percent in Jordan (GSMA 2019).

There are several factors beyond the benefits of scale and cost efficiency that make television programs, particularly Sesame Workshop programs, a unique platform for promoting children's positive outcomes. The first is their ability to promote representation. Because Sesame Workshop's model includes community-based research and the participation of local stakeholders, our programs introduce children to characters, scenes, and stories that reflect their own lives and experiences. Meaningful representation has been shown to have a powerful impact on children's self-esteem (Martins and Harrison 2012). Because of this connection, the program characters garner children's affection and empathy, which promotes recall of the educational messaging they promote. Second, television programs have the ability to influence behavior. In keeping with social learning theory (Bandura 1962) and the theoretical underpinning of entertainment education (Singhal and Rogers 1999), the program can support not just imitation but the application of much-needed skills. Finally, television has the ability to affect the multiple levels of influence in a child's life. Drawing from social and behavioral change communication concepts and grounded in Bronfenbrenner's (1979) social ecological model, Sesame Workshop's programs target individual children and also address the greater influences (e.g., family and peer networks, community, and social and cultural structures) that affect a child's growth and development. Media too are a part of a child's broader cultural ecosystem, and one strength of the Sesame Workshop model is its appeal to and ability to encourage the

1 While these citations refer to penetration rates before the conflict in Syria began, smaller-scale media landscape studies conducted by Sesame Workshop suggest that the rates remain similar.

adults in children's lives to adopt a model of responsive care. Since its inception, *Sesame Street's* creators have intentionally included celebrities, music, and humor to engage adults as well as children because they believed, accurately, that a child would learn more if an adult was engaged with them.

USE OF A PARTICIPATORY APPROACH IN THE CREATION OF *AHLAN SIMSIM*

As with all Sesame Workshop coproductions, the development of *Ahlan Simsim* began with research and consultation on the ground. Recognizing the important role caregivers and cultural context play in a child's early development (Bronfenbrenner 1979; Lagercrantz 2016), Sesame Workshop deliberately turned to caregivers, practitioners, and children living in the region to inform the curriculum design for season one. This participatory approach helped Sesame Workshop identify critical focus areas for *Ahlan Simsim* and helped us create a culturally appropriate and locally resonant product.

Below we report on our development process for *Ahlan Simsim*, and on key findings from each arm of the participatory approach: a needs assessment, consultation with expert advisors, work with online research communities, and a formative study of children's emotional literacy (see Appendix).² We used these studies to ensure that we included the voices of relevant communities in the program development and design.

NEEDS ASSESSMENT FINDINGS: THE NEED FOR SUPPORT WITH SEL SKILLS

Sesame Workshop begins the development of any new television production with a needs assessment to evaluate the existing gaps for young children in a given context, as articulated by their caregivers and practitioners (Foulds et al. 2021) We conducted the *Ahlan Simsim* needs assessment in Jordan, Lebanon, Iraq, and Syria in August 2018 to help us understand the joys and challenges of being a parent of a child three to eight years old in that region, to identify parents' trusted advisors, and to determine which issues they sought advice for. We also

² Our online research communities were a group of displaced Syrians and host community families living in Iraq, Jordan, and Lebanon, whom we invited to participate in quarterly 30-minute research panels covering a range of topics, such as various aspects of brand awareness, feedback on delivery mechanisms, and reactions to preliminary concepts under development for the television show. All participating families had consistent access to a smartphone and could participate in as many of the research panels as they liked. They received an honorarium for their participation to cover their mobile data usage and time. The Appendix includes an overview of the needs assessment, online research communities, and formative research we conducted. While a consultation with expert advisors was part of this participatory approach, we have limited our definition of research in this article to our engagement with *Ahlan Simsim's* direct users.

asked caregivers' and practitioners' opinion on potential focus topics for a new television show for children.³

In response to questions on parenthood and parenting challenges, caregivers reported that the best parts of being a parent were the feeling of parenthood, being responsible for children's wants and needs, giving and receiving love and compassion from their kids, and watching their children grow. The caregivers reported that the most difficult elements of being a parent were being unable to fulfill their children's needs because of financial issues, securing safe living conditions in current circumstances and/or an uncertain security situation, responsibility, single parenthood, and dealing with siblings' different personalities.

The caregivers most commonly reported seeking parenting guidance on children's health issues, discipline and managing children's behavior, parenting methods and childrearing, and schooling issues.

The caregivers were excited about the prospect of a new children's television show and expressed the need for a program that would reflect cultural values and teach children a range of skills, attitudes, and information. When asked what topics they would like a new television show for children to focus on, caregivers in the region said they wanted literacy and numeracy in Arabic and English, and a range of other skills. They emphasized the need for a particular focus on social-emotional skills, including respecting others, good manners, teamwork and cooperation, sharing and helping, expressing feelings and emotions, self-confidence, forgiveness, dealing with differences, honesty, and communication skills. Caregivers' articulation of the specific need for help in teaching social-emotional skills was likely connected to the increased stress of living in environments of conflict and crisis.

Practitioners' ideas about topics for a new children's television show echoed many of the caregivers' recommendations, although the practitioners prioritized social-emotional skills over academic skills. They recommended that the series focus on principles and values, health and hygiene, environmental conservation, critical thinking and decisionmaking, communication skills and expressing feelings and emotions, the importance of attending and staying in school, technology skills, social skills (e.g., mutual respect and understanding, self-confidence, self-regulation, and making friends), self-protection from sexual harassment, rejecting violence, and literacy, numeracy, and language arts (Arabic and English).

³ We use "parent" and "caregiver" interchangeably throughout this paper to refer to a child's primary caregiver.

The *Ahlan Simsim* needs assessment with caregivers and practitioners yielded this most important takeaway: children in the Syrian response region need SEL above all else. The needs assessment also highlighted the importance of considering the sensitivities of children who have experienced trauma when we created the content. With this in mind, we proceeded with the next stage of our development process—the educational content seminar—with a commitment to a trauma-informed approach and a sharper focus on SEL as the critical focus of our show.

FEEDBACK FROM EXPERT ADVISORS

In September 2018, following our decision to focus the *Ahlan Simsim* content on SEL, we refined our curricular approach by meeting with education advisors, including academics, practitioners, psychologists, and trauma, risk, and resilience specialists. These meetings, which took place in Amman, Jordan, and Beirut, Lebanon, included advisors from Syria and Iraq. They stated that identifying and regulating emotions was a skill gap among children three to eight years old who were living in Jordan, Lebanon, Syria, and Iraq—in particular among children who had experienced trauma.

We heard repeatedly from advisors that the emotional vocabularies of children in the region are often limited to two to three words to describe the broad range of human emotion. Beyond what we know from the research on exposure to traumatic events and the strong association with physical and mental health outcomes (Panter-Brick et al. 2009; Shonkoff et al. 2012), it is worth noting that limited engagement with the terminology of emotions may be a behavior learned from caregivers who experienced trauma during their own development.⁴ Children learn by observing and interacting with those around them, beginning at a very early age; consequently, their own ability to process and express emotions may be affected. We also heard that expressing emotion is typically limited to two behaviors: a deadpan or flat affect, and a physical, sometimes violent reaction, particularly among children who have experienced trauma (Cole et al. 2018).

The trauma, risk, and resilience specialists speaking at the workshops articulated specific considerations for children who have experienced certain adverse experiences, such as violence and displacement. For example, when we introduced one of the common strategies we use with populations in the United States who

4 The Lebanese Civil War (1975-1990), the conflict in Iraq (2003-present), and various conflicts involving Jordan over the last several decades have affected the lives of parents with children in our target age group, including the way they understand, respond to, and cope with feelings. Disassociation from intense emotions, particularly during periods of severe stress and trauma, is itself a coping strategy.

have experienced trauma—“Imagine a safe space”—the advisors pointed out that imagining a safe space could be retraumatizing for displaced children if the place they imagine is the home they have lost. They also explained that imagining a safe space requires a level of abstract thinking that is too advanced for children who have experienced developmental delays due to trauma. The advisors noted that trauma affects people of any age at a basic and sensory level and can also impact speech. They encouraged us to model nonverbal and sensory modes of expression, such as using music, colors, and shapes to portray various feelings. Finally, the specialists advised us not to underestimate the importance of giving the children a fun and entertaining respite from lives that for many of them are very stressful. Children need to play to learn, yet many children living in emergency contexts have limited opportunity to do so. They advised that a Sesame Street experience that gives children permission to be playful, joyful, and hopeful would do much to support their positive development.

THE IMPORTANCE OF IDENTIFYING AND REGULATING EMOTIONS

After receiving the advisors’ guidance, we returned to the caregivers to gain a better understanding of their ideas about the importance of identifying and regulating emotions. We also asked the caregivers to identify which terms for emotions were the easiest and most challenging for their children to comprehend, and which emotions children express most often.

The caregivers put a high priority on their children’s ability to understand and regulate their emotions. A large number of the caregivers reported that they believe in the power of television shows to support their children’s development because they said their children often imitate what they see on television. The caregivers also felt that an educational television show could give their children an opportunity to learn while also having fun.

When presented with a set of basic emotions (based on Shaver et al. 2001) and asked about the highest and lowest priority emotions for the show to address, the caregivers reported that

- the top three emotions their children were already familiar with, in descending order, were happiness, anger, and caring;
- the top two emotions that would be unfamiliar to their children, in descending order, were jealousy and awe; and

- the five emotions they wish their children had more knowledge about were anger, self-confidence, sadness, love for helping others, and courage.

Out of these findings, we formulated a curriculum for the first season of *Ahlan Simsim* that focused on defining and exploring nine primary emotions that were both familiar and unfamiliar to children, and which were specifically tied to the caregivers' requests (see Table 1). These findings helped us craft a curriculum that we believed would be meaningful and responsive to our audience's articulated needs.

FORMATIVE RESEARCH FINDINGS: HOW CHILDREN COMMUNICATE ABOUT THEIR FEELINGS

Simultaneous to the online research described in a footnote above and in the Appendix, we conducted individual interviews with caregiver-child dyads in Jordan and Lebanon to understand the terms children use to communicate their feelings. During the interviews, the researchers read Arabic-language children's storybooks aloud and stopped at certain points to ask the children questions related to the emotions displayed in that moment in the story. We found that their responses supported the caregivers', practitioners', and advisors' assertions that children struggle to find the emotional vocabulary needed to identify and regulate complicated feelings, such as anxiety, frustration, guilt, or jealousy. The children's articulated emotional range was limited, and they often resorted to phrases or actions to describe a feeling (see Kohn et al. 2020 for a full summary of the study results). The results of these interviews corroborated what we heard from expert advisors and from the online research surveys—that the children needed support to build a vocabulary of emotions in order to help them identify and regulate their emotions.

APPLICATION OF OUR FINDINGS: A NEW SEL-FOCUSED ARABIC-LANGUAGE SESAME STREET

The findings from our educational content seminar and the follow-up research studies corroborated the practitioners' and caregivers' assertions during the needs assessment—that is, that SEL is important for young children in the Syrian response context. Based on input from caregivers, educators, child-protection officers, social workers, and others who expressed the immense importance of this topic area, we focused *Ahlan Simsim* on SEL; season one focused specifically on identifying and regulating emotions. We focused on nine carefully selected emotions during the 26-episode season. Our selection was based

on recommendations from the caregivers about which emotional terms would be least familiar to their children and most important for them to learn (see Table 1). In response to the caregivers' call for help in equipping their children to handle or cope with their emotions, we also selected and introduced six coping strategies that we repeated throughout the season (see Table 2). We selected the strategies based on the advisors' recommendations for coping strategies that were culturally relevant, developmentally appropriate for early childhood, trauma informed, and relevant to the region. We presented the strategies through short, catchy songs that the children and caregivers could readily remember, and they were delivered by characters they connected with and cared about.

Table 1: Nine Emotions Presented in Season 1

-
1. **Anger:** غضب *Ghadab*
 2. **Caring:** مامتة! *Ihtimam*
 3. **Fear:** فوخ *Khowf*
 4. **Frustration:** طابح! *Ihbat*
 5. **Nervousness:** رتوت *Tawator*
 6. **Hope/Determination:** لمأ مي مصت *Amal/Tasmeem*
 7. **Jealousy:** غيرة *Gheerah*
 8. **Loneliness:** ودحو *Wihdah*
 9. **Sadness:** نوح *Hozon*
-

Table 2: Season 1 Strategy Bank

1. **Belly Breathe:** Put your hands on your belly and say “stop.” Take a deep, slow breath in through your nose while expanding your belly, then let it out through your mouth while letting your belly deflate.
 2. **Count to 5:** Take a deep breath and slowly count from one to five.
 3. **Move It Out:** Let your feelings out through physical movement: shake your hands, stomp your feet, and dance around to express how you feel.
 4. **Draw It Out:** Draw a picture of how you feel. Think about the color, shape, and texture of your feeling.
 5. **Ask for Help:** Talk to a trusted adult about how you feel. Ask for help and support.
 6. **Make a Plan:** First identify your goal. Next, identify the steps you need to take to get to your goal.
-

Arising from our research with local communities and the advice of local child-development experts on program content, the theory of change that supports our approach to the *Ahlan Simsim* show is based on the idea that learning to identify and constructively express emotions is a fundamental aspect of healthy development, resilience, and effective coping. Research shows that being better able to identify, express, and regulate emotions (important components of emotional and social competence) helps to reduce children’s internalizing and externalizing of behaviors and ultimately provides a foundation for their later ability to function across peer and school contexts (Izard et al. 2001). As is true of the acquisition of any skill, children need to learn the basics before they can master complexities. By focusing on labeling, identifying, and appropriately expressing a limited number of core emotions, *Ahlan Simsim* provides a tool for mastering a foundational element of broader SEL skills, such as perspective-taking and conflict resolution.

To support this approach, the structure, set, characters, and language of the show are also informed by research and guidance from local advisors, who emphasized that reflecting the rich diversity of the Middle East would be important to engaging children. The show's set—a middle-class home and garden that is a composite of typical homes found in Syria, Iraq, Jordan, and Lebanon—aims to reflect an environment that is realistic and relatable to a large cross-section of our viewers, which includes families of all income levels across the Middle East. Our choice of languages was also intentional. The show is produced primarily in Levantine Arabic and features characters who speak Syrian, Jordanian, Iraqi, and Lebanese dialects. We take special care to use words and phrases that have the broadest comprehension and shared meaning among people living in the four target countries, thereby creating, as closely as possible, what our language advisors refer to as a neutral dialect. To identify and express key curricular concepts, the characters use vocabulary in modern standard Arabic, which is familiar to and spoken by people throughout the region and is the language most children use in school. The variety show portion of *Ahlan Simsim* features adult and child guests from the broader Middle East, including Egypt, Saudi Arabia, and beyond, who speak in their native dialects. Feedback from language experts and local advisors assured us that featuring characters from our four target countries and the broader Middle East would reflect the reality of diaspora and help to promote exposure to and appreciation for the rich variety of cultures, countries, and dialects across the region, which will ultimately increase resonance and children's engagement.

The participatory approach we took in developing *Ahlan Simsim* revealed that children struggle to find the vocabulary to express their complex emotions and often resort to action terms, such as "I'm done." For this reason, we intentionally built definitions of emotion terms into the episodes. The focus emotions are defined at the open and close of every episode, and the program characters describe the physical manifestations of emotions throughout the programs. To enhance children's comprehension, we also incorporated visual supports associated with each emotion, such as using a frown to indicate sadness.

We learned that caregivers seek guidance on how to show empathy and responsiveness to their children's emotions. The show features two adult human characters, Teta Noor and Hadi, who are the caregivers for Basma and Jad, *Ahlan Simsim*'s two central characters. We heard repeatedly from the advisors that adults in the region also needed support in developing social-emotional skills, so we intentionally included caregivers in the show as a way to model healthy emotional expression and empathy for the caregivers in the region who view *Ahlan Simsim* with their children. We built in frameworks to guide caregivers in responding to

difficult emotional scenarios. For example, Teta Noor and Hadi modeled simple steps in how to respond to emotions, such as helping a child identify a feeling using facial expressions and physiological signs, helping them identify the reasons they are feeling as they do, validating their feelings, and helping them find an appropriate coping strategy.

Equally important was portraying child characters in a manner that gave them agency in identifying and coping with their emotions. For example, the episodes model children going to the adult characters and asking for assistance. This choice responded to our finding that children often do not know how to express or cope with their emotions in a healthy manner.

Finally, taking into consideration the findings from our needs assessment and the advisors' recommendations, we took several steps to ensure that the program's content was trauma informed. First, the coping strategies we selected as the focus of the first season avoided imagination exercises that required children to think of abstract, potentially retraumatizing concepts of safety. Our strategies focused instead on concrete concepts, such as counting to five, taking deep breaths, and moving one's body to work through an emotion. Second, we intentionally used nonverbal communication strategies to model expression. For example, main character Jad uses a paintbrush to express his feelings when he can't find the words for his ideas. Lastly, we took a play-based and humorous approach to the storylines, which gave children and their caregivers an opportunity to engage with the content in a joyful and positive manner.

Once rough cuts of the *Ahlan Simsim* episodes were complete, the researchers showed them to viewer focus groups that they had explicitly created to gain feedback from minority groups and the caregivers of children with disabilities about the coping strategies used, and on the storylines in general. IRC field staff in Iraq and Jordan conducted six focus groups. The sample of 37 female and male caregivers included Iraqis, Syrians, and Jordanians from host communities and camp settings; seven of them cared for children with disabilities. Our protocols focused on appeal, comprehension, and recommendations to improve the show, and the team incorporated their findings into ongoing revisions of the content as it was being finalized. This included tweaking the language to improve clarity, adding graphic elements to make the key emotion vocabulary words "pop," and adding a music score and sound effects throughout each episode to increase viewers' engagement.

CONCLUSION

Despite mounting evidence of the importance of SEL programming for young children affected by conflict and crisis, and the best practices espoused by the humanitarian aid community around using a participatory approach, few existing ECD programs in emergency settings focus on SEL, and even fewer use a participatory approach to content development. The development process for the *Ahlan Simsim* television production provides a case study in how to create contextually relevant, trauma-informed SEL content employing a participatory approach. Employing an approach that is foundational to the Sesame Workshop model (Cooney 1966; Lesser 1974), Sesame Workshop embarked on the development of *Ahlan Simsim* with the core objective of creating a program that was responsive to the articulated needs of children and families affected by conflict and displacement. Our needs assessment yielded the clear result that families needed help with their SEL skills, and that the program's content must be sensitive to the effects of persistent stress and trauma. Feedback from expert advisors during our educational content seminar helped us further narrow our focus on identifying and regulating emotions and trauma-informed content design. Additional findings from our online research communities and formative research helped us craft our educational framework—that is, the core emotions and self-regulation strategies that we repeated throughout the season. The final product is a culturally relevant SEL program specifically tailored to the needs of children and families in the target region.

Without meaningful feedback from caregivers, practitioners, expert advisors, and the children themselves in the creation of *Ahlan Simsim*, we might have created a program that both overreached and failed to meet the children's explicit need for foundational SEL skills, and that did not reflect children's and parents' existing knowledge of these skills. Even worse, without carefully integrating knowledge about the effects of trauma into our content and practices, we risked retraumatizing some children. The implications of these lessons for the education in emergencies community are serious, particularly given how critically important participatory research and community buy-in are to designing responsive, sustainable, effective ECD in emergencies programming (Foulds et al. 2021). Our hope is that the process and approach that we undertook in developing the *Ahlan Simsim* show can be adapted across emergency contexts and various types of SEL interventions, thereby creating a new standard for community-centered, innovative educational programming for children around the world.

REFERENCES

- Bandura, Albert. 1962. "Social Learning through Imitation." In *Nebraska Symposium on Motivation, 1962*, edited by M. R. Jones, 211-74. Lincoln: University of Nebraska Press.
- Bouchane, Kolleen, Hirokazu Yoshikawa, Katie Maeve Murphy, and Joan Lombardi. 2018. *Early Childhood Development and Early Learning for Children in Crisis and Conflict*. Background paper prepared for the 2019 Global Education Monitoring Report. <https://nurturing-care.org/wp-content/uploads/2018/12/UNESCO.pdf>.
- Britto, Pia R., Stephen J. Lye, Kerrie Proulx, Aisha K. Yousafzai, Stephen G. Matthews, Tyler Vaivada, Rafael Perez-Escamilla et al. 2016. "Nurturing Care: Promoting Early Childhood Development." *The Lancet* 389: 91-102. [https://doi.org/10.1016/S0140-6736\(16\)31390-3](https://doi.org/10.1016/S0140-6736(16)31390-3).
- Bronfenbrenner, Urie. 1979. *The Ecology of Human Development*. Cambridge, MA: Harvard University Press.
- Cole, Charlotte. 2016. "The Global Sesame Effect." In *The Sesame Effect: The Global Impact of the Longest Street in the World*, edited by Charlotte Cole and June H. Lee, 9-39. New York: Routledge.
- Cole, Charlotte, Scott Cameron, Estee Bardanashvili, and Shanna Kohn. 2018. "Results of Content & Creative Workshop for Mass Media." Unpublished Sesame Workshop report.
- Cook, Alexandra, Joseph Spinazzola, Julian Ford, Cheryl Lanktree, Margaret Blaustein, Marylene Cloitre, Ruth DeRosa et al. 2005. "Complex Trauma in Children and Adolescents." *Psychiatric Annals* 35 (5): 390-98. <https://doi.org/10.3928/00485713-20050501-05>.
- Cooney, Joan. 1966. "A Report to Carnegie Corporation of New York: The Potential Uses of Television in Preschool Education." Unpublished Sesame Workshop report.
- Cornwall, Andrea. 2002. *Beneficiary, Consumer, Citizen: Perspectives on Participation for Poverty Reduction*. Stockholm: Swedish International Development Cooperation Agency. https://www.participatorymethods.org/sites/participatorymethods.org/files/beneficiary%20consumer%20citizens_cornwall.pdf.

- Durlak, Joseph A., Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, and Kriston B. Schellinger. 2011. "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions." *Child Development* 82 (1): 405-32. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>.
- Felitti, Vincent J., Robert F. Anda, Dale Nordenberg, David F. Williamson, Alison M. Spitz, Valerie Edwards, Mary P. Koss, and James S. Marks. 1998. "Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study." *American Journal of Preventive Medicine* 14 (4): 245-58. [https://doi.org/10.1016/S0749-3797\(98\)00017-8](https://doi.org/10.1016/S0749-3797(98)00017-8).
- Forman-Hoffman, Valerie, Stefanie Knauer, Joni McKeeman, Adam Zolotor, Roberto Blanco, Stacy Lloyd, Elizabeth Tant, and Meera Viswanathan. 2013. "Child and Adolescent Exposure to Trauma: Comparative Effectiveness of Interventions Addressing Trauma Other Than Maltreatment or Family Violence." *Comparative Effectiveness Review* 107: ES1-ES3. https://www.ncbi.nlm.nih.gov/books/NBK126092/pdf/Bookshelf_NBK126092.pdf.
- Foulds, Kim, Naureen Khan, Sneha Subramanian, and Ashraful Haque. 2021. "Implementing a Humanitarian Needs Assessment Framework for Early Childhood Development: Informing Intervention Design for Displaced Rohingya Communities in Bangladesh." *Journal on Education in Emergencies* 7 (1): 112-32. <https://doi.org/10.33682/1uqv-kn0f>.
- GSMA (Global System for Mobile Communications Association). 2019. *The Mobile Economy: Middle East and North Africa*. London: GSMA. <https://www.gsmaintelligence.com/research/?file=87bc4fd841cb69e2fae9e313dcdcc45b&download>.
- INEE (Inter-agency Network for Education in Emergencies). 2010. *INEE Minimum Standards for Education: Preparedness, Response, Recovery*. New York: INEE. https://inee.org/system/files/resources/INEE_Minimum_Standards_Handbook_2010%28HSP%29_EN.pdf.
- Izard, Carroll, Sarah Fine, David Schultz, Allison Mostow, Brian Ackerman, and Eric Youngstrom. 2001. "Emotion Knowledge as a Predictor of Social Behavior and Academic Competence in Children at Risk." *Psychological Science* 12 (1): 18-23. <https://doi.org/10.1111/1467-9280.00304>.

- Jabbar, Sinaria Abdel, and Haidar Ibrahim Zaza. 2014. "Impact of Conflict in Syria on Syrian Children at the Zaatari Refugee Camp in Jordan." *Early Child Development and Care* 184 (9-10): 1507-30. <https://doi.org/10.1080/03004430.2014.916074>.
- Kohn, Shanna., Kim Foulds, Katie Maeve Murphy, and Charlotte Cole. 2020. "Creating a Sesame Street for the Syrian Response Region: How Media Can Help Address the Social and Emotional Needs of Children Affected by Conflict." *Young Children* 75 (1): 32-41. <https://www.naeyc.org/resources/pubs/yc/mar2020/creating-sesame-street-syrian-response>.
- Lagercrantz, Hugo. 2016. *Infant Brain Development: Formation of the Mind and the Emergence of Consciousness*. Solna, Switzerland: Springer International.
- Lahiri, Shaon, Mark van Ommeren, and Bayard Roberts. 2017. "The Influence of Humanitarian Crises on Social Functioning among Civilians in Low- and Middle-Income Countries: A Systematic Review." *Global Public Health* 12 (12): 1461-78. <https://doi.org/10.1080/17441692.2016.1154585>.
- Lesser, Gerald S. 1974. *Children and Television: Lessons from Sesame Street*. New York: Random House.
- Mares, Marie-Louise, and Zhongdang Pan. 2013. "Effects of Sesame Street: A Meta-Analysis of Children's Learning in 15 Countries." *Journal of Applied Developmental Psychology* 34 (3): 140-51. <https://doi.org/10.1016/j.appdev.2013.01.001>.
- Martins, Nicole, and Kristen Harrison. 2012. "Racial and Gender Differences in the Relationship Between Children's Television Use and Self-Esteem: A Longitudinal Panel Study." *Communication Research* 39 (3): 338-57. <https://doi.org/10.1177/0093650211401376>.
- Measham, Toby, Jaswant Guzder, Cécile Rousseau, Laura Pacione, Morganne Blais-McPherson, and Lucie Nadeau. 2014. "Refugee Children and Their Families: Supporting Psychological Well-Being and Positive Adaptation Following Migration." *Current Problems in Pediatric and Adolescent Health Care* 44 (7): 208-15. <https://doi.org/10.1016/j.cppeds.2014.03.005>.

- Melki, Jad, Yasmine Dabbous, Khaled Nasser, Sarah Mallat, Maysa Shawwa, Michael Oghia, Diana Bachoura et al. 2012. *Mapping Digital Media: Lebanon*. London: Open Society Foundations. <https://www.opensocietyfoundations.org/uploads/94b12bbe-399d-4aa8-8890-62494a986de8/mapping-digital-media-lebanon-20120506.pdf>.
- MICT (Media in Cooperation and Transition). 2014. "Syria Audience Research." Berlin: MICT. https://static1.squarespace.com/static/5ecbde016a831b2551433606/t/5efe17ae2845322eabd3371e/1593710511481/syrienstudie_20140814.pdf.
- Panter-Brick, Catherine, Mark Eggerman, Viani Gonzalez, and Sarah Safdar. 2009. "Violence, Suffering, and Mental Health in Afghanistan: A School-Based Survey." *The Lancet* 374 (9692): 807-16. [https://doi.org/10.1016/S0140-6736\(09\)61080-1](https://doi.org/10.1016/S0140-6736(09)61080-1).
- Shaver, Phillip, Judith Schwartz, Donald Kirson, and Cary O'Connor. 2001. "Emotion Knowledge: Further Exploration of a Prototype Approach. In *Emotions in Social Psychology: Essential Readings*, edited by W. Gerrod Parrott, 26-56. Philadelphia: Psychology Press.
- Shonkoff, Jack P., Andrew S. Garner, the Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, and Section on Developmental and Behavioral Pediatrics et al. 2012. "The Lifelong Effects of Early Childhood Adversity and Toxic Stress." *Pediatrics* 129 (1): e232-46. <https://doi.org/10.1542/peds.2011-2663>.
- Singhal, Arvind, and Everett M. Rogers. 1999. *Entertainment-Education: A Communication Strategy for Social Change*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Substance Abuse and Mental Health Services Administration. 2014. *SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach*. HHS Publication No. (SMA) 14-4884. Rockville, MD: Substance Abuse and Mental Health Services Administration. <https://store.samhsa.gov/system/files/sma14-4884.pdf>.
- Sweis, Rana F., and Dina Baslan. 2013. *Mapping Digital Media: Jordan*. London: Open Society Foundations. <https://www.opensocietyfoundations.org/uploads/3480b336-3bf1-4b70-b657-4706a4dab0c4/mapping-digital-media-jordan-20131121.pdf>.

- Tufte, Thomas, and Paolo Mefalopulos. 2009. *Participatory Communication: A Practical Guide*. Washington, DC: World Bank. <https://doi.org/10.1596/978-0-8213-8008-6>.
- Umiltà, Maria Alessandra, Rachel Wood, Francesca Loffredo, Roberto Ravera, and Vittorio Gallese. 2013. "Impact of Civil War on Emotion Recognition: The Denial of Sadness in Sierra Leone." *Frontiers in Psychology* 4: 523. <https://doi.org/10.3389/fpsyg.2013.00523>.
- UNESCO (UN Educational, Scientific and Cultural Organization). 2017. *Global Education Monitoring Report: Aid to Education Is Stagnating and Not Going to Countries Most in Need*. Policy Paper 31. Paris: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000249568>.
- UNICEF (UN Children's Fund). 2018. "Half of Syria's Children Have Grown Up Only Seeing Violence as Conflict Nears Eight-Year Mark: UNICEF." New York: UNICEF. <https://www.unicef.org/press-releases/half-syrias-children-have-grown-only-seeing-violence-conflict-nears-eight-year-mark>.
- Wagner, Peter, Andrea Glioti, Hussein Dawood. 2021. "Iraq." Maastricht, The Netherlands: European Journalism Centre. <https://medialandscapes.org/country/iraq>.
- World Bank. 2016. "How Many Years Do Refugees Stay in Exile?" <https://blogs.worldbank.org/dev4peace/how-many-years-do-refugees-stay-exile>.
- World Health Organization, UNICEF, and World Bank. 2018. *The Nurturing Care Framework for Early Childhood Development: A Framework for Helping Children Survive and Thrive to Transform Health and Human Potential*. Geneva: World Health Organization. <https://nurturing-care.org/ncf-for-eecd>.

APPENDIX

Table A1: Study Methodologies and Sampling⁵

Study	Research Methodology	Purpose/Research Question	Number of Participants	Countries
Study 1: Needs assessment	Individual in-depth interviews with caregivers and practitioners	Learn about children’s academic and social-emotional needs Understand caregivers’ parenting needs Identify and compare caregivers’ and practitioners’ educational priorities for young children Learn more about the professional needs and challenges of people working directly with young children and their families	265 respondents: 195 caregivers of children ages 3-8 70 practitioners who work with young children (e.g., teachers, teaching assistants, facilitators, social workers, health-care workers, and protection of-ficers)	Iraq, Jordan, Lebanon, Syria

⁵ While the Ahlan Simsim program covers Iraq, Jordan, Lebanon, and Syria, not all countries were included in all the formative studies, due to budget constraints, access, and security concerns for the research team and respondents. Given the nature of our methodologies and data-collection processes, sampling presents a limitation. While Sesame Workshop has sought feedback through a variety of methods with diverse populations, our samples were not nationally or regionally representative. Selection bias in sampling for the above-mentioned studies restricts our ability to generalize to the population at large, given that some studies over-indexed on displaced populations while others required respondents to have access to a smartphone and the internet, and in some cases, it was not possible to enter certain parts of the region.

Study	Research Methodology	Purpose/Research Question	Number of Participants	Countries
Study 2: Online research communities	Online surveys, accessible by computer, tablet, or smartphone	The online research communities were used to give constant and ongoing feedback to the Sesame Workshop team as they were developing different pieces of the production, strategy, marketing, and education and research frameworks	11 studies. ⁶ Sample sizes per online study ranged from 75-110 caregivers of children ages 3-6: 20-30 respondents from the host community in each country 15-20 displaced respondents per instance of reaching out in each country	Iraq, Jordan, Lebanon
Study 3: Formative study on children's emotional literacy	Individual interviews conducted with caregiver-child dyads: During the interview, researchers read <i>First Day of School</i> , stopping at certain points to ask the child questions related to the emotional spectrum depicted in the story at that moment. The process was repeated with <i>The Hare and the Tortoise</i> . After both storybooks were read, the researcher interviewed the caregiver while the child colored.	Identify terms children use to communicate their feelings Understand how children express their feelings using those terms Understand from parents and caregivers how children express their emotions	60 caregiver-child dyads: 50% displaced Syrians 50% host community 50% children ages 5-6 50% children ages 7-8	Jordan, Lebanon

⁶ Each study was distinct and framed by relevant research questions specific to the content being tested.

DEVELOPING A CULTURALLY RELEVANT MEASURE OF RESILIENCE FOR WAR-AFFECTED ADOLESCENTS IN EASTERN UKRAINE

SERGIY BOGDANOV, ANDRIY GIRNYK, VIRA CHERNOBROVKINA,
VOLODYMYR CHERNOBROVKIN, ALEXANDER VINOGRADOV, KATERYNA
HARBAR, YULIYA KOVALEVSKAYA, OKSANA BASENKO, IRINA IVANYUK,
KIMBERLY HOOK, AND MIKE WESSELLS

ABSTRACT

Psychosocial support in education that is provided during emergencies frequently aims to support children's resilience, but strong, contextual measures of resilience are in short supply in Eastern Europe. In this article, our aim is to describe the development and psychometric properties of the first measure of resilience for war-affected adolescents in Eastern Ukraine. We used qualitative methods to identify the main cultural characteristics of resiliency and then used these constructs to develop the measure. We used exploratory structural equation modeling to extract five factors that showed high internal consistency: family support ($\omega=0.89$), optimism ($\omega=0.87$), persistence ($\omega=0.87$), health ($\omega=0.86$), and social networking ($\omega=0.87$). Confirmatory factor analysis suggested that a concise model of resiliency fit the data almost as well as the exploratory structural equation modeling model. The measure demonstrated good test-retest reliability. In this article, we also discuss the importance of development, validation, and the use of culturally relevant measures of resilience for strengthening psychosocial support programs in schools, particularly in Ukraine.

Received January 27, 2020; revised September 18, 2020, November 17, 2020, and March 17, 2021; accepted August 2, 2021; electronically published December 2021.

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

INTRODUCTION

In March 2014, amidst political upheaval in Ukraine, Russia invaded and annexed Crimea; these actions were followed by ongoing military conflict among Russia-supported separatists and Ukrainian military forces in Eastern Ukraine. These events resulted in significant social and economic disruption and dislocation (United Nations Refugee Agency n.d.; World Health Organization 2016), with many individuals and families experiencing significant trauma and violence. In addition to significant veteran casualties and injuries, 3,344 civilians (including children) have been killed and more than 7,000 injured since the start of the conflict (United Nations Office of the High Commissioner for Human Rights 2019). Currently, more than 1.5 million people in Ukraine are identified as internally displaced persons (United Nations Refugee Agency: Ukraine 2019). When children grow up in such a context of instability, their ability to overcome exposure to ongoing stress is negatively affected and they may experience long-term health consequences (Ehnholt and Yule 2006; Miller and Rasmussen 2010; Shonkoff et al. 2011; Werner 2012).

A recent survey (Bogdanov, Kovalevskaya et al. 2017) of 466 randomly selected schoolchildren living within five kilometers of the frontline zone in the Donetsk region of Ukraine revealed that the children were very negatively affected by exposure to trauma, which included seeing tanks and other military machines and experiencing shooting or explosions. As a result, a prevailing number of the children surveyed reported suffering effects such as feeling slightly or deeply frightened, having difficulty concentrating, and experiencing impaired sleep. Approximately a quarter of the students reported needing adult support to feel more secure and less worried.

When our work began, there was little scientific evidence that could point to resilience assets specific to the children living in the war zone in Eastern Ukraine. Despite the resilience scales already developed and tested in different cultural contexts (He and van de Vijver 2015; Tol, Song, and Jordans 2013; Windle, Bennett, and Noyes 2011), only a few that measure certain aspects of resiliency, such as the Strength and Difficulties Questionnaire (Goodman, Lamping, and Ploubidis 2010), have been used to measure the effectiveness of psychosocial programs in Ukraine (Bogdanov, Kovalevskaya et al., 2017; Bogdanov, Zalesska, and Basenko 2019). The CYRM-28, one of the most promising resilience measures for youth that embeds the process of cultural adaptation, which we considered adapting to the Ukrainian context, has been successfully adapted and validated with various cultures across the world (Liebenberg, Ungar, and Van De Vijver

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS IN EASTERN UKRAINE

2012; van Rensburg, Theron, and Ungar 2019), specifically in low- and middle-income countries (Kaunda-Khangamwa et al. 2020; van Rensburg et al. 2019). This measure, which has a three-factor structure (individual, relational, and contextual), includes a local functioning scale that offers the possibility of contextualizing it to specific cultures and environments (Ungar 2016). Despite the value of using the CYRM-28, the results of its factor analysis demonstrate some inconsistencies in the underlining constructs (Liebenberg, Joubert, and Foucault 2017; van Rensburg et al. 2019). Using this measure may have resulted in unanswered questions about silent resilience factors, which could considerably change our understanding of locally defined concepts of resilience. Another problem of measuring resilience is that, even with well-developed measures such as the CYRM-28, there is the possibility of bias that comes from not including children's perspectives (Ungar 2016), even though resilience researchers are called on to develop local measures from the users' perspective (Betancourt et al. 2013).

Accurate psychosocial research and appropriate service delivery require reliable, valid, and useful measurement tools. The literature repeatedly calls attention to the high need for validated resiliency measures for the purposes of program implementation (Clauss-Ehlers 2008; Windle, Bennett, and Noyes 2011), especially in emergency settings. These kinds of measures for use with conflict-affected populations are frequently lacking, owing to the difficulty and cost of local adaptation and testing. To address this literature gap, we describe in this article the academic-practitioner work of the National University of the Kyiv-Mohyla Academy and its diverse partners in Ukraine. Our aim was to develop a contextualized concept and measure of resilience that applies to war-affected adolescents in Eastern Ukraine. In the remainder of this article, we describe our use of qualitative and statistical methods to create and validate a local resilience questionnaire; discuss a local concept of resilience; and make suggestions for the development of future psychosocial programs in Ukraine.

DEFINING AND MEASURING RESILIENCE

Education is a fundamental right and a high priority for children in all humanitarian settings. Because many children living in conflict zones are affected by their war experiences, losses, displacement, and insecurity (e.g., Sagi-Schwartz 2008), school-based psychosocial support (PSS) serves to enhance the wellbeing of children and adolescents while also enabling them to learn (Ager et al. 2011; Jordans, Pigott, and Tol 2016; Jordans et al. 2010). The field of PSS in emergency settings increasingly recognizes the importance of avoiding deficit-focused

approaches to education and working to strengthen the resilience of war-affected children, as well as their families, teachers, school personnel, etc.

One challenge in modern resilience science is the inconsistency in the literature regarding how resilience is defined and measured (Masten 2018). Many definitions of resilience are continually revised, indicating the continuing development of this field. Most of these definitions feature the roles of adaptive functioning and problem-solving in the face of significant adversity. When confronted with adversity, individuals rely on their resources and strengths to actively navigate and engage with their social environment. Generally speaking, resiliency is defined as a trait, an outcome, a process, or an overall domain that reflects all of these concepts. Others have defined resilience as the ability to adapt positively to hardship or challenges (Masten 2018). Based on extensive empirical research, Rutter (2006, 2012) discussed children's resilience as a positive psychological outcome in various risk situations. These findings suggest that resilience is not exemplified by the superior functioning of a "superkid"; instead, they emphasize that different environments lead to differences in children's responses to adversity.

Masten (2007, 2018) offered a different conceptualization, describing resilience as individual- and family-centered concepts from the framework of systems theory. From Masten's theoretical point of view, childhood resilience is not a single and stable personal trait but the result of dynamic interactions across and between interacting systems (Masten 2007), which could be extended to include community systems. Other authors similarly note the importance of environment: Werner (1996) took a socioecological perspective of resilience, describing protective factors at the individual, family, and community levels, and Luthar (Luthar, Cicchetti, and Becker 2000; Luthar and Brown 2007) identified resilience as an interaction between a child and their environment. Ungar (2013), who discussed resilience as a phenomenon and a process that arise from the interaction between the individual and their environment, also specified the critical role the environment plays, due to its ability to provide individuals and groups with access to protective factors and support resources. Ungar (2008, 2011, 2013) emphasizes that these interactions are always influenced by the specific context and culture.

The current literature attempts to identify factors of child and adolescent resilience that are similar across situations and cultures. While some commonalities exist, it is also clear that resilience is affected by specific cultural contexts and by a dynamic interplay of many variables (Tol et al. 2013; Ungar 2008). The dominance of Western-based ideas about what resilience is and how it can be measured limits our understanding of silent cultural factors and creates bias toward psychosocial

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS IN EASTERN UKRAINE

intervention goals in war-affected regions (Kaunda-Khangamwa et al. 2020; Ungar 2013). Building consensus on key definitions of resilience, specifically understanding the effects exposure to trauma has on developmental outcomes, will enable researchers to build models that differentiate the main protective and promotive factors at different socioecological levels, which is crucial to promoting best outcomes in high-risk situations (Betancourt et al. 2013; Garmezy 1991; Masten 2007; Peltonen et al. 2014). On the one hand, existing evidence conclusively supports the notion that children exposed to conflict are able to negotiate negative outcomes using similar strategies across contexts and cultures, such as seeking parent and peer support, relying on their cognitive skills and self-regulation, a sense of faith and hope, and on their self-efficacy, acculturation, and prosocial skills (Cicchetti 2010; Masten 2007; Tol et al. 2013). Previous exposure to trauma also must be considered, in addition to gender and developmental periods, by studying individual differences in responding to adverse events (Cicchetti 2010; Masten 2018). On the other hand, researchers still look for explanations of how different learning environments, as represented by parenting beliefs, behavioral strategies, and socialization goals, can shape the developmental consequences of a growing child (Keller and Otto 2009).

The challenges associated with operationalizing the construct of resiliency is closely linked with the development of valid, culturally relevant, and practicable psychometric scales for measuring resilience. Although the development and validation of such measure are critical in diverse contexts to ensure that constructs are appropriately described and identified (van de Vijver and Leung 2000), few contextualized measures of resilience have as yet been developed.

CONCEPT AND INSTRUMENT DEVELOPMENT

MIXED METHODS APPROACH

In order to understand the concept of resilience in Ukraine from the ground up (an inductive rather than an a priori approach), we designed a mixed methods study to first identify local understanding of resilience and subsequently to inform the novel resilience instrument.

CONTEXTUALIZATION OF A LOCAL RESILIENCE CONSTRUCT:
QUALITATIVE METHODS

We conducted an initial qualitative study that aimed to explore assets in a local resilience construct and inform the validation study. We conducted the study in Pokrovsk (formerly Krasnoarmeysk), which is located within the security buffer zone controlled by the Ukrainian government (Bogdanov, Girnyk, Lasorenko et al. 2017). Sixty-seven children between the ages of 9 and 18 participated. We obtained parental consent and participant (i.e., child) assent before conducting the interviews. We used free listing (i.e., listing the different responses to each question and associating interviewee identification numbers next to each response) (Weller and Romney 1988) to produce brief textual responses via semistructured interviews (Table 1), and we completed our analysis by identifying the central themes (i.e., coding) revealed by grouping texts and running frequencies for each code. Specifically, respondents listed everything related to a child's characteristics, resources, or abilities that help them overcome stressful life events. Our analysis of these descriptions revealed the following resiliency categories: happy (n=32), communicative (n=21), optimistic (n=10), family support (n=10), helpful (n=10), curiosity and intelligence (n=8), and persistence (n=7). These Ukrainian children described attempting to cope with traumatic stress by engaging in pleasant activities such as developing hobbies, looking for support from their parents and friends, and using positive thinking. We also identified two negative strategies for coping with stress, namely, isolation and conflict with others. These categories were subsequently included in the measure's items. We also conducted ten focus groups (total N=53) with young adolescents in order to gain detailed information about the selected resilience categories identified in the free-listing interviews.

Table 1: Questions Included in the Qualitative Study

Open-Ended Questions	
1	What are the main problems facing families living here in the conflict area?
2	How would you describe children who are feeling well and are growing and developing well, in spite of the many problems they face?
3	How would you describe a child who is in a state of sadness or despondency?
4	What do children (such as you) who live in the buffer zone do to take care of themselves and others?
5	How do children (such as you) cope with sadness? What else can help children in a state of sadness or despondency?
6	What can adults do to alleviate grief or sadness in children?"

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

The resulting model of local adolescent resiliency included 12 characteristics: family support, helping others, communicating with friends, conflict with others, networking skills, isolation, happiness, optimism, physical health, self-confidence, persistence, and curiosity. All the categories represented two levels of socioecology, relationship and individual; family support, helping others, communicating with friends, and conflict with others were at the relationship level, and the rest of the components were specific to the individual level. Interestingly, their relationships with teachers and school performance were not mentioned during either the free-listing interviews or the focus group discussions; thus, we did not include these categories in the model. Any macro-level characteristics, such as community traditions, rituals, religion, or spirituality, did not appear in the qualitative results. We can speculate that the absence of contextual variables common in other countries affected by war (Cortes and Buchanan 2007; Eggerman and Panter-Brick 2010) is a result of the local culture in Eastern Ukraine. Historically, this geographic region was shaped as individuals of various Slavic nationalities moved into the Donbass region. Over time, the local spiritual traditions were replaced with an impersonal proletarian culture, especially during the time of the Soviet Union (Kusina n.d.). The physical health factor expressed in one child's comment that they were "in strong body, strong spirit" might be the closest representation of perceived spirituality.

We developed a pool of 146 items, based on the initial model of youth resiliency. We added 11 questions to test consistency in the participants' responses. We used a Likert-type scale (5=not at all true, 4=somewhat not true, 3=true as much as false, 2=somewhat true, 1=completely true).

VALIDATION STUDY SAMPLE AND PARTICIPANT DEMOGRAPHICS:
QUANTITATIVE METHODS

In the initial validation study, 218 participants living in the military zones in Ukraine completed the measure. The mean age of participants was 13.98 years old (SD=1.3; age range 12-17). More than half (57.3%) of the sample identified as female. All the children were recruited from five schools located in the city of Pokrovsk. The validation study sample was different from the qualitative survey sample, but both represented a similar population that included children from Pokrovsk and children of internally displaced persons. To limit the burden of an already lengthy assessment, we did not ask these children to complete a checklist of adverse life events.

FACTOR ANALYSIS

The initial number of categorical variables was too large to conduct an exploratory factor analysis (sample size of 218 observations), so there was a clear need for a feature selection procedure to reduce it substantially. To this end, we computed a 146x146 matrix of Pearson product-moment correlations of items, and for each item we counted the number of correlations with other items that exceeded a value of 0.5 (the threshold for the large correlational effect size proposed by Jacob Cohen (1992)). For the subsequent analysis, we retained only items with two or more large correlations. This decision assumed that at least three variables are required to identify a factor. Variables with fewer and weaker correlations were less likely to form a stable interpretable latent dimension.

We performed all confirmatory factor analyses (CFA) with Mplus 8.4 using a weighted least square mean and variance adjusted (WLSMV) estimator. We tested a series of four competing models (see Table 2). Model 1 was a unidimensional model, where all items were indicators of a general factor of resilience. Model 2 was based on five factors that were allowed to correlate. Model 3 was the conventional higher order model, with five first-order factors and one second-order general factor. Model 4 was the corresponding bifactor model, which allowed each item to load simultaneously onto the general factor and onto its corresponding index factor. Correlations of general and specific factors were fixed at 0 in the model.

Table 2: Comparison of CFA Models

Model	df	χ^2	RMSEA	CFI	TLI	SRMR
Unidimensional	324	1510.9	.131	.776	.758	.126
Correlated Factors	311	509.6	.055	.963	.958	.060
Second Order	316	542.6	.058	.957	.953	.067
Bifactor	294	524.3	.061	.957	.948	.062

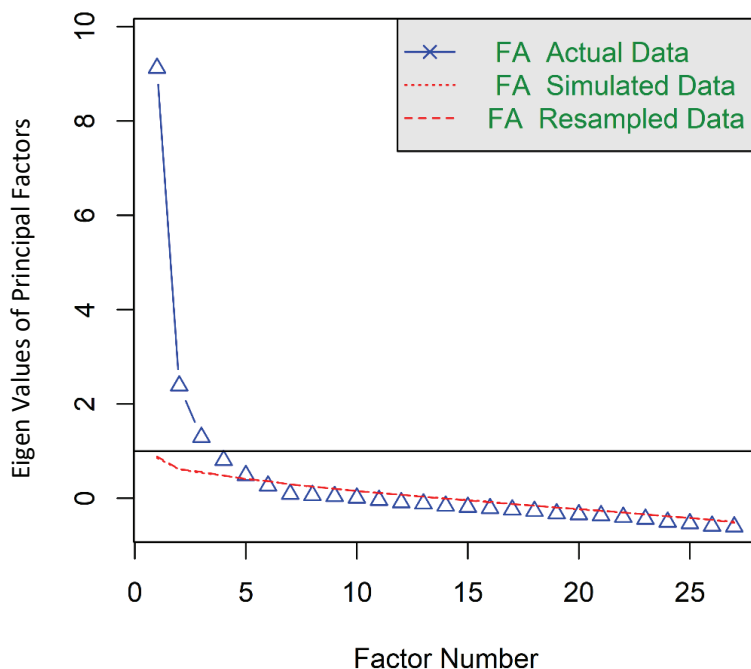
We calculated estimates of internal consistency for each factor using McDonald's ω , and we assessed test-retest reliability on a separate small sample of 31 respondents.

RESULTS

Analysis of the correlation matrix showed that, of 146 variables, 104 did not have a single high ($r > 0.5$) correlation with the rest, and 15 had only one such relationship. The set of 27 highly correlated variables consisted of items with two to eight correlations greater than 0.5.

The results of dimensionality analysis were uncertain: index very simple structure complexity 1 achieved a maximum of 0.86 with one factor; very simple structure complexity 2 achieved a maximum of 0.93 with two factors; the Velicer's minimum average partial and empirical Bayesian information criterion achieved a minimum of 0.03 and -995.53, respectively, thus indicating that four factors should be retained. Sample size-adjusted Bayesian information criterion achieved a minimum of 3609.45 with 20 factors. Based on the results of a parallel analysis, we decided to extract five latent dimensions (see the scree plot in Figure 1).

Figure 1: Parallel Analysis Scree Plots



We performed the actual extraction of 5 factors for 27 categorical observed variables with Mplus, using the exploratory structural equation modeling procedure. Exploratory structural equation modeling is equivalent to the usual exploratory factor analysis with an WLSMV estimator and oblique Geomin rotation, but it allows for an estimation of additional parameters, such as error covariances. The obtained solution was good, with a small number of statistically significant cross-loadings, small modification indices for error covariances, and good model fit indices (RMSEA=.058 with 90% CI [.048, .068], SRMR=.036, CFI=.969, TLI=.952, $\chi^2=391.9$, $df=226$, $p < .0001$).

The estimates of reliability for the summated scales using model-based McDonald's ω were as follows: F1=0.89, 95% CI [.85, .92], F2=0.87, 95% CI [.84, .91], F3=0.87, 95% CI [.84, .90], F4=0.86, 95% CI [.82, .90], F5=0.87, 95% CI [.83, .90].

To test a simpler model with a minimal number of crossloadings, we conducted a CFA where each latent factor was measured by a limited number of the most important items. This more concise CFA model included only underscored loadings (see Table 3 for goodness-of-fit indices). Since exploratory factors have substantial correlations, their relationship can be explained in various ways. First, there can be one common factor behind all indicators of the resilience construct. To test this hypothesis, goodness-of-fit indices for the unidimensional model should be checked. Second, there may be a second-order factor that affects first-order factors. Finally, there may be a bifactor model that allows all items to load onto both the general factor and their respective specific factor directly. We present goodness-of-fit indices for these models in Table 3. We present parameter estimates and reliabilities (McDonald's ω and ω_h) for the model in Table 4.

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

Table 3: ESEM Standardized Loadings and Factor Correlations of 27 Items

Item Number	Items	F1	F2	F3	F4	F5	R ²
Item 1	I know that my parents love me (Я знаю, что мои родители любят меня)	.93	.20	-.12	-.25	.02	.88
Item 2	My family is interested in what my favorite games and activities are (Моим родным интересно, какие у меня любимые игры и занятия)	.88	-.20	.11	-.03	.06	.73
Item 3	If I have problems, my parents usually talk to me, find out the reason, and help in a difficult situation (Если у меня проблемы, мои родители обычно говорят со мной, узнают причину и помогают в трудной ситуации)	.85	.00	.16	.11	-.31	.75
Item 4	Adults communicate with me, they can support and calm me (Взрослые общаются со мной, могут меня поддержать и успокоить)	.84	-.19	.23	.14	-.02	.81
Item 5	I have a good relationship with my family (У меня хорошие отношения в семье)	.77	.24	-.10	-.06	.01	.72
Item 6	My parents can listen to me without criticizing me (Мои родители могут выслушать меня и не критиковать при этом)	.72	.05	-.05	.09	.00	.59
Item 7	In any, even difficult situations, I find something to be happy about (В любой, даже сложной ситуации, я нахожу, чему можно порадоваться)	-.01	.69	.31	.03	-.14	.63
Item 8	I am cheerful (Я жизнерадостный)	.03	.56	-.04	.47	-.04	.72
Item 9	I love life despite the difficulties (Несмотря на трудности, я люблю жизнь)	.18	.55	.01	.31	.02	.72

Item Number	Items	F1	F2	F3	F4	F5	R ²
Item 10	I can enjoy everything and often smile (Я умею радоваться всему, часто улыбаюсь)	-.02	.51	.17	.37	.04	.73
Item 11	Overall, I'm a happy person (В целом, я счастливый человек)	.12	.53	.20	.12	.06	.61
Item 12	I know that everything will be fine (Я знаю, что все будет хорошо)	.23	.48	.06	.26	.01	.63
Item 13	I understand how to help myself when I see how others overcome difficulties (Я понимаю, как себе помочь, когда вижу, как другие преодолевают трудности)	.04	.00	.75	.21	-.10	.66
Item 14	When I strive for my goals, I am able to overcome difficulties (Когда я стремлюсь к своим целям, то я способен преодолевать трудности)	-.04	-.06	.69	.27	.03	.66
Item 15	I can unobtrusively get a person to talk, find out what happened to him (Я умею ненавязчиво разговаривать человека, узнать, что с ним случилось)	-.02	.00	.69	-.01	.17	.60
Item 16	When the problem is too difficult, I try another approach (Когда проблема оказывается слишком трудной, я пробую к ней другой подход)	.21	.07	.77	-.13	.01	.68
Item 17	I am not afraid of criticism and always ready to substantiate my position (Я не боюсь критики и всегда готов обосновать свою позицию)	.10	.07	.63	.02	.10	.58
Item 18	I believe that I can find a way out of any situation (Я считаю, что из любой ситуации можно найти выход)	-.05	.36	.61	-.18	.15	.66

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

Item Number	Items	F1	F2	F3	F4	F5	R ²
Item 19	I easily get acquainted with new people (Я легко знакожусь с новыми людьми)	-.09	.01	.08	.00	.91	.86
Item 20	I easily make new friends (Я легко нахожу новых друзей)	.10	.02	-.01	.17	.71	.73
Item 21	I easily make contact with peers (Я легко иду на контакт со сверстниками)	.08	-.01	.17	.21	.52	.60
Item 22	In general, I can find common language with many people (В целом, я могу найти общий язык со многими)	.02	-.01	.33	.16	.49	.64
Item 23	Overall, I am confident in my abilities (В целом, я уверен в своих силах)	.04	.11	.41	.23	.31	.69
Item 24	I am agile and energetic (Я подвижный и энергичный)	-.01	.00	-.03	.85	.12	.80
Item 25	I feel physically healthy (Я чувствую себя физически здоровым)	.11	.07	-.13	.68	.09	.59
Item 26	I often play agile games (Я часто играю в подвижные игры)	-.08	.11	.13	.66	.01	.56
Item 27	I look after my health because I believe that “in a healthy body there is a healthy mind” (Я слежу за своим здоровьем, потому что верю, что «в здоровом теле – здоровый дух»)	.01	.15	.11	.55	.04	.51

Note: Statistically significant ($p < 0.05$) loadings and factor correlations are bolded.

Conceptually, the preliminary classification of the measure items helped us to identify relevant factors: family support, optimism, persistence, physical health, and social networking. The concise model with only three crossloadings fit the data almost as well as the full exploratory model. All CFA models fit the data equally well. However, the simplest model with correlated factors had slightly better characteristics. The use of the bifactor model enables us to interpret the

total score and the score for the family support subfactor, while the remaining four factors are more likely a specifically formulated general factor (see Table 4).

Table 4: Bifactor Solution: Loadings on the General (GF) and Specific Factors (SF1-SF5)

	GF	SF1	SF2	SF3	SF4	SF5
Item 1	0.38	0.84				
Item 2	0.41	0.74				
Item 3	0.43	0.71				
Item 4	0.57	0.67				
Item 5	0.45	0.69				
Item 6	0.45	0.63				
Item 7	0.59		0.42			
Item 8	0.62		0.51		0.42	
Item 9	0.73		0.47			
Item 10	0.75		0.35		0.25	
Item 11	0.71		0.30			
Item 12	0.71		0.40			
Item 13	0.60			0.56		
Item 14	0.63			0.52		
Item 15	0.57			0.51		
Item 16	0.59			0.47		
Item 17	0.62			0.46		
Item 18	0.62			0.39		
Item 19	0.61					0.75
Item 20	0.69					0.46
Item 21	0.70					0.31
Item 22	0.74					0.28
Item 23	0.77			0.25		0.19
Item 24	0.67				0.55	
Item 25	0.60				0.43	
Item 26	0.60				0.46	
Item 27	0.63				0.37	
ω	0.97	0.94	0.92	0.92	0.91	0.92
ω_h	0.83	0.67	0.24	0.31	0.26	0.22

Note: All loadings are statistically significant ($p < .05$).

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

The questionnaire demonstrated good internal consistency and test-retest reliability, with the latter assessed by a Pearson correlation coefficient between the first and second administration of the test to 31 subjects, with an interval of one week (see Table 5).

Table 5: Test-Retest Reliability of Scales

Scale	<i>r</i>
Family Support (F1, SF1)	.89
Optimism (F2, SF2)	.79
Persistence (F3, SF3)	.68
Social Networking (F5, SF5)	.70
Physical Health (F4, SF4)	.78
Resilience, Total	.82

DISCUSSION

In this article, we describe the inductive development of a contextualized resilience construct and psychometric testing of a contextually validated resiliency scale in a sample of young adolescents who are living in the frontline area of Eastern Ukraine. By incorporating locally relevant items developed from prior qualitative data in the same population, we generated a brief, reliable, and valid measure of resilience factors on different socioecological levels. The measure, which for convenience we refer to as the Kyiv-Mohyla Academy Resiliency Screener for Youth, can be used with conflict-affected male and female adolescents in Ukraine. The time needed to administer the 27-item scale is approximately ten minutes, which makes the Kyiv-Mohyla Academy Resiliency Screener for Youth an easy-to-use tool for everyday PSS praxis. This can significantly reduce the burden on children when testing different resilience models through a combination of events checklists and outcome measures.

The local Ukrainian resilience construct, operationalized within a bifactor model with a “general” resilience factor and “family support” as the only significant subfactors, mostly fits with Masten’s (2018) theoretical framework of resilience as an interactional process between individual and family variables. The meaning of family support as a resilience factor is represented by parenting styles such as deep listening, showing a real interest in a child’s life, having a positive attitude toward a child, and not being critical. In many war-affected countries around the globe, family support is one of the most distinct and significant resilience assets

that has been studied (Cicchetti 2010; Masten 2018; Tol et al. 2013). Unfortunately, ethnographic studies that explore the specific cultural psychology of Eastern Ukraine are rare. One qualitative study, conducted among miners and their families living in the eastern and western cities of Ukraine, described a socialization phenomenon within the miner culture, namely, the importance of the family unit within the local community values hierarchy. These authors (Kusina n.d.) suggested that concern about the family and its welfare allows the miners to justify the daily risks they face when going underground. From this perspective, we can perhaps better understand the hidden emotional ties within the family unit and the extraordinary role they play as a protective factor in the life-threatening experiences of both adults and children in Eastern Ukraine. We consider family factors universal, but at the same time unique, in their manifestation as one of the significant relational and contextual factors within the local ecosystem of child resilience.

Other locally specific assets of a general resilience factor, such as optimism, physical health, social networking, and persistence, are not unique characteristics of resilience per se and have been already named by many authors as ego-resilience assets that help youth to regulate their emotions and navigate pathways to existing external resources (Betancourt et al. 2013; Cicchetti 2010; Jefferies et al. 2019; Tol et al. 2013). We also found that the social networking factor represented an essential relational component of adolescents' resilience and children's ability to access important social resources, which can be understood as a process of establishing social contacts in new place.

The local resilience construct, which compares to Ungar's (2011) sociocultural model, is mostly characterized by the absence of important relational and context variables, such as community support (Kuterovac-Jagodić 2003) and spirituality (Duraković-Belko, Kulenovic, and Dapic 2003; Klasen et al. 2010). This may inform PSS programs in terms of understanding family support and developing a social network of peers as the most significant external resources for young adolescents, but it also might raise some concern, in light of the increased number of family violence cases reported in Ukraine (UNICEF 2018).

The lack of value children and adolescents place on other supportive relationships with caring adults (e.g., schoolteachers) also raises questions about the perceived quality of these relationships in Eastern Ukraine. This issue has already been recognized by the Ukrainian ministry of education in its New Ukrainian School reform concept, which aims to change teachers' attitudes from authoritarian to more child focused (Government of Ukraine n.d.).

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS IN EASTERN UKRAINE

We suggest that PSS programs in Ukraine focus on building more open-school ecosystems that engage parents as active actors in the education process. Another promising way to develop more resilient responses may include focusing psychosocial programs that use cognitive-behavioral approaches, such as positive thinking and communicative skills, that have already been effective in the education setting (Bogdanov, Gnida et al. 2017; Bogdanov et al. 2019). Physical literacy and sports programs could benefit psychosocial programs through their positive impact on resilience, but also as a contextually relevant component of positive coping (Jefferies et al. 2019). Finally, receiving additional attention from PSS program staff members might be helpful to introverted children.

A significant feature of the tool described above is its grounding in Ukraine, and thus its contextual relevance. A mixed method approach helped to limit researcher bias about key assets of a local resilience concept by instead investigating this concept through a qualitative exploration of children's subjective perceptions. Our systematic procedure of qualitative data analysis ensured a triangulation of selected resilience assets that were further operationalized in a set of 27 questionnaire items. Rigorous statistical analysis enabled us to identify five factors of local resilience and to evaluate the internal consistency and reliability of the new measure. Our hope is that the approach we have outlined may inspire researcher-practitioner teams in other non-Western countries to conceptualize and measure resilience in ways that fit the local context and culture, and to use these measures to gauge the effectiveness of school-based PSS in emergencies.

LIMITATIONS

We note several limitations of this study. First, although the results suggest the usefulness and appropriateness of this scale, we sampled individual adolescents from only one urban area and the sample size was relatively small. Future work to confirm and expand these findings would add to this work. Second, we did not interview parents or schoolteachers, which may have enabled us to more fully understand the context of the young children in our sample and the processes related to how the local context informs developmental outcomes.

CONCLUSION

To our knowledge, this is the first validity study of instruments to evaluate resiliency among Ukrainian adolescents affected by the country's current conflict. Our hope is that this instrument will be used to measure the effects

various psychosocial programs have on children living near the conflict line, and to strengthen the ability of government and international institutions to select appropriate and effective psychosocial programs. Therefore, it will be freely available to other researchers and clinical workers in Ukraine. The authors, in cooperation with the Ukrainian Ministry of Education and Science, plan to disseminate the Kyiv-Mohyla Academy Resiliency Screener for Youth within the national school psychologists' network and to international organizations working in Ukraine with war-affected children. This will expand our ongoing efforts to implement safe-space interventions for school psychologists (Bogdanov, Gnida et al. 2017) and teachers (Bogdanov, Girnyk, Chernobrovkina et al., 2017) in the eastern regions of Ukraine.

ACKNOWLEDGMENTS

All study procedures were approved by the Expert and Ethics Review Board of the Scientific-Methodological Center of Practical Psychology of the National Academy of Pedagogical Science. Data for this study were collected as part of a project founded by UNICEF Ukraine ("Comprehensive Psychosocial Support and Assistance to Mobile Team Members Who Work in the 'Grey Zone'"). Kimberly Hook is supported by the National Institute of Health (T32MH116140). The sponsors had no role in the study design; in the collection, analysis, and interpretation of data; in writing the articles; or in the decision to submit it for publication.

REFERENCES

- Ager, Alastair, Bree Akesson, Lindsay Stark, Eirini Flouri, Braxton Okot, Faith McCollister, and Neil Boothby. 2011. "The Impact of the School-Based Psychosocial Structured Activities (PSSA) Program on Conflict-Affected Children in Northern Uganda." *Journal of Child Psychology and Psychiatry* 52 (11): 1124-33. <https://doi.org/10.1111/j.1469-7610.2011.02407.x>.
- Betancourt, Theresa, Sarah Meyers-Ohki, Alexandra Charrow, and Tol Wietse. 2013. "Interventions for Children Affected by War: An Ecological Perspective on Psychosocial Support and Mental Health Care." *Harvard Review of Psychiatry* 21 (2): 70-91. <https://doi.org/10.1097/HRP.0b013e318283bf8f>.

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

- Bogdanov, Sergiy, Andriy Girnyk, Boris Lasorenko, Volodymir Savinov, and Viktoriya Solovjova. 2017. "Psychosocial Factors of Resilience among Children Who Are Living in the Front-Line Zone in the East of Ukraine." *Проблеми політичної психології [Journal of Political Psychology]* 4 (18).
- Bogdanov, Sergiy, J. Kovalevskaya, Andriy Girnyk, Volodymir Chernobrovkin, Vira Chernobrovkina, Vitaly Panok, Nadiya Lunchenko et al. 2017. *Research Summary: UNICEF Psychosocial Support Programs for School Children in Donetsk and Luhanska Oblasts*. http://ekmair.ukma.edu.ua/bitstream/handle/123456789/11861/UNICEF_KMA_Research%20summary_en_2017.pdf?sequence=1.
- Bogdanov, Sergiy, Andriy Girnyk, Vira Chernobrovkina, Volodymir Chernobrovkin, and Viktoriya Solovjova. 2017. Підготовка вчителів до розвитку життєстійкості/стресостійкості у дітей в освітніх навчальних закладах: навчально-методичний посібник [Pedagogical Training Curriculum on Resilience Strengthening Among School Children: Methodological Book]. Kyiv, Ukraine: NaUKMA.
- Bogdanov, Sergiy, Tetiana Gnida, Oksana Zalesska, Vitaly Panok, and Viktoriya Solovjova. 2017. Корекційно-розвиткова програма формування стійкості до стресу в дітей дошкільного віку та школярів "Безпечний простір": навчально-методичний посібник. [Psychological Program "SAFE SPACE" on Resilience Strengthening among Pre-School and School Children]. Kyiv, Ukraine: NaUKMA.
- Bogdanov, Sergiy, Oksana Zalesska, and Oksana Basenko. 2019. "Effectiveness Testing of a Multi-Layered Model of Psychosocial Support within Schools in Conflict Affected Eastern Ukraine." *NORRAG Special Issue* 02.
- Cicchetti, Dante. 2010. "Resilience under Conditions of Extreme Stress: A Multilevel Perspective." *World Psychiatry* 9 (3): 145-54. <https://doi.org/10.1002/j.2051-5545.2010.tb00297.x>.
- Clauss-Ehlers, Caroline S. 2008. "Sociocultural Factors, Resilience, and Coping: Support for a Culturally Sensitive Measure of Resilience." *Journal of Applied Developmental Psychology* 29 (3): 197-212. <https://doi.org/10.1016/j.appdev.2008.02.004>.

- Cohen, Jacob. 1992. "A Power Primer." *Psychological Bulletin* 112 (1): 155-59. <https://doi.org/10.1037//0033-2909.112.1.155>
- Cortes, Liliana, Marla Jean Buchanan. 2007. "The Experience of Columbian Child Soldiers from a Resilience Perspective." *International Journal for the Advancement of Counselling* 29: 43-55. <https://doi.org/10.1007/s10447-006-9027-0>.
- Duraković-Belko, Elvira, Alija Kulenovic, and Renko Dapic. 2003. "Determinants of Posttraumatic Adjustment in Adolescents from Sarajevo Who Experienced War." *Journal of Clinical Psychology* 59 (1): 27-40. <https://doi.org/10.1002/jclp.10115>.
- Eggerman, Mark, and Catherine Panter-Brick. 2010. "Suffering, Hope, and Entrapment: Resilience and Cultural Values in Afghanistan." *Social Science and Medicine* 71 (1): 71-83. <https://doi.org/10.1016/j.socscimed.2010.03.023>.
- Ehnholt, Kimberly A., and William Yule. 2006. "Practitioner Review: Assessment and Treatment of Refugee Children and Adolescents Who Have Experienced War-Related Trauma." *Journal of Child Psychology and Psychiatry* 47 (12): 1197-1210. <https://doi.org/10.1111/j.1469-7610.2006.01638.x>.
- Garmezy, Norman. 1991. "Resilience in Children's Adaptation to Negative Life Events and Stressed Environments." *Pediatric Annals* 20 (9): 459-66. <https://doi.org/10.3928/0090-4481-19910901-05>.
- Goodman, Anna, Donna L. Lamping, & George B. Ploubidis. 2010. "When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children." *Journal of Abnormal Child Psychology* 38: 1179-91. <https://doi.org/10.1007/s10802-010-9434-x>.
- Government of Ukraine. n.d. *Education and Science Reform*. <https://www.kmu.gov.ua/en/reformi/rozvitok-lyudskogo-kapitalu/reforma-osviti>.
- He, Jia, and Fons J. R. van de Vijver. 2015. "The Value of Keeping an Open Eye for Methodological Issues in Research on Resilience and Culture." In *Youth Resilience and Culture: Commonalities and Complexities*, edited by Linda Theron, Linda Liebenberg, and Michael Ungar, 189-201. New York: Springer. https://doi.org/10.1007/978-94-017-9415-2_14.

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

- Jefferies, Philip, Michael Ungar, Patrice Aubertin, and Dean Kriellaars. 2019. "Physical Literacy and Resilience in Children and Youth." *Frontiers in Public Health* 7: 346. <https://doi.org/10.3389/fpubh.2019.00346>.
- Jordans, Mark J.D., Ivan H. Komproe, Wietse A. Tol, Brandon A. Kohrt, Nagendra P. Luitel, Robert D. Macy, and Joop T.V.M. De Jong. 2010. "Evaluation of a Classroom-Based Psychosocial Intervention in Conflict-Affected Nepal: A Cluster Randomized Controlled Trial." *Journal of Child Psychology and Psychiatry* 51 (7): 818-26. doi: <https://doi.org/10.1111/j.1469-7610.2010.02209.x>.
- Jordans, Mark J. D., Hugo Pigott, and Wietse A. Tol. 2016. "Interventions for Children Affected by Armed Conflict: A Systematic Review of Mental Health and Psychosocial Support in Low- and Middle-Income Countries." *Current Psychiatry Reports* 18 (1): 9. <https://doi.org/10.1007/s11920-015-0648-z>.
- Kaunda-Khangamwa, Blessings N., Innocent Maposa, Rosalia Dambe, Kennedy Malisita, Emmanuel Mtagalume, and Lalio Chigaru, et al. 2020. "Validating a Child Youth Resilience Measurement (CYRM-28) for Adolescents Living with HIV (ALHIV) in Urban Malawi." *Frontiers in Psychology* 11. doi: <https://doi.org/10.3389/fpsyg.2020.01896>.
- Keller, Heidi, and Hiltrud Otto. 2009. "The Cultural Socialization of Emotion Regulation during Infancy." *Journal of Cross-Cultural Psychology* 40 (6): 996-1011. <https://doi.org/10.1177/0022022109348576>.
- Klasen, Fionna, Gabriele Oettingen, Judith Daniels, Manuela Post, Catrin Hoyer, and Hubertus Adam. 2010. "Posttraumatic Resilience in Former Ugandan Child Soldiers." *Child Development* 81: 1096-1113. <https://doi.org/10.1111/j.1467-8624.2010.01456.x>.
- Kusina, Ksenia, n.d. "Miners Stories from Eastern and Western Ukraine." Donbas Studies. *International Journal for the Advancement of Counselling*. <https://donbasstudies.org/en/miners-stories>.
- Kuterovac-Jagodić, Gordana. 2003. "Posttraumatic Stress Symptoms in Croatian Children Exposed to War: A Prospective Study." *Journal of Clinical Psychology* 59 (1): 925. <https://doi.org/10.1002/jclp.10114>.

- Liebenberg, Linda, Natacha Joubert, and Marie-Lynne Foucault. 2017. *Understanding Core Resilience Elements and Indicators a Comprehensive Review of the Literature*. Public Health Agency of Canada. <http://lindaliebenberg.com/wp-content/uploads/2017/11/PHAC-Resilience-Report-Final-Version-November-2017.pdf>.
- Liebenberg, Linda, Michael Ungar, and Fons J.R. van de Vijver. 2012. "Validation of the Child and Youth Resilience Measure-28 (CYRM-28) among Canadian Youth." *SAGE Open* 22 (2): 219-26. <https://doi.org/10.1177/1049731511428619>.
- Luthar, Suniya S., and Pamela J. Brown. 2007. "Maximizing Resilience through Diverse Levels of Inquiry: Prevailing Paradigms, Possibilities, and Priorities for the Future." *Development and Psychopathology* 19 (3): 931-55. <https://doi.org/10.1017/S0954579407000454>.
- Luthar, Suniya. S., Dante Cicchetti, and Bronwyn Becker. 2000. "The Construct of Resilience: A Critical Evaluation and Guidelines for Future Work." *Child Development* 71 (3): 543-62.
- Masten, Ann. 2007. "Resilience in Developing Systems: Progress and Promise as the Fourth Wave Rises." *Development and Psychopathology* 19 (3): 921-30. <https://doi.org/10.1017/S0954579407000442>.
- Masten, Ann. S. 2018. "Resilience Theory and Research on Children and Families: Past, Present, and Promise." *Journal of Family Theory & Review* 10: 12-31 <https://doi.org/10.1111/jftr.12255>.
- Miller, Kenneth E., and Andrew Rasmussen. 2010. "War Exposure, Daily Stressors, and Mental Health in Conflict and Post-Conflict Settings: Bridging the Divide Between Trauma-Focused and Psychosocial Frameworks." *Social Science & Medicine* 70 (1): 7-16. <https://doi.org/10.1016/j.socscimed.2009.09.029>.
- Peltonen, Kirsi, Samir Qouta, Marwan Diab, and Raija-Leena Punamäki. 2014. "Resilience among Children in War: The Role of Multilevel Social Factors." *Traumatology* 20 (4): 232-40. <https://doi.org/10.1037/h0099830>.

DEVELOPING A RESILIENCY MEASURE FOR ADOLESCENTS
IN EASTERN UKRAINE

- Rutter, Michael. 2006. "The Promotion of Resilience in the Face of Adversity." In *Families Count: Effects on Child and Adolescent Development*, edited by A. Clarke-Stewart and J. Dunn, 26-52 (The Jacobs Foundation Series on Adolescence). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511616259.003>.
- Rutter, Michael. 2012. "Resilience as a Dynamic Concept." *Development and Psychopathology* 24 (2): 335-44. <https://doi.org/10.1017/S0954579412000028>.
- Sagi-Schwartz, Abraham. 2008. "The Well Being of Children Living in Chronic War Zones: The Palestinian-Israeli Case." *International Journal of Behavioral Development* 32 (4): 322-36. <https://doi.org/10.1177/0165025408090974>.
- Shonkoff, Jack P., Andrew S. Garner, Benjamin. S. Siegel, Mary I. Dobbins, Marian F. Earls, Andrew. S. Garner, Laura McGuinn, John Pascoe, and David L. Wood, et al. 2011. "The Lifelong Effects of Early Childhood Adversity and Toxic Stress." *Pediatrics* 129 (1): e232-46. <https://doi.org/10.1542/peds.2011-2663>.
- Tol, Wietse, Suzan Song, and Mark J. D. Jordans. 2013. "Resilience and Mental Health in Children and Adolescents Living in Areas of Armed Conflict: A Systematic Review of Findings in Low- and Middle-Income Countries." *Journal of Child Psychology and Psychiatry* 54 (4): 445-60. <https://doi.org/10.1111/jcpp.12053>.
- Ungar, Michael. 2008. "Resilience across Cultures." *British Journal of Social Work* 38 (2): 218-35. <https://doi.org/10.1093/bjsw/bcl343>.
- Ungar, Michael. 2011. "The Social Ecology of Resilience: Addressing Contextual and Cultural Ambiguity of a Nascent Construct." *American Journal of Orthopsychiatry* 81 (1): 1-17. <https://doi.org/10.1111/j.1939-0025.2010.01067.x>.
- Ungar, Michael. 2013. "Resilience, Trauma, Context, and Culture." *Trauma, Violence, & Abuse*, 14 (3): 255-66. <https://doi.org/10.1177/1524838013487805>.
- Ungar, Michael. 2016. *The Child and Youth Resilience Measure (CYRM) Child Version: Users Manual*. Halifax, NS: Resilience Research Centre.
- UNICEF. 2018. *Ukraine Research on Domestic and Gender-Based Violence*. New York: UNICEF.

- United Nations Office of the High Commissioner for Human Rights. 2019. *Report on the Human Rights Situation in Ukraine 16 August to 15 November 2019*. Geneva: UNHCR. https://www.ohchr.org/Documents/Countries/UA/28thReportUkraine_EN.pdf.
- United Nations Refugee Agency. n.d. *Ukraine Refugee Crisis*. Geneva: UN Refugee Agency. <https://www.unrefugees.org/emergencies/ukraine/>.
- United Nations Refugee Agency: Ukraine. 2019. *Registration of Internal Displacement*. Geneva: UN Refugee Agency. http://bit.ly/IDPs_UA.
- van de Vijver, Fons J. R., and Leung, K. 2000. "Methodological Issues in Psychological Research on Culture." *Journal of Cross-Cultural Psychology* 31 (1): 33-51. <https://doi.org/10.1177/0022022100031001004>.
- van Rensburg, Angelique, Linda C. Theron, and Michael Ungar. 2019. "Using the CYRM-28 with South African Young People: A Factor Structure Analysis." *Research on Social Work Practice* 29 (1): 93-102. <https://doi.org/10.1177/1049731517710326>.
- Weller, Susan, and A. Kimball Romney. 1988. *Systematic Data Collection*. Newbury Park, CA: SAGE.
- Werner, Emmy E. 1996. "Vulnerable but Invincible: High Risk Children from Birth to Adulthood." *European Child & Adolescent Psychiatry* 5: 47-51. <https://doi.org/10.1007/BF00538544>.
- Werner, Emmy E. 2012. "Children and War: Risk, Resilience, and Recovery." *Development and Psychopathology* 24 (2): 553-58. <https://doi.org/10.1017/s0954579412000156>.
- Windle, Gill, Kate Mary Bennett, and Jane P. Noyes. 2011. "A Methodological Review of Resilience Measurement Scales." *Health and Quality of Life Outcomes* 9 (1): 8. <https://doi.org/10.1186/1477-7525-9-8>.
- World Health Organization. 2016. "Bringing Mental Health Care to Conflict-Affected Ukraine." <https://www.who.int/news-room/feature-stories/detail/bringing-mental-health-care-to-conflict-affected-ukraine>.

DEVELOPING THE GROUP FACILITATION ASSESSMENT OF COMPETENCIES TOOL FOR GROUP-BASED MENTAL HEALTH AND PSYCHOSOCIAL SUPPORT INTERVENTIONS IN HUMANITARIAN AND LOW-RESOURCE SETTINGS

GLORIA A. PEDERSEN, MANASWI SANGRAULA, PRAGYA SHRESTHA, POOJA LAKSHMIN, ALISON SCHAFER, RENASHA GHIMIRE, NAGENDRA P. LUITEL, MARK J. D. JORDANS, AND BRANDON A. KOHRT

ABSTRACT

In humanitarian settings, mental health and psychosocial support services (MHPSS) are often delivered in group-based formats. Group interventions enable providers to reach more individuals when resources and technical expertise are limited. Group-based programs also foster social support, empathy, and collective problem-solving among the participants. To remedy the current lack of tools available to assess the group facilitation competencies of individuals delivering group-based MHPSS, we made it our objective to develop such a tool. Our approach, which focused on adults, complimented a similar initiative underway for children and adolescents. We reviewed MHPSS manuals to identify key group facilitation competencies, which include developing and reviewing group ground rules, facilitating participation among all group members, fostering empathy between members, encouraging collaborative problem-solving, addressing barriers to attendance, time management, and ensuring group confidentiality. We then developed the Group Facilitation Assessment of Competencies Tool (GroupACT) The GroupACT is a structured observational tool for assessing these competencies during standardized role-plays with actor clients, or in vivo during the delivery of group sessions with actual

Received May 11, 2020; revised December 16, 2020; accepted August 2, 2021; electronically published December 2021

Journal on Education in Emergencies, Vol. 7, No. 2

Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).

ISSN 2518-6833

clients. We conclude this article with guidance for using the GroupACT to assess facilitators' competencies in providing group-based MHPSS in the health, education, protection, and other sectors in humanitarian settings.

INTRODUCTION

At the start of the COVID-19 pandemic, 168 million people around the globe needed humanitarian assistance and protection due to war, forced displacement, natural disasters, and other crises. This represents 1 in 45 people, the highest figure in decades (OCHA 2019), and the number has increased due to the COVID-19 pandemic (United Nations 2020). Among the most serious effects of conflict, pandemics, and other humanitarian crises is how they disrupt people's mental health and psychosocial wellbeing (Charlson et al. 2019; Lahiri, van Ommeren, and Roberts 2017). Recent reviews of psychosocial support programs in humanitarian emergency settings and low-and-middle-income countries (LMIC) suggest that adults who experience depression and posttraumatic stress disorder as a result of humanitarian crises can be helped by receiving mental health and psychosocial support services (MHPSS) (Bangpan, Felix, and Dickson 2019; Barbui et al. 2020). However, the research on sustainable and scalable approaches for such treatment in both ongoing and post-humanitarian contexts is limited.

When offering aid during humanitarian crises, many actors (e.g., international nonprofit organizations, regional organizations, the private sector, governments, community and religious groups) use a group modality to deliver programs and interventions. Group-based interventions have the potential to be more cost-effective and scalable than individualized services. For example, groups are an effective modality for offering protective spaces in which to provide education in humanitarian settings. They also can be a successful vehicle for a variety of other activities that address a multitude of community needs, including health care, social support, livelihoods, sports, and other recreation (Wood and Kallestrup 2018). Participants in group-based initiatives have identified a number of advantages, including an increased feeling of social togetherness, family bonding, and having a safe space in which to share feelings (Dickson and Bangpan 2018). Moreover, in humanitarian settings, it is often community members and institutions, such as educators in formal and informal education programs, livelihood training programs, and health programs, that act as frontline service providers and deliver services in a group-based format (Galappatti and Richardson 2016; Hendrickx et al. 2019). For example, building educators' capacity to "provide life-saving knowledge, skills and psychosocial support to those

affected by crisis” (UNESCO and Bokova 2017, 19) is the third strategic goal of the UNESCO Strategic Framework for Education in Emergencies.

To tackle the global mental health burden, nonspecialists (e.g., community health workers, teachers, organization staff, community members, professionals without formal or with limited mental health training) are increasingly being trained to deliver interventions, including acting as group facilitators across humanitarian and low-resource settings (Kohrt, Asher et al. 2018). This includes facilitators who are delivering and protecting education in humanitarian settings, such as child-friendly spaces that support education, physical activity, and children’s psychosocial needs (Hermosilla et al. 2019; Save the Children 2009), and early childhood development programs that deliver psychosocial support to distressed caregivers (Murphy and Hutton 2018). Educators in humanitarian settings often are the first and consistent contact with young people, and therefore, they are an important group to be trained in MHPSS skills (IASC Reference Group for Mental Health and Psychosocial Support in Emergency Settings 2017).

Increasing evidence shows that trained and supervised nonspecialists can provide support effectively and deliver brief psychosocial and psychological interventions in humanitarian and low-resource settings (Barbui et al. 2020; Pedersen et al. 2019; Singla et al. 2017). This includes individual interventions (e.g., one-on-one sessions between a facilitator and a person seeking services) and group-based interventions (e.g., one or two facilitators with groups of three or more people who have a common experience or need or who live in the same community, including family members). As such, increasing attention is being given to the competencies nonspecialists require to deliver MHPSS successfully. Competencies refers to the key observable skills needed to deliver care effectively (Fairburn and Cooper 2011), which can be assessed through structured role-plays, with actors playing the role of clients (Ottman et al. 2020).

With the aim of improving the quality of the implementation of psychological and psychosocial support interventions worldwide, the World Health Organization (WHO) is developing a platform called Ensuring Quality in Psychological Support, or EQUIP (see www.who.int/mental_health/emergencies/equip/en/). Once launched, EQUIP will be an open-access online platform offering resources to help trainers, supervisors, and organization staff members facilitate competency-based training, including easy-to-use rating tools for assessing and evaluating competency (Kohrt et al. 2020). Tools to measure the competency of nonspecialists who deliver psychological and psychosocial care to adults have already been developed and implemented in humanitarian settings, including a

tool developed to assess the common factors of psychological support, such as nonverbal and verbal communication. This tool, the Enhancing Assessment of Common Therapeutic factors (ENACT) rating scale, is typically used in one-on-one role-play scenarios (Kohrt, Jordans et al. 2015; Kohrt, Mutamba et al. 2018; Kohrt, Ramaiya et al. 2015). The WeACT tool, which builds on the ENACT, is a competencies assessment tool tailored to nonspecialists and educators who are delivering care to children and adolescents in the child protection, education, and mental health and psychosocial support sectors (Jordans, Coetzee et al. 2021).

Facilitators of group-based interventions use a set of processes and skills that differs from those used by providers of individual services. For example, to avoid one member dominating a session, group facilitators must manage time and turn-taking effectively. Also, to ensure confidentiality in group sessions, facilitators must require members to agree not to share another member's stories or experiences with anyone outside the group. It is essential to apply a minimum competency standard for group facilitators to ensure the quality of the adult group-based interventions or support programs provided by nonspecialists in humanitarian settings. However, there is a gap in the resources available for assessing the competencies a group facilitator needs in order to achieve the desired outcome for their group.

To meet this need, in this study we drew from the group-format psychological and psychosocial intervention literature to identify the competencies that are fundamental to the facilitation of group-based interventions, and which differ from the processes used in individual-based interventions. We also developed the Group Facilitation Assessment of Competencies Tool (GroupACT), which supports the evaluation of group facilitator competencies.¹ Here we describe the development of the GroupACT and lay out an agenda for future work to determine its feasibility, acceptability, and psychometric properties for use in developing competency in MHPSS.

METHODS

The GroupACT was developed using a three-step process: (1) creating a competency codebook, which involved conducting a literature search to obtain content for the codebook and codes; (2) coding of competencies, which involved applying the final competency codebook to group-based psychological and psychosocial intervention manuals; and (3) generating items, which involved using the content extracted

¹ Examining its psychometric properties is a future step for the GroupACT, as it is beyond the scope of this field note. To date, the authors have not assessed the properties.

from the competencies coding process to define and describe the competency assessment items to be used in the GroupACT. The following methods describe the development of the codebook, the coding of the competencies, and the generation of items as they specifically relate to the development of the GroupACT. Further details on the full methodology, including the identification and inclusion of psychological and psychosocial manuals for this review, can be found in a separate publication (Pedersen et al. 2020).

CREATING A COMPETENCY CODEBOOK

In April 2018, to identify content for the competency codebook and codes, we conducted a literature search for current published competencies and competency frameworks, as well as a systematic review and a literature review for competencies related to the delivery of psychological and psychosocial interventions in LMIC. We conducted searches in Google Scholar, PubMed, and PsycINFO to identify sources.²

The literature search produced nine global competency frameworks and published competencies (Pedersen et al. 2020): (1) Cognitive Therapy Scale-Revised (Blackburn et al. 2000); (2) the ENACT rating scale and its evidenced resources for item generation (Kohrt, Jordans et al. 2015; Kohrt, Ramaiya et al. 2015); (3) the e-Problem Solving Therapy training and assessment tool (Cartreine et al. 2012); (4) the Global Social Service Workforce Alliance (2017) competencies; (5) the Improving Access to Psychological Therapies competence framework (National Collaborating Centre for Mental Health 2019); (6) the Let's Get Talking Practice Support, which includes competencies, training, and supervision for the delivery of talking therapies (Te Pou o te Whaakaro Nui 2016); (7) the Motivational Interviewing Rating Scale, which is based on Miller and colleagues' (2003) manual for motivational interviewing; (8) the PracticeWise Psychosocial and Combined Treatments Coding Manual (Chorpita, Daleiden, and Weisz 2005, 2008); and the and Yale Adherence and Competence Scale guidelines (Carroll et al. 2000; Nuro et al. 2005). We then did an additional search specifically for "group facilitation," "group competency," and "group competencies." We did not identify any group facilitation competency assessment tools or frameworks in the initial or secondary search.

2 Search terms we used for PubMed: review AND competenc* and mental; search terms for PsycINFO: review AND subject: competenc* AND TX mental AND subject: health worker* OR lay health worker OR mental health worker OR health professional; and review AND SU: competenc* AND TX mental. In Google Scholar, we also included the term "competency framework."

CODING COMPETENCIES AND GENERATING ITEMS

CODING COMPETENCIES

We transferred the competency codebook and psychological and psychosocial manuals into NVivo 12 (QSR International 2012). Two coders (GAP, PL) piloted the codebook and assessed interrater reliability, with discrepancies being resolved by a third coder (BAK). The same manual was coded independently by each coder and repeated until 80 percent interrater reliability was established. We then divided and coded 17 manuals separately; if the manual sections covered multiple competencies, double coding was permitted. Codes frequently double coded were collapsed. We recorded process notes on a shared Google Doc to aid discussion and agreement around coding for each manual. The code “group facilitation skills” was iteratively added during the coding process, as eight of the 17 (47%) manuals clearly denoted the importance of adhering to group facilitation competencies and guidelines in order to run a group session successfully. We applied the group facilitation skills code generously to the text, and captured any content (e.g., words with “group”), techniques, and instructions related to group facilitation, using a minimum of 5-8 sentences and a maximum of 1 page per coding reference to ensure that we obtained all relevant information related to the skill. Codes were exported from NVivo to create summary descriptions, for each manual, of the “group facilitation skills” code content.

GENERATING ITEMS

We distilled the “group facilitation skills” code summaries into subthemes to support the generation of competency assessment items. Next, we applied our “group facilitation skills” code to 973 tool items extracted from a separate competency tool item review (Ottman et al. 2020); we identified only 4 of the 973 tool items related to group facilitation. Three items were from the Mindfulness-Based Interventions Teaching Assessment Criteria (Crane et al. 2016) and one from the Fidelity of Implementation Rating System (Knutson et al. 2009). These items were overarching group facilitation concepts that were typically used to assess a facilitator’s achievement after the implementation of an intervention or sessions was complete (e.g., raters assessed recorded sections of completed sessions). They included group management and active teaching themes, including teacher facilitation skills such as “breaks into teachable units,” “guides review of material,” “guides practice in a way that makes the key learning available to participants,” and “careful management of issues such as group rules, boundaries and confidentiality, but which is simultaneously a place in which participants can

explore and take risks.” Due to the specific nature of these tools, and to ensure successful psychometric evaluation of the GroupACT in the future, we extracted the descriptions of these four items and added them to our subthemes to refine our generation process, rather than using the items directly in the tool. Finally, a field team in Nepal who were developing and adapting a local group version of the Problem Management Plus (Group PM+) manual for adults (Sangraula et al. 2020) and Group Interpersonal Therapy (Group IPT) for adolescents (Rose-Clarke et al. 2020) reviewed and refined all items proposed for use in the GroupACT.

RESULTS

Eight of the 17 psychological and psychosocial intervention manuals addressed skills related specifically to group facilitation, all of which provided content for the GroupACT. The final list of manuals included

1. Cognitive Processing Therapy (CPT) for people dealing with mental health problems following traumatic events (e.g., rape, torture, combat) (Bass, Bolton et al. 2013; Bass, Annan et al. 2013);
2. Caregiver Skills Training (CST) to support families living in LMIC who have children with developmental disorders (Hamdani et al. 2017; WHO 2017);
3. Friendship Bench, an approach to treating common mental disorders in low-resource settings (Chibanda et al. 2011; Chibanda et al. 2016, Singh 2017);
4. Happy Families Program, a family skills intervention to support displaced Burmese migrant families, which could be adapted to multiple settings (Annan et al. 2017; Puffer et al. 2017);
5. Group IPT, which is a group-based intervention that had been evaluated for adults and adolescents for treatment of depression, with studies conducted in northern Uganda (Bass et al. 2006; Bolton et al. 2007; Bolton et al. 2003; Mutamba, Kane et al. 2018; Mutamba, Kohrt et al. 2018; Verdeli et al. 2008; Verdeli et al. 2016);
6. Group PM+, a group version of the Problem Management Plus intervention for adults in humanitarian settings; PM+ incorporates stress management, behavioral activation, problem solving, and strengthening social support (Dawson et al. 2015; Sangraula et al. 2020; Jordans, Kohrt et al. 2021);

7. Parenting Program Uganda, a parenting skills intervention to support healthy child development, including psychosocial and nutritional needs (Singla and Kumbakumba 2015; Singla, Kumbakumba, and Aboud 2015); and
8. Self-Help Plus, designed based on acceptance and commitment therapy principles for people with high levels of stress and psychological distress, particularly those living in humanitarian settings (Brown et al. 2018; Epping-Jordan et al. 2016; Tol et al. 2018; Tol et al. 2020).

Table 1 displays manual titles, abbreviations, and brief descriptions of the intervention for the manuals that had an element of group-based delivery.

Table 1: Mental Health and Psychosocial Support Intervention Manuals Including a Group-Delivery Format

Manual	Abbreviation	Intervention Description
Cognitive Processing Therapy	CPT	Cognitive Processing Therapy aims to support people with mental health problems following traumatic events, including rape, domestic violence, combat, torture, and child sexual abuse. This manual was created for delivery by nonspecialists in the Democratic Republic of the Congo. Additional information on CPT is available online: https://www.jhsph.edu/research/centers-and-institutes/global-mental-health/our-projects/by-intervention/
Caregiver Skills Training	CST	Caregiver Skills Training is an open-access program that supports families of children with developmental delays or disorders, including intellectual and pervasive developmental disorders (e.g., autism), and it may be implemented in LMIC. The program uses a family-centered approach and is designed to be delivered by nonspecialists (nurses, community-based workers, or peer caregivers) as part of a network of health and social services for children and families. WHO CST materials are currently being assessed and will be made available, pending the evaluation results; description of the WHO CST program development can be found online: https://www.frontiersin.org/articles/10.3389/fpsy.2019.00769/full

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

Manual	Abbreviation	Intervention Description
Friendship Bench	FB	The Friendship Bench is a transdiagnostic treatment that uses cognitive behavioral therapy techniques specifically related to problem-solving and behavioral activation. It is written in English and Shona and was created to be supported by the use of tablets. FB materials are freely available online: https://www.friendshipbenchzimbabwe.org/
Group Interpersonal Therapy	Group IPT	The WHO Group Interpersonal Therapy adapts traditional individual IPT therapy into a simplified version designed for group treatment of depression in a variety of settings. The therapy covers four main problem areas that are common to individual IPT, including grief, disputes/conflict, life changes, and loneliness/isolation. This model teaches that one or more of these problem areas can trigger depression. WHO Group IPT materials are available online: https://www.who.int/mental_health/mhgap/interpersonal_therapy/en/
Group Problem Management Plus	Group PM+	The WHO Group Problem Management Plus intervention is a five-session group therapy for adults in humanitarian settings. The intervention includes stress management, behavioral activation, problem solving, and strengthening social support. The intervention has been evaluated in Pakistan and Nepal. The manual is available online: https://www.who.int/publications/i/item/9789240008106
Happy Families	HF	The Happy Families caregiver and family skills intervention is meant for children ages 7 to 15 and their caregivers. It was adapted from the Strengthening Families program and includes topics on parenting and skills for better family functioning. It was developed for implementation with displaced Burmese migrant families living in Thailand. Additional information on HF is available online: https://www.rescue.org/report/building-happy-families-irc-research-brief

Manual	Abbreviation	Intervention Description
Parenting Program Uganda	PPU	<p>The Parenting Program Uganda is aimed at encouraging parents to adopt and practice parenting skills that support their children’s healthy development (“develop into strong, healthy, and smart people”). It comprises five main messages: (a) diversify the child’s diet with animal-source foods and provide three to four meals daily; (b) hand-wash with soap and use latrines; (c) engage in two-way talk with the child, using pictures; (d) provide play materials in the home; and (e) love and respect yourself, your child, and your spouse.</p> <p>More information on PPU is available online: https://plan-international.org/publications/parenting-impact-study-lira-uganda</p>
Self-Help Plus	SH+	<p>Self-Help Plus was developed to help people with high levels of stress and psychological distress (e.g., symptoms of depression, anxiety), especially in areas where there are many people needing support (i.e., a humanitarian setting), or where there are difficulties in the provision and/or supervision of psychological interventions. Facilitators and cofacilitators use prerecorded audio, pictures, and support materials to conduct each session.</p> <p>WHO SH+ materials are currently being assessed and will be made available pending results. The Juba Arabic version for use in South Sudan is available on request.</p>

The item-generation process resulted in eight competencies related to group facilitation. Table 2 lists the items, their respective descriptions, and corresponding sample content from the manuals that supported the selection and refinement of items. Although the manuals tended to emphasize similar competencies, the manner in which they expressed it occasionally varied. For example, the CST, CPT, and Group IPT manuals all emphasized fostering empathy among group members but with different strategies. In CST, facilitators fostered empathy by encouraging other members to praise and support each other for handling both positive and challenging experiences. In CPT, facilitators were encouraged to model empathic responses: “Hearing a group leader acknowledge a comment or respond with encouragement to something a group member says can be very helpful for group members.” In Group IPT, the facilitator was expected to build a “feeling of closeness between members of the group,” which included recommendations for nonverbal communication skills: “You should show understanding and help

to build this closeness using body language such as nodding and showing interest while group members are talking.” The manuals also presented different strategies for managing participants who monopolized a discussion, such as this from the CPT manual:

Ask group members who are quick to answer a question or make a comment to count to 10 before they talk so other members can voice their thoughts. If necessary, group leaders can ask that once a group member has participated three times they wait until other group members speak before they add to the discussion. These suggestions should be made to the whole group so that one member is not singled out or embarrassed.

The Friendship Bench uses a bell to help participants pay attention and support taking turns: “If someone is talking for too long the bell can be rung . . . If people interrupt or talk amongst themselves the bells will be rung to establish order.”

The competencies were then organized into a format that could be scored by observing either role-plays, or in vivo sessions with actual clients. Four levels, which follow the ENACT scoring framework and incorporate feedback from practitioners and researchers implementing group- and individual-based interventions, are included in the GroupACT: Level 1: “Any unhelpful behavior”; Level 2: “No basic skills or some but not all basic skills”; Level 3: “All basic skills”; and Level 4: “All basic helping skills plus any advanced skill.” This structure, which also aligns with similar competency tools being implemented on the WHO EQUIP platform, will help to harmonize reporting if organizations use multiple competency tools for their programs. The final format of the GroupACT assessment tool can be found in the Appendix.

Table 2: GroupACT Competency Items (n=8), Item Descriptions, and Corresponding Manual Themes and Samples

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
<p>Develop group guidelines and ground rules collaboratively (first session)</p>	<ul style="list-style-type: none"> • Elicit ground rules for the group while attending to the groups’ cultural and religious practices • Establish guidelines with the group in the initial session and implement ground rules • Elicit feedback from group members, making sure to ask for agreement on the guidelines and adjusting rules depending on group needs 	<p>Instructions for developing or reviewing group guidelines was identified across the group-based manuals. Typically described as “guidelines,” “ground rules,” or “rules,” sample instructions cover attendance, what the sessions will and will not offer (e.g., material goods), and the general structure and length of the program. For example, the Group IPT manual states, “Cover group rules: During the first group session, facilitators should discuss not giving material goods, attendance and dropping out of the group,” and give instructions for facilitators to cover the overall structure, length, and format of the sessions. Other manuals, such as the PPU and FB manuals, follow a similar structure that offers instructions to standardize expectations and encourage a sense of respect among the group members. The Group PM+ manual encourages the group members to participate by making group guidelines that will help them feel comfortable in the group setting, “Ask participants to suggest rules: What are other important rules to help you feel comfortable participating in the group?”</p>

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
Review group guidelines and ground rules (subsequent sessions)	<ul style="list-style-type: none"> • Review guidelines at the beginning of each session, determine whether all group members are still in agreement, elicit feedback, and adjust rules in accordance with the context or need • Remind members of ground rules and ensure that they are clear and agreed to • Address rule violations with members individually when appropriate 	<p>Reviewing group guidelines before the beginning of each session was recommended in all the manuals. For example, the CST manual includes instructions to introduce and review group guidelines before each session and insists that facilitators start each session by “asking all participants if they agree with the group guidelines.”</p>
Encourage participation of all group members	<ul style="list-style-type: none"> • Encourage all members to discuss and be involved in sessions • Provide reflection and support a sense of belonging for members • Consolidate group members’ learning • Use techniques such as turn-taking to ensure that each member has an opportunity to speak and share • Attend to any literacy, numeracy, or technical skills so that all members have an equal opportunity to participate 	<p>The manuals highlighted managing group participation in a variety of ways. For instance, the Group IPT manual has the instructions, “You should not force anyone to speak. However, gentle prompting can be fine.” The FB manual suggests using bells to remind group members to take turns: “If someone is talking for too long the bell can be rung . . . If people interrupt or talk amongst themselves the bells will be rung to establish order.” The CPT manual suggests to readers that “questions or making statements that point out common problems [is] . . . one of the best ways for the group leader to encourage group members to share with each other.”</p>

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
<p>Foster empathy between group members</p>	<ul style="list-style-type: none"> • Encourage members to display their understanding of their peers' experiences • Prompt members to relate to one another and discuss how they are feeling during sharing • Exemplify empathy for others to follow, verbally and nonverbally • Provide summarizing statements and interpretations about members' emotional or situational similarities, but do not make these connections if the member has only shared these feelings in a private discussion • Recognize members who have displayed empathetic behavior in group by offering encouragement 	<p>Fostering empathy between group members to promote positive sharing of feelings and emotions during sessions was distinct in most of the manuals. For example, the Group IPT manual instructs facilitators to bring "a feeling of closeness between members of the group," and it includes example scripts for facilitators to follow: "I noticed that many of you shared a similar reaction to Rita's story. It sounds like you would have felt similarly in that situation. Rita, how does it make you feel that others had feelings similar to yours?"</p>

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
<p>Guide collaborative problem-solving among group members</p>	<ul style="list-style-type: none"> • Guide group members to share problem-solving ideas, and encourage members to praise and support each other for both positive and challenging experiences • Facilitate groups so they can work together to address barriers to problem-solving while normalizing barriers • Work with the group to eliminate unrealistic solutions and identify solutions that are timely, realistic, and attainable, and consider potential challenges/barriers that may arise • Suggest that members find a group partner to discuss solutions and to check in on one another for support 	<p>In group interventions or support groups, it is often the members’ commonalities that bring them together and help them to feel less alone by sharing a similar problem or feeling (Dickson and Bangpan 2018). We identified a few manuals that include instructions for facilitators to encourage group problem-solving. For example, the CST manual has instructions for facilitators to ask questions such as, “Did anyone try a different activity than they planned? Why?” The Group IPT manual has similar instructions, offering scripted examples for facilitators to follow, such as, “FACILITATOR: Mary, I recall you mentioned to the group that you have a problem similar to Jasmine’s. She just told us that she has tried everything to make things better. Would you like to tell us a little about your struggle and what you’ve tried? If not, that is OK too.” In the case examples of the Group PM+ manual, during the Managing Problems strategy, the facilitator encourages group members to brainstorm possible solutions for solving the scenario character’s problem.</p>

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
Mitigate barriers to attendance	<ul style="list-style-type: none"> • Address barriers to members’ attendance, including location and time of sessions • Provide a safe, comfortable, and calm setting, with a comfortable temperature, minimal noise, and privacy—if setting is outdoors, it should be covered • Actively address potential barriers, such as lack of childcare or nursing, employment schedule, transportation, disabilities, religious observances, physical health, menstrual practices, etc., and gather information from members • Engage the group in brainstorming ways that all members can attend sessions • Schedule or adjust sessions to accommodate most group members and encourage members to support each other’s attendance, while respecting limits to confidentiality (e.g., shared childcare, traveling together, etc.); this includes participation via phone/virtually • Update members on logistical changes to sessions in a timely manner; when appropriate, encourage attendance even if a member cannot make every session 	<p>Facilitators should promote accessibility to group sessions whenever feasible. This can be done in a number of ways, including respecting the group members’ cultures. For example, the SH+ manual includes instructions for facilitators to attend to the groups’ cultural and religious practices to ensure accessibility for all members, and, “if possible, [to have a space] with a good temperature, minimal noise, and privacy.” This may be challenging in some settings, particularly in contexts where indoor space is difficult to locate. If outdoors, it is suggested that facilitators should attempt to meet under cover (SH+ manual), such as in the shade of a tree or in tents. Reminding group members of sessions and updating them as soon as possible if a meeting location or time has changed is also important. For instance, during the SH+ pilot with South Sudanese refugees in Uganda, facilitators visited group members’ homes a day prior to the intervention to remind them of the group session. The Group PM+ manual recommends addressing the challenges to attend the group; this is done through an activity in the first session, “Reasons for joining Group PM+ (advantages) and Challenges to join Group PM+ (disadvantages).</p>

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
<p>Ensure confidentiality among group members</p>	<ul style="list-style-type: none"> • Explain what confidentiality is and outline when facilitators may break confidentiality • Elicit from the group how to manage confidentiality, including an agreed to response to what to do if members see each other outside of sessions • Address confidentiality issues when unexpected or uninvited people come to sessions • Explain how confidentiality relates to issues of respect, the importance of valuing each other's experiences, and the feelings associated with violation of one's confidentiality • Address when confidentiality is broken by other members during group without targeting or blaming group members 	<p>Group confidentiality is often addressed in the manuals along with the group ground rules. For example, the PPU manual includes “Confidentiality” as one of the four keywords that facilitators are instructed to write on a flipchart when establishing group ground rules. Instructions from other manuals include having facilitators tell groups that “we will respect each other’s privacy, so don’t share things you learn about other families outside of this room, especially their private challenges or problems” (CST manual); and to “remind group members at the start of each session (and if needed during a session) that their conversations will remain confidential and that they have a right to privacy” (Group IPT manual). The Group PM+ manual also addresses the three aspects to confidentiality: 1. Breaching confidentiality when a group member’s life is believed to be at risk; 2. The facilitator sharing group discussions with her supervisor to assure the best care; and 3. The responsibility of the group facilitator and the participants for maintaining confidentiality outside of the group setting.</p>

GroupACT Competency	Competency Item Description	Corresponding Manual Subthemes and Samples
<p>Manage time effectively with breaks, energizers, and pacing</p>	<ul style="list-style-type: none"> • Demonstrate punctual timekeeping, including having and reviewing a schedule of activities for sessions with group • Include adequate time for group members to ask questions • Monitor time and communicate with members about changes to the schedule • Consolidate participant learning and pace activities • Give appropriate, timed breaks between activities, with instructions to signal the beginning and end of breaks • If sessions are >45 minutes, facilitators should do an energizer or other activity to keep group members engaged 	<p>The manuals recognize the importance of time management, breaks, and brief activities to keep groups energized, particularly for long sessions or multiple sessions within a day. For example, the CST manual offers a sample schedule of activities for a 2.5-hour session, including a 15-minute break. Similarly, the SH+ manual has 10-minute breaks throughout the sessions, with the option of running an energizer during this time to keep group members engaged. To support time management, facilitators can use tools, such as a bell that signals when breaks or activities have ended or time is running out (SH+ manual; FB manual). The Group PM+ manual contains session plans with time allocation and breaks during all 5 sessions.</p>

DISCUSSION

Many group-based interventions and programs are delivered in humanitarian and other low-resource settings. A recent umbrella review (Barbui et al. 2020) reports significant evidence for nonspecialists to effectively deliver psychological and psychosocial interventions in for adults with depression and posttraumatic stress disorder in humanitarian settings. It advises that future programs focus on ethical and sustainable implementation approaches, such as group-based programming and increasing nonspecialists' capacity to deliver quality care. Moreover, the IASC Framework for MHPSS in Emergency Settings suggests that nonspecialists, such as teachers and other educators, can deliver appropriate focused care, which could provide people in humanitarian settings who are

dealing with mental health problems with access to supportive care (e.g., help in improving functionality, increasing social supports, reducing symptoms, etc.) (IASC 2017). To make progress in this direction, we developed the GroupACT to support the assessment of competencies relevant to group facilitation, with the aim of creating a minimum competency standard for group facilitators who deliver psychological and psychosocial support. This tool includes eight items that address key group-facilitation competencies identified in the eight psychological and psychosocial group-based manuals mentioned above.

Table 3 offers a summary of potential applications for the GroupACT. A systematic approach to adapting the GroupACT culturally and contextually is recommended when facilitating such groups as an international nongovernmental organization. The transcultural translation and adaptation procedure used in global mental health field work offers an adaptation of the tool in five domains: comprehensibility, acceptability, relevance, completeness, and technical equivalence (Van Ommeren et al. 1999). An organization may choose to include key community stakeholders in supporting the adaptation process.

It is preferable to adapt the role-play scripts to represent typical members of the group program. For example, a group-based MHPSS program supporting women 30-40 years old who are in a state of distress would adapt role-play scripts to represent these women's characteristics. Implementing groups can train raters and actors from their organization (e.g., trainers, supervisors, program managers, research assistants) or recruit local community members who are interested in the work and may benefit from capacity-building. Organizations can video- or audio-record standardized role-plays using staff members who have experience working with the target population.

These videos can support rater training and agreement, and help the organization establish a training library of GroupACT materials in their local programming office. Role-play assessments with trainees or group facilitators can be implemented in person or remotely (e.g., telephone, video conferencing) and, when possible, it is preferable to record the video or audio of an assessment so that ratings can be completed in a timely manner for the implementing team. To support the development of group facilitation competencies and give the trainees tailored feedback, the trainer or supervisor may use a competency-based approach and assess competencies during training. This can be done using typical role-play practice during a training, such as observing peer-to-peer role-plays or having a trainer play a part in the practice role-plays. With “on-the-go” ratings, the trainer or supervisor has the option of focusing on select competencies specific to that session, or on those that may be more advanced or take longer to build capacity in. Immediate feedback is then available to support skills strengthening—including adapting the training as needed—and to address any potentially harmful behaviors in a controlled environment before facilitators deliver real-world group-based interventions. GroupACT assessments may be implemented pre- and posttraining and postsupervision for a range of objectives: gauge the effectiveness of a training program; guide selection of potential facilitators; and identify trainees’ competencies that need remediation.

Table 3: Suggested Applications for the Group Facilitation Assessment of Competencies Tool (GroupACT) for Research and Implementation

When	Why	How	Raters	Actors
Pretraining assessment	Assess trainees' competency levels to inform training plan; record a baseline of competency to track trainees' progress and performance across training; guide organizations or program managers in the selection of nonspecialists to participate in training	<i>Modality:</i> Standardized role plays with mock group members <i>Formats:</i> Live observations, video recordings, audio recordings, transcripts	Trainers with experience in group facilitation and/or GroupACT; trained external raters with experience in group facilitation and/or GroupACT	Trainers with experience delivering group-based interventions; other organizational staff (e.g., research assistant) trained to play role; external actors (e.g., local actor troops) trained to play role
During training	Formally or informally track and record trainees' progress during training; measure maintenance or drift in skills; inform any needed adjustments to training activities	<i>Modality:</i> Periodic role-plays; single-competency role-plays <i>Formats:</i> Live observations, video recordings, audio recordings, transcripts	Trainers with experience in group facilitation and/or GroupACT; trained external raters with experience in group facilitation and/or GroupACT	Peer trainees; trainers with experience in delivery of group-based interventions
Posttraining assessment	Evaluate trainees to confirm minimum competency levels are met; compare pre- and postassessments to examine effectiveness of training program; inform remediation needs and activities; highlight supervision needs; inform selection of trainees' as future trainers, supervisors, or participants in intervention trials	<i>Modality:</i> Standardized role plays with mock group members <i>Formats:</i> Live observations, video recordings, audio recordings, transcripts	Trainers with experience in group facilitation and/or GroupACT; trained external raters with experience in group facilitation and/or GroupACT	Trainers with experience in delivery of group-based interventions; other organizational staff trained to play role; external actors trained to play role

When	Why	How	Raters	Actors
Supervision	Track improvement and maintenance of competencies over time; monitor group sessions to provide feedback	<p><i>Modality:</i> Standardized role plays with mock group members; with actual group members</p> <p><i>Formats:</i> Live observations, video recordings, audio recordings, transcripts</p>	Supervisors, including mental health experts or peers; trained external raters with experience in group facilitation and/or GroupACT	Peer nonspecialists with experience in group facilitation; supervisors, including mental health experts or peers; other organizational staff; external actors trained to play role

Psychometrics for the GroupACT must be established to determine whether the tool increases group facilitator competency. The reliability and validity of the items should be tested in multiple settings across a variety of group-based interventions and support programs, and feedback should be obtained from the group members about the facilitator's effectiveness and their perceptions of the quality of the care they received, and of the program they are participating in. The feasibility, acceptability, and perceived utility of this tool, including the number of competencies and items, should be assessed by facilitators, group members, supervisors, and implementing staff. The application of the GroupACT should extend beyond psychological and psychosocial interventions. It can be adapted and piloted with diverse populations, needs, and settings, including varied ethnic groups (mixed or single), languages, and religions. Minimum of criteria for group facilitator competency should be established for a range of domains: education, nutrition, gender-based violence, substance use, and microfinance programs.

When running group education in humanitarian settings, where people often are distressed, it is imperative that the educator has group facilitation skills and is able to address the social inclusion and emotional wellbeing of group members (e.g., fostering empathy and participation among the group members, establishing group confidentiality and problem-solving). Moreover, standardized assessment ratings and data collection will enhance the monitoring and evaluating of evidence-based capacity-building among humanitarian programs that are using group facilitation methods, including MHPSS, education, health, and finance, and will support national policy to strengthen capacity between education and psychosocial support systems (Jordans and Kohrt 2020).

Future researchers also could examine the tool's ability to inform training and supervision practices among those delivering group-based interventions and support programs. Using the GroupACT to assess facilitators' strengths preservice may enable trainers and supervisors to tailor their training approaches to the facilitators' abilities prior to the training and to continue using competency-based training techniques (Frank et al. 2010; Kohrt et al. 2020). Future research also could address the adequacy of these competences in diverse emergency settings, where and why group facilitation guidelines need to differ, what distinguishes "minimum" competences from other important ones in MHPSS group processes, and how well mainstreamed group process guidelines account for gender dynamics, cultural norms, and power relations in diverse emergency contexts.

CONCLUSION

To address the capacity-building needs of nonspecialists delivering group interventions and support programs in humanitarian and other low-resource settings, it is essential to identify which competencies are needed to facilitate groups appropriately and successfully. In response to this need, we have developed the GroupACT, a tool for evaluating nonspecialists' group facilitation skills that can be implemented using structured role-plays with group members. Further research is needed to establish the tool's psychometrics and test for acceptability, feasibility, and utility in multiple group-based interventions and support programs. The GroupACT could be adapted to and piloted in multiple contexts where support groups are addressing a variety of psychosocial, educational, monetary, and other community needs.

REFERENCES

- Annan, Jeannie., Amanada Sim, Eve S. Puffer, Carmel Salhi, and Theresa S. Betancourt. 2017. "Improving Mental Health Outcomes of Burmese Migrant and Displaced Children in Thailand: A Community-Based Randomized Controlled Trial of a Parenting and Family Skills Intervention." *Prevention Science* 18: 793-803. <https://doi.org/10.1007/s11121-016-0728-2>.
- Bangpan, Mukdarut, Lambert Felix, and Kelly Dickson. 2019. "Mental Health and Psychosocial Support Programmes for Adults in Humanitarian Emergencies: A Systematic Review and Meta-Analysis in Low and Middle-Income Countries." *BMJ Global Health* 4 (5): e001484. <https://doi.org/10.1136/bmjgh-2019-001484>.
- Barbui, Corrado., Marianna Purgato, Jibril Abdulmalik, Ceren Acarturk, Julian Eaton, Chiara Gastaldon, Oye Gureje, et al. 2020. "Efficacy of Psychosocial Interventions for Mental Health Outcomes in Low-Income and Middle-Income Countries: An Umbrella Review." *The Lancet Psychiatry* 7 (2): 162-72. [https://doi.org/10.1016/s2215-0366\(19\)30511-5](https://doi.org/10.1016/s2215-0366(19)30511-5).
- Bass, Judith. K., Jeannie Annan, Sarah McIvor Murray, Debra Kaysen, Shelly Griffiths, Talita Cetinoglu, Karin Wachter, Laura K. Murray, and Paul A. Bolton. 2013b. "Controlled Trial of Psychotherapy for Congolese Survivors of Sexual Violence." *New England Journal of Medicine* 368 (23): 2182-91. <https://doi.org/10.1056/NEJMoa1211853>.

- Bass, Judith, Paul Bolton, Debra Kaysen, Shelly Griffith, Sarah Murray, Talita Cetinoglu, Katie Robinette, Marie-France Guimond, Karin Wachter, and Jeannie Annan. 2013a. *Addressing Sexual Violence Related Trauma in Eastern DRC with Cognitive Processing Therapy*. Washington, DC: World Bank.
- Bass, Judith, Richard Neugebauer, Kathleen F. Clougherty, Helen Verdeli, Priya Wickramaratne, Lincoln Ndogoni, Liesbeth Speelman, Myrna Weissman, and Paul Bolton. 2006. "Group Interpersonal Psychotherapy for Depression in Rural Uganda: 6-Month Outcomes-Randomised Controlled Trial." *British Journal of Psychiatry* 188: 567-73. <https://doi.org/10.1192/bjp.188.6.567>.
- Blackburn, Ivy Marie, Ian Andrew James, Derek L. Milne, and F. Katharina Reichelt. 2001. "The revised Cognitive Therapy Scale-Revised (CTS-R) psychometric properties." *Behavioural and Cognitive Psychotherapy* 29 (4): 431-446. <https://doi.org/10.1017/S1352465801004040>.
- Bolton, Paul, Judith Bass, Richard Neugebauer, Helen Verdeli, Kathleen F. Clougherty, Priya Wickramaratne, Liesbeth Speelman, Lincoln Ndogoni, and Myrna Weissman. 2003. "Group Interpersonal Psychotherapy for Depression in Rural Uganda: A Randomized Controlled Trial." *Journal of the American Medical Association* 289 (23): 3117-24. <https://doi.org/10.1001/jama.289.23.3117>.
- Bolton, Paul, Judith Bass, Theresa Betancourt, Liesbeth Speelman, Grace Onyango, Kathleen F. Clougherty, Richard Neugebauer, Laura Murray, and Helen Verdeli. 2007. "Interventions for Depression Symptoms among Adolescent Survivors of War and Displacement in Northern Uganda: A Randomized Controlled Trial." *Journal of the American Medical Association* 298 (5): 519-27. <https://doi.org/10.1001/jama.298.5.519>.
- Brown, Felicity L., Kenneth Carswell, Jura Augustinavicius, Alex Adaku, Marx R. Leku, Ross G. White, Peter Ventevogel et al. 2018. "Self Help Plus: Study Protocol for a Cluster-Randomised Controlled Trial of Guided Self-help with South Sudanese Refugee Women in Uganda." *Global Mental Health* 5: e27. <https://doi.org/10.1017/gmh.2018.17>.

- Brown, Felicity L., Anne M. de Graaff, Jeannie Annan, and Theresa S. Betancourt. 2017. "Annual Research Review: Breaking Cycles of Violence. A Systematic Review and Common Practice Elements Analysis of Psychosocial Interventions for Children and Youth Affected by Armed Conflict." *Journal of Child Psychology and Psychiatry* 58 (4): 507-24. <https://doi.org/10.1111/jcpp.12671>.
- Carroll, Kathleen M., Charla Nich, Rachel L. Sifry, Kathryn F. Nuro, Tami L. Frankforter, Samuel A. Ball, Lisa Fenton, and Bruce J. Rounsaville. 2000. "A General System for Evaluating Therapist Adherence and Competence in Psychotherapy Research in the Addictions." *Drug & Alcohol Dependence* 57 (3): 225-38. [https://doi.org/10.1016/S0376-8716\(99\)00049-6](https://doi.org/10.1016/S0376-8716(99)00049-6).
- Cartreine, James A., Trina E. Chang, Janette L. Seville, Luis Sandoval, John B. Moore, Shuai Xu, and Mark T. Hegel. 2012. "Using Self-Guided Treatment software (ePST) to Teach Clinicians How to Deliver Problem-Solving Treatment for Depression." *Depression Research and Treatment Journal of the American Medical Association* 2012: 309094. <https://doi.org/10.1155/2012/309094>.
- Charlson, Fiona, Mark van Ommeren, Abraham Flaxman, Joseph Cornett, Harvey Whiteford, and Shekhar Saxena. 2019. "New WHO Prevalence Estimates of Mental Disorders in Conflict Settings: A Systematic Review and Meta-Analysis." *The Lancet* 394 (10194): 240-48. [https://doi.org/10.1016/S0140-6736\(19\)30934-1](https://doi.org/10.1016/S0140-6736(19)30934-1).
- Chibanda, Dixon, Petra Mesu, Lazarus Kajawu, Frances Cowan, Ricardo Araya, and Mealnie Abas. 2011. "Problem-Solving Therapy for Depression and Common Mental Disorders in Zimbabwe: Piloting a Task-Shifting Primary Mental Health Care Intervention in a Population with a High Prevalence of People Living with HIV." *BMC Public Health* 11 (1): 828. <https://doi.org/10.1186/1471-2458-11-828>.
- Chibanda, Dixon, Helen A. Weiss, Ruth Verhey, Victoria Simms, Ronald Munjoma, Simbarashe Rusakaniko, Alfred Chingono et al. 2016. "Effect of a Primary Care-Based Psychological Intervention on Symptoms of Common Mental Disorders in Zimbabwe: A Randomized Clinical Trial." *Journal of the American Medical Association* 316 (24): 2618-26. <https://doi.org/10.1001/jama.2016.19102>.

- Chorpita, Bruce F., Eric L. Daleiden, and John R. Weisz. 2005. "Identifying and Selecting the Common Elements of Evidence Based Interventions: A Distillation and Matching Model." *Mental Health Services Research* 7 (1): 5-20. <https://doi.org/10.1007/s11020-005-1962-6>.
- Chorpita, Bruce F., Eric L. Daleiden, and John R. Weisz. 2008. *Psychosocial and Combined Treatments Coding Manual*. University of California, Los Angeles: PracticeWise LLC.
- Crane, Rebecca S., Judith G. Soulsby, Willem Kuyken, J. Mark G. Williams, Catrin Earnes, Trish Bartley, Lucinda Cooper et al. 2016. "Mindfulness-Based Interventions Teaching Assessment Criteria (MBI:TAC): Manual, Summary, and Addendum." The Universities of Bangor, Exeter and Oxford, Bangor, United Kingdom.
- Dawson, Katie S., Richard A. Bryant, Melissa Harper, Alvin Kuowei Tay, Atif Rahman, Alison Schafer, and Mark van Ommeren. 2015. "Problem Management Plus (PM+): a WHO transdiagnostic psychological intervention for common mental health problems." *World Psychiatry* 14 (3): 354-357. <https://doi.org/10.1002/wps.20255>.
- Dickson, Kelly, and Mukdarut Bangpan. 2018. "What Are the Barriers to, and Facilitators of, implementing and receiving MHPSS Programmes Delivered to Populations Affected by Humanitarian Emergencies? A Qualitative Evidence Synthesis." *Global Mental Health* 5: e21. <https://doi.org/10.1017/gmh.2018.12>.
- Epping-Jordan, JoAnne E., Russ Harris, Felicity L. Brown, Kenneth Carswell, Claudette Foley, Claudia García-Moreno, Cary Kogan, and Mark van Ommeren. 2016. "Self-Help Plus (SH+): A New WHO Stress Management Package." *World Psychiatry* 15 (3): 295-296. <https://doi.org/10.1002/wps.20355>.
- Fairburn, Christopher G., and Zafra Cooper. 2011. "Therapist Competence, Therapy Quality, and Therapist Training." *Behaviour Research and Therapy* 49 (6-7): 373-78. <https://doi.org/10.1016/j.brat.2011.03.005>.
- Frank, J. R., L. S. Snell, O. T. Cate, E. S. Holmboe, C. Carraccio, S. R. Swing, P. Harris et al. 2010. "Competency-Based Medical Education: Theory to Practice." *Medical Teacher* 32 (8): 638-45. <https://doi.org/10.3109/0142159X.2010.501190>.

- Galappatti, Amanda, and Stephen M. Richardson. 2016. "Linking Mental Health and Psychosocial Support and Disaster Risk Reduction: Applying a Wellbeing Lens to Disaster Risk Reduction." *Intervention* 14 (3): 223-31.
- Global Social Service Workforce Alliance Interest Group on Para Professionals. 2017. *Para Professionals in the Social Service Workforce: Guiding Principles, Functions and Competencies*. Washington, DC: Global Social Service Workforce Alliance. <https://www.socialserviceworkforce.org/resources/para-professionals-social-service-workforce-guiding-principles-functions-and-competencies>.
- Hamdani, Usman, Parveen Akhtar, Huma Zill e, Huma Nazir, Fareed Minhas, Siham Sikander, Duolao Wang, Chiara Servilli, and Atif Rahman. 2017. "WHO Parents Skills Training (PST) Programme for Children with Developmental Disorders and delays Delivered by Family Volunteers in Rural Pakistan: Study Protocol for Effectiveness Implementation Hybrid Cluster Randomized Controlled Trial." *Global Mental Health* 4, e11. <https://doi.org/10.1017/gmh.2017.7>.
- Hendrickx, Maarten., Aniek Woodward, Daniella C. Fuhr, Egbert Sondorp, and Bayard Roberts. 2019. "The Burden of Mental Disorders and Access to Mental Health and Psychosocial Support Services in Syria and Among Syrian Refugees in Neighboring Countries: A Systematic Review." *Journal of Public Health* 42 (3): e299-e310. <https://doi.org/10.1093/pubmed/fdz097>.
- Hermosilla, Sabrina, Janna Metzler, Kevin Savage, Miriam Musa, and Alastair Ager. 2019. "Child Friendly Spaces Impact across Five Humanitarian Settings: A Meta-Analysis." *BMC Public Health* 19: e576. <https://doi.org/10.1186/s12889-019-6939-2>.
- IASC (Inter-Agency Standing Committee Reference Group for Mental Health and Psychosocial Support in Emergency Settings). 2017. *A Common Monitoring and Evaluation Framework for Mental Health and Psychosocial Support in Emergency Settings*. Geneva: IASC.
- Jordans, Mark J. D., and Brandon A. Kohrt. 2020. "Scaling Up Mental Health Care and Psychosocial Support in Low-Resource Settings: A Roadmap to Impact." *Epidemiology and Psychiatric Sciences* 29: e189. <https://doi.org/10.1017/S2045796020001018>.

- Jordans, Mark J. D., April Coetzee, H. Frederick Steen, Gabriela V. Koppenol-Gonzalez, Heba Galayini, Safwat Y. Diab, Samira A. Aisha, and Brandon A. Kohrt. 2021. "Assessment of service provider competency for child and adolescent psychological treatments and psychosocial services in global mental health: evaluation of feasibility and reliability of the WeACT tool in Gaza, Palestine." *Global Mental Health* 8: e7. <https://doi.org/10.1017/gmh.2021.6>.
- Jordans, Mark J. D., Brandon A. Kohrt, Manaswi Sangraula, Elizabeth L. Turner, Xueqi Wang, Pragya Shrestha, Renasha Ghimire, Edith van't Hof, Richard A. Bryant, Katie S. Dawson, Kedar Marahatta, Nagendra P. Luitel, and Mark van Ommeren. 2021. "Effectiveness of Group Problem Management Plus, a brief psychological intervention for adults affected by humanitarian disasters in Nepal: A cluster randomized controlled trial." *PLOS Medicine* 18 (6): e1003621. <https://doi.org/10.1371/journal.pmed.1003621>.
- Knutson, Nancy M., Marion S. Forgatch, Laura A. Rains, and Margrét Sigmarsdóttir. 2009. *Fidelity of Implementation Rating System (FIMP): The Manual for PMTO™*. Eugene, OR: Implementation Sciences International, Inc.
- Kohrt, Brandon, Laura Asher, Anvita Bhardwaj, Mina Fazel, Mark J. D. Jordans, Byamah Mutamba, Abhijit Nadkarni, Gloria Pedersen, Daisy Singla, and Vikram Patel. 2018a. "The Role of Communities in Mental Health Care in Low- and Middle-Income Countries: A Meta-Review of Components and Competencies." *International Journal of Environmental Research and Public Health* 15 (6): 1279. <https://doi.org/10.3390/ijerph15061279>.
- Kohrt, Brandon A., Mark J. D. Jordans, Sauharda Rai, Pragya Shrestha, Nagendra P. Luitel, Megan Ramaiya, Daisy Singla, and Vikram Patel. 2015. "Therapist Competence in Global Mental Health: Development of the Enhancing Assessment of Common Therapeutic Factors (ENACT) Rating Scale." *Behaviour Research and Therapy* 69: 11-21. <https://doi.org/10.1016/j.brat.2015.03.009>.
- Kohrt, Brandon A., Byamah B. Mutamba, Nagendra P. Luitel, Wilfred Gwaikolo, Patrick Onyango Mangen, Juliet Nakku, Kisa Rose, Janice Cooper, Mark J. D. Jordans, and Florence Baingana. 2018b. "How Competent Are Non-Specialists Trained to Integrate Mental Health Services in Primary Care? Global Health Perspectives from Uganda, Liberia, and Nepal." *International Review of Psychiatry* 30 (6): 182-98. <https://doi.org/10.1080/09540261.2019.1566116>.

- Kohrt, Brandon A., Megan K. Ramaiya, Sauharda Rai, Anvita Bhardwaj, and Mark J. D. Jordans. 2015b. "Development of a Scoring System for Non-Specialist Ratings of Clinical Competence in Global Mental Health: A Qualitative Process Evaluation of the Enhancing Assessment of Common Therapeutic Factors (ENACT) Scale." *Global Mental Health* 2: e23. <https://doi.org/10.1017/gmh.2015.21>.
- Kohrt, Brandon A., Alison Schafer, Ann Willhoite, Edith van't Hof, Gloria A. Pedersen, S. Watts, Katherine Ottman, Kenneth Carswell, and Mark van Ommeren. 2020. "Ensuring Quality in Psychological Support (WHO EQUIP): Developing a competent Global Workforce." *World Psychiatry* 19 (1): 115-16. <https://doi.org/10.1002/wps.20704>.
- Lahiri, Shaon, Mark van Ommeren, and B. Roberts. 2017. "The Influence of Humanitarian Crises on Social Functioning among Civilians in Low- And Middle-Income Countries: A Systematic Review." *Global Public Health* 12 (12): 1461-78. <https://doi.org/10.1080/17441692.2016.1154585>.
- Miller, William R., Theresa B. Moyers, Denise Ernst, and Paul Amrhein. 2003. Manual for Motivational Interviewing Skill Code (MISC). Unpublished manuscript. Albuquerque: Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico (2003). <https://casaa.unm.edu/download/misc.pdf>.
- Murphy, Regina, and Paul Hutton. 2018. "Practitioner Review: Therapist Variability, Patient-Reported Therapeutic Alliance, and Clinical Outcomes in Adolescents Undergoing Mental Health Treatment. A Systematic Review and Meta-Analysis." *Journal of Child Psychology and Psychiatry* 59 (1): 5-19. <https://doi.org/10.1111/jcpp.12767>.
- Mutamba, Byamah B., Jeremy C. Kane, Joop De Jong, James Okello, Seggane Musisi, and Brandon A. Kohrt. 2018a. "Psychological Treatments Delivered by community Health workers in Low-Resource Government Health Systems: Effectiveness of Group Interpersonal Psychotherapy for Caregivers of Children Affected by Nodding Syndrome in Uganda." *Psychological Medicine* 48 (15): 2573-83. <https://doi.org/10.1017/S0033291718000193>.

- Mutamba, Byamah B., Brandon A. Kohrt, James Okello, Janet Nakigudde, Bernard Opar, Seggane Musisi, William Bazeyo, and Joop de Jong. 2018. "Contextualization of Psychological Treatments for Government Health Systems in Low-Resource Settings: Group Interpersonal Psychotherapy for Caregivers of Children with Nodding Syndrome in Uganda." *Implementation Science* 13 (1): 90. <https://doi.org/10.1186/s13012-018-0785-y>.
- National Collaborating Centre for Mental Health, ed. 2019. *The Improving Access to Psychological Therapies Manual*. London: National Health Service.
- Nuro, Kathryn F., Lisa Maccarelli, Steve Martino, Samuel A. Ball, Stuart M. Baker, Bruce J. Rounsaville, and Kathleen M. Carroll. 2005. *Yale Adherence and Competence Scale (YACSII) Guidelines*. Yale University Psychotherapy Development Center, Manual-Guided Training Series No. 7. New Haven, CT: Yale University Psychotherapy Development Center Substance Abuse Center Department of Psychiatry
- OCHA (United Nations Office for the Coordination of Humanitarian Affairs). 2019. *2020 Global Humanitarian Overview*. Geneva: OCHA.
- Ottman, Katherine E., Brandon A. Kohrt, Gloria A. Pedersen, and Alison Schafer. 2020. "Use of Role Plays to Assess Therapist Competency and Its Association with Client Outcomes in Psychological Interventions: A Scoping Review and Competency Research Agenda." *Behaviour Research and Therapy* 130: 103531. <https://doi.org/10.1016/j.brat.2019.103531>.
- Pedersen, Gloria A., Pooja Lakshmin, Alison Schafer, Sarah Watts, Kenneth Carswell, Ann Willhoite, Katherine Ottman, Edith van't Hof, and Brandon A. Kohrt. 2020. "Common Factors in Psychological Treatments Delivered by Non-Specialists in Low- and Middle-Income Countries: Manual Review of Competencies." *Journal of Behavioral and Cognitive Therapy* 30 (3): 165-86. <https://doi.org/10.1016/j.jbct.2020.06.001>.
- Pedersen, Gloria A., Eva Smallegange, April Coetzee, Kim Hartog, Jasmine Turner, Mark J. D. Jordans, and Felicity L. Brown. 2019. "A Systematic Review of the Evidence for Family and Parenting Interventions in Low- and Middle-Income Countries: Child and Youth Mental Health Outcomes." *Journal of Child and Family Studies* 28 (8): 2036-55. <https://doi.org/10.1007/s10826-019-01399-4>.

- Puffer, Eev S., Jeannie Annan, Amanada L. Sim, Carmel Salhi, and Theresa S. Betancourt. 2017. "The Impact of a Family Skills Training Intervention among Burmese Migrant Families in Thailand: A Randomized Controlled Trial." *PLoS One* 12 (3): e0172611. <https://doi.org/10.1371/journal.pone.0172611>.
- QSR International. 2012. *NVIVO Qualitative Data Analysis Software 10*. Doncaster, Australia: QSR International.
- Rose-Clarke, Kelly, Indira Pradhan, Pragma Shrestha, Prakash B.K, Jananee Magar, Nagendra P. Luitel, Delan Devakumar et al. 2020. "Culturally and developmentally adapting group interpersonal therapy for adolescents with depression in rural Nepal." *BMC Psychology* 8 (1): e83. <https://doi.org/10.1186/s40359-020-00452-y>.
- Sangraula, Manaswi, Elizabeth L. Turner, Nagendra P. Luitel, Edith van Hof, Pragma Shrestha, Renasha Ghimire, Richard Bryant et al. 2020. "Feasibility of Group Problem Management Plus (PM+) to Improve Mental Health and Functioning of Adults in Earthquake-Affected Communities in Nepal." *Epidemiology and Psychiatric Sciences* 29: e130. <https://doi.org/10.1017/S2045796020000414>.
- Save the Children. 2009. *Child Friendly Spaces Facilitator Training Manual*. New York: Save the Children. <https://resourcecentre.savethechildren.net/library/child-friendly-spaces-facilitator-training-manual>.
- Singh, Maanvi 2017. "The Friendship Bench Can Help Chase the Blues Away." *Goats and Soda* [NPR podcast], January 10, 2017. <http://www.npr.org/sections/goatsandsoda/2017/01/10/508588401/the-friendship-bench-can-help-chase-the-blues-away>.
- Singla, Daisy R., Brandon A. Kohrt, Laura K. Murray, Arpita Anand, Bruce F. Chorpita, and Vikram Patel. 2017. "Psychological treatments for the world: Lessons from low- and middle-income Countries." *Annual Review of Clinical Psychology* 13: 149-81. <https://doi.org/10.1146/annurev-clinpsy-032816-045217>.
- Singla, Daisy R., and Elias Kumbakumba. 2015. "The Development and Implementation of a Theory-Informed, Integrated Mother-Child Intervention in Rural Uganda." *Social Science & Medicine* 147: 242-51. <https://doi.org/10.1016/j.socscimed.2015.10.069>.

- Singla, Daisy R., Elias Kumbakumba, and Frances E. Aboud. 2015. "Effects of a Parenting Intervention to Address Maternal Psychological Wellbeing and Child Development and Growth in Rural Uganda: A Community-Based, Cluster-Randomised Trial." *The Lancet Global Health* 3 (8): e458-69. [http://doi.org/10.1016/S2214-109X\(15\)00099-6](http://doi.org/10.1016/S2214-109X(15)00099-6).
- Te Pou o te Whaakaro Nui. 2016. *Practice support: Competencies, training and supervision for talking therapies delivery*. Auckland, New Zealand: Te Pou o te Whakaaro Nui.
- Tol, Wieste A., Jura Augustinavicius, Kenneth Carswell, Felicity L. Brown, Alex Adaku, Marx R. Leku, Claudia García-Moreno, Peter Ventevogel, Ross G. White, and Mark van Ommeren. 2018. "Translation, Adaptation, and Pilot of a Guided Self-Help Intervention to Reduce Psychological Distress in South Sudanese Refugees in Uganda." *Global Mental Health* 5: e25. <https://doi.org/10.1017/gmh.2018.14>.
- Tol, Wieste A., Marx R. Leku, Daniel P. Lakin, Kenneth Carswell, Jura Augustinavicius, Alex Adaku, Teresa M. Au et al. 2020. "Guided Self-Help to reduce Psychological Distress in South Sudanese Female Refugees in Uganda: A Cluster Randomised Trial." *Lancet Global Health* 8: e254-63. [https://doi.org/10.1016/S2214-109X\(19\)30504-2](https://doi.org/10.1016/S2214-109X(19)30504-2).
- UNESCO (United Nations Educational, Scientific and Cultural Organization) and Irina Bokova. 2017. *UNESCO Strategic Framework for Education in Emergencies in the Arab Region (2018-2021)*. Beirut: UNESCO.
- United Nations. 2020. *Amid Threat of Catastrophic Global Famine, COVID-19 Response Must Prioritize Food Security, Humanitarian Needs, Experts Tell General Assembly*. 31st Special Session, GA/12294. New York: United Nations General Assembly.
- Van Ommeren, Mark, Bhogendra Sharma, Suraj Thapa, Ramesh Makaju, Dinesh Prasain, Rabindra Bhattaria, and Joop T. V. M. de Jong. 1999. "Preparing Instruments for Transcultural Research: Use of the Translation Monitoring Form with Nepali-Speaking Bhutanese." *Transcultural Psychiatry* 36 (3): 285-301. <https://doi.org/10.1177/136346159903600304>.

- Verdeli, Helen, Kathleen Clougherty, Grace Onyango, Eric Lewandowski, Liesbeth Speelman, Theresa S. Betancourt, Richard Neugebauer, Traci R. Stein, and Paul Bolton. 2008. "Group Interpersonal Psychotherapy for Depressed Youth in IDP Camps in Northern Uganda: Adaptation and Training." *Child & Adolescent Psychiatric Clinics of North America* 17 (3): 605-24. <https://doi.org/10.1016/j.chc.2008.03.002>.
- Verdeli, Helen, Tatiana Therosme, Eddy Eustache, Olissaint S. Hilaire, Benissois Joseph, Cemile C. Sönmez, and Giuseppe Raviola. 2016. "Community Norms and Human Rights: Supervising Haitian Colleagues on Interpersonal Psychotherapy (IPT) with a Depressed and Abused Pregnant Woman." *Journal of Clinical Psychology* 72 (8): 847-55. <https://doi.org/10.1002/jclp.22366>.
- Wood, Benjamin M., and Per Kallestrup. 2018. "A Review of Non-Specialised, Group-Based Mental Health and Psychosocial Interventions in Displaced Populations." *International Journal of Migration, Health and Social Care* 14 (3): 347-59. <https://doi.org/10.1108/IJMHSC-02-2018-0010>.
- World Health Organization. 2017. *Caregiver Skills Training for the Management of Developmental Disorders: Evidence Base*. Geneva: World Health Organization.

APPENDIX

GROUP FACILITATION ASSESSMENT OF COMPETENCIES TOOL (GROUPACT)

#1 GROUP GUIDELINES AND/OR GROUNDRULES

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Violates group guidelines (i.e., answers phone in-session, interrupts members, etc.) <input type="checkbox"/> Allows members to violate guidelines without correction or acknowledgement	<input type="checkbox"/> Introduces concept of group session guidelines in first session <input type="checkbox"/> Establishes ground rules (e.g., respect, listen, and pay attention to each other, “everything that gets discussed stays here,” no phones, etc.) <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Asks for agreement from the group on guidelines <input type="checkbox"/> Elicits group feedback, providing interpretations and reflections <input type="checkbox"/> Adjusts rules depending on need or context

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

#2 REVIEW OF GUIDELINES AND/OR GROUNDRULES IN SUBSEQUENT SESSIONS

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Violates group guidelines (i.e., answers phone in-session, interrupts members, etc.) <input type="checkbox"/> Allows members to violate guidelines without correction or acknowledgement <input type="checkbox"/> Shames participant for breaking ground rules	<input type="checkbox"/> Reviews and encourages adherence to ground rules <input type="checkbox"/> Acknowledges when ground rules are being broken and addresses it <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Asks for agreement from the group on guidelines <input type="checkbox"/> Elicits group feedback, providing interpretations and reflections <input type="checkbox"/> Adjusts rules depending on need or context

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

#3 GROUP PARTICIPATION

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Displays favoritism to specific members <input type="checkbox"/> Excludes other members (e.g., ignores input) <input type="checkbox"/> Forces unwilling participant to join discussion <input type="checkbox"/> Scolds participant(s) for under- or over-sharing	<input type="checkbox"/> Uses timely techniques (e.g., turn taking; “gentle prompting,” etc.) to encourage fair participation <input type="checkbox"/> Clarifies discussion points for members struggling with literacy, numeracy, or tech skills <input type="checkbox"/> Addresses participation barriers (e.g., interruptions) <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Provides reflection on discussion <input type="checkbox"/> Discusses ways members can support one another to participate <input type="checkbox"/> Checks in on comfort with sharing for all members

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

#4 FOSTERING EMPATHY AMONG GROUP MEMBERS

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Does not intervene when group members are unempathetic, hurtful, or hostile toward one another <input type="checkbox"/> Shares member information from private session as an example of empathy or lack of empathy	<input type="checkbox"/> Encourages and fosters empathy among group members (e.g., points out expressions of empathy toward one another) <input type="checkbox"/> Uses prompts (e.g., “How do you feel after you shared with us?”) <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Provides summarizing statements and interpretations (see example above) <input type="checkbox"/> Demonstrates empathy for others to follow (e.g., nods head, says “uh huh”)

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

#5 COLLABORATIVE PROBLEM-SOLVING

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Gives direct suggestions without group input <input type="checkbox"/> Does not intervene or address harmful or unrealistic solutions (e.g., “quit job,” “avoid husband,” etc.) <input type="checkbox"/> Judges solutions created by group <input type="checkbox"/> Allows judgement from members (e.g., “That will never work,” “That’s stupid,” etc.)	<input type="checkbox"/> Equally encourages all members to share how they addressed similar problems <input type="checkbox"/> Encourages members to praise and support each other for positive and challenging experiences <input type="checkbox"/> Unrealistic and unhelpful solutions are eliminated <input type="checkbox"/> Encourages brainstorming of solutions in the group <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Solicits feedback from group to ensure solutions are attainable, realistic, and timely <input type="checkbox"/> Addresses potential barriers <input type="checkbox"/> Validates challenges (e.g., “Not all solutions work for everyone”) <input type="checkbox"/> Suggests member pairing to support each other (if applicable to the intervention)

Check the level that best applies (only one level should be checked)

Level 1
any unhelpful behavior

Level 2
no basic skills,
or some but not all basic skills

Level 3
all basic skills

Level 4
all basic helping skills
plus any advanced skill

Notes:

#6 ADDRESSING BARRIERS TO ATTENDANCE

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Actively excludes members' attendance (e.g., does not make schedule adjustments) <input type="checkbox"/> Ignores feedback on barriers to attendance <input type="checkbox"/> Rejects/ignores sociodemographic and minority barriers (e.g., religious observances, menstrual practices, disabilities, etc.)	<input type="checkbox"/> Actively solicits information to address potential barriers to attendance (e.g., work/farming schedule, transportation, etc.) <input type="checkbox"/> Works to reschedule sessions or adjusts schedules accordingly <input type="checkbox"/> Encourages members to attend even if previous sessions were missed <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Engages group in problem-solving about how all members can attend sessions <input type="checkbox"/> Encourages group members to support one another in attending group sessions <input type="checkbox"/> Supports access to sessions (e.g., convenient session locations, traveling together, etc.)

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

DEVELOPING GROUPACT FOR HUMANITARIAN SETTINGS

#7 GROUP CONFIDENTIALITY

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Shares a member's confidential information with the group (e.g., a member's trauma history disclosed in private) <input type="checkbox"/> Threatens to share group information with community or family members <input type="checkbox"/> Targets or blames members when confidentiality is broken <input type="checkbox"/> Others break confidentiality	<input type="checkbox"/> Explains rationale for confidentiality, including situations when confidentiality can be broken <input type="checkbox"/> Confidentiality issues are addressed when unexpected/uninvited individuals arrive <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Explains issues of respect, valuing others' experiences, and feelings associated with violation of confidentiality <input type="checkbox"/> Appropriately addresses times when confidentiality is broken during group sessions

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

#8 TIME MANAGEMENT: APPROPRIATE BREAKS, ENERGIZERS, AND PACING

Check all behaviors that are demonstrated in each category.		
Unhelpful or potentially harmful behaviors	Basic helping skills	Advanced helping skills
<input type="checkbox"/> Prevents clients from taking a break <input type="checkbox"/> Forces group to continue when emotionally exhausted or distressed <input type="checkbox"/> Targets or blames participant(s) when requesting breaks or energizers	<input type="checkbox"/> Reviews schedule for the day <input type="checkbox"/> Includes and explains timed breaks with instructions for start/stop signals <input type="checkbox"/> Includes time for questions in schedule <input type="checkbox"/> <i>None of the above</i>	<input type="checkbox"/> <i>Completes all basic helping skills (Level 3)</i> <input type="checkbox"/> Elicits feedback and checks in with learning (e.g., has well-spaced summarizing and “checking in” activities) <input type="checkbox"/> Checks in with group to see when breaks are needed/preferred <input type="checkbox"/> Conducts group energizers

Check the level that best applies (only one level should be checked)

<input type="checkbox"/> Level 1 <i>any unhelpful behavior</i>	<input type="checkbox"/> Level 2 <i>no basic skills, or some but not all basic skills</i>	<input type="checkbox"/> Level 3 <i>all basic skills</i>	<input type="checkbox"/> Level 4 <i>all basic helping skills plus any advanced skill</i>
--	---	--	--

Notes:

BOOK REVIEW

CAN BIG BIRD FIGHT TERRORISM?
CHILDREN'S TELEVISION AND GLOBALIZED MULTICULTURAL EDUCATION
BY NAOMI A. MOLAND
OXFORD UNIVERSITY PRESS, 2020. XI + 277 PAGES
\$39.95 (HARDCOVER), \$31.95 (E-BOOK)
ISBN 978-0-19-090395-4

Founded in 1969, Sesame Workshop, the nonprofit behind *Sesame Street*, celebrated its 50th birthday in 2019.¹ The groundbreaking educational television program has given birth to a huge body of research about educational television. As a child in small-town New England, *Sesame Street* was my reference point for city life: the first time I remember seeing an African American person on television and the first time my Deaf sister and I saw an adult who wasn't a teacher or parent of a Deaf child use sign language was on the program. Treating diversity as a given and encouraging curiosity in young children across generations is *Sesame Street's* great strength. Over the decades, that spirit of inclusiveness and progressive values in education has infused Sesame Workshop's international projects in more than 30 countries around the world.

Naomi Moland's addition to the literature on Sesame Workshop, *Can Big Bird Fight Terrorism? Children's Television and Globalized Multicultural Education*, is especially relevant in the shift to remote learning during this moment of COVID-19. Moland provides an in-depth look at *Sesame Square*, a children's television program offered in Nigeria through a partnership of Sesame Workshop and the US Agency for International Development. Nigeria was chosen as the site of this partnership because simmering regional conflict and religious differences were fueling extremist groups, like Boko Haram. Moland uses a combination of interviews, ethnographic observations, and episode analysis to develop a case study focused on two questions about the potential contribution multicultural education can make to peacebuilding. First, she seeks to understand how educators, in this case the writers and producers of *Sesame Square*, in their efforts to localize an externally developed curriculum can inadvertently re-create or reinforce the very stereotypes and divisions they want to break down. Second, she interrogates whether a "public curriculum" of conflict, violence, division, and discrimination renders multicultural education's messages of peace and tolerance ineffective or even offensive.

1 The original name was the Children's Television Workshop.

Moland draws on the work of Cynthia Miller-Idriss to provide a compelling conceptual critique of multicultural education as the basis for peacebuilding, cautioning us to “temper our expectations for what education can do” (p. 190). She explains that the medium of television is especially prone to the pitfalls of multiculturalism. This includes reducing differences to fixed identities using recognizable symbolic images that leave little room for viewer interaction, and little space to capture people’s “messy, shifting identities” (p. 197) with the level of complexity needed to provide a critically inclusive experience. In Nigeria, the reduction of identity was a particular challenge with viewers from the north where, in an effort to ensure that clerics and community leaders would not forbid people to watch, such depictions were reduced to a static picture of the most conservative iteration of identity.

In chapters 2-4, Moland sets up the fundamental challenge of multicultural education implemented through international development. Since the goal of multicultural education programs will always be to catalyze change in attitudes and behaviors, program developers will always face the dilemma of balancing local values with values from elsewhere. In this case, in an environment rife with sharp divisions along religious, regional, and ethnic lines, Sesame Workshop’s Western expression of the values of peace and inclusion struck an uneasy balance with local conceptions of the same values. As educators and development agencies localize projects, they focus on the “needs” of their audience. In this process, they often take over the power to define their audience’s differences, often framed as deficits, and thus tend to echo the discourses of “orientalism they intend to correct” (p. 194). Moland describes the power dynamics of international organizations choosing which locals to privilege as they construct both problem and solution. Ensuring a collaborative and largely equal partnership between the writers and developers from Sesame Workshop in the United States and the *Sesame Square* production team in Nigeria was a large part of the effort to get this balance right. The exploration of this theme is useful beyond the current case study because the producers tried so hard to get the balance right, yet still faced instructive challenges.

Further using the theory of nesting orientalisms—a variation of Said’s work explaining that a group “orientalized” in the North and West can, in turn, “orientalize” another group, resulting in multiple or nested orientalized identities (Bakić-Hayden 1995)—to interrogate the possibility of reinforcing stereotypes through multicultural education, Moland concludes that the target audience in multicultural education and international development projects will always be “othered,” or treated as intrinsically different and alien to their true selves, because, even when they work closely with host governments or communities,

the outside agencies identify both the project goals and the target audience. This is especially true in conflict-affected contexts, where divisions already run deep (p. 191) and outside actors' access to communities is often curtailed. Moland uses the example of *Sesame Square's* portrayal of Nigerians from the north of the country as a compelling example, documenting writing room conversations showing that the Nigerian team included mostly Christian Yoruba writers, with only two relatively elite writers from northern Nigeria. The needs and concerns of the Muslim Hausa target audience were filtered through this group. She cites religious differences as especially difficult to bridge, particularly when people see their religion, in fundamentalist terms, as the only "correct" way of being in the world.

Moland's findings and response to the central question on the inadvertent reproduction of stereotypes provide a well-researched and thoughtful critique that applies to international education projects. However, they might not have the inevitability in terms of children's television that she assumes. Othering is likely to occur when the diversity of the writing room does not reflect the diversity of the audience, so whether the writing team could have taken more risks in their portrayal of religious diversity if its members had been more representative or if they had been able to conduct more pilots and focus groups in communities in northern Nigeria is an important question. Moland suggests that managing multicultural initiatives with respect to the problem of othering the target audience, and its downstream effects, is an area for further research. She also recommends that future efforts focus on hybrid and fluid identities by showing, for example, characters who speak some degree of several languages, as many Nigerians do (p. 204). This recommendation was not taken up by the producers, who were worried about confusing or alienating part of their audience.

In chapters 5-6, Moland finds that Big Bird might be able to fight terrorism, but he cannot do it alone. He will inevitably be undermined when a public curriculum or surrounding narrative of violence coexists with the lasting structural violence of colonialism and a government that is "incompetent and corrupt such that people cannot know whether state institutions exist to provide services or prey on them" (p. 198). The program creators must hope that *Sesame Square's* messages of peace and tolerance can help lay the "foundation for the ongoing battle for hearts and minds" (p. 200). Moland explores this point in relation to violent conflict and the path that young people might follow—from social and religious networks that provide social services and protection to radical terrorist organizations like Boko Haram. Showing an alternative to the existing public curriculum of state corruption could also open a pathway to demanding more accountable

government. In framing her pessimism about multicultural education projects, Moland assumes the typical two- to four-year timeline of a typical international development project. Interrogating the relationship between the timeline, those driving the project, and support for surrounding or ancillary activities is a shortcoming in her analysis.

As Moland rightly points out, the multicultural education project's offering cannot be so utopian that it becomes unrelatable, which returns to the highly contextual question of how to balance the curricula of informal public education programs like *Sesame Square* and formal education. Moland cites the need for remedies beyond this single television program, such as expanding access to academically relevant and inclusive opportunities in local school contexts. Other suggestions, which may be more workable for a program like *Sesame Square* than developing different versions of the program or expanding formal educational opportunities, are to emphasize commonalities, such as the fact that the entire audience is affected by conflict, and to focus on teaching children the skills and processes they need to understand diverse others, to develop fluid, layered identities, and to resolve conflict. Overall, Moland has made a valuable and well-written addition to the literature on children's television, multicultural education, and Sesame Workshop by clearly naming the successes and pitfalls of *Sesame Square* in Nigeria, so that others can learn from this experience. Her research and recommendations for developing multicultural education programs delivered through television take on new relevance in light of the school closures in the era of COVID-19 as we intensify our exploration of the possibilities of remote education beyond basic academic skills.

KATE LAPHAM

Open Society Foundations

*The views expressed here are the author's
and do not represent the Open Society Foundations.*

For permission to reuse, please contact journal@inee.org.

REFERENCE

Bakić-Hayden, Milica. 1995. "Nesting Orientalisms: The Case of Former Yugoslavia." *Slavic Review* 54 (4): 917-31. <https://doi.org/10.2307/2501399>.

BOOK REVIEW

*NISSEM GLOBAL BRIEFS: EDUCATING FOR THE SOCIAL,
THE EMOTIONAL AND THE SUSTAINABLE*
EDITED BY ANDY SMART, MARGARET SINCLAIR, AARON BENAVIDES,
JEAN BERNARD, COLETTE CHABBOTT, S. GARNETT RUSSELL,
AND JAMES WILLIAMS
NISSEM, 2019. 453 PAGES
FREE ONLINE ACCESS: [HTTPS://WWW.NISSEM.ORG/NGB1](https://www.nissem.org/NGB1)

NISSEM Global Briefs: Educating for the Social, the Emotional and the Sustainable addresses how to educate children to have important skills: skills for life, skills for the 21st century, skills needed in a modern, unequal, globalized, and polluted world; skills needed to reach the 17 UN Sustainable Development Goals (SDGs). In its 10 sections and 42 chapters, this edited volume offers the diverse perspectives of more than 60 contributors.

Education in emergencies audiences will be interested in this inspirational book because it provides knowledge on how to use education to achieve sustainable development and sustainable lifestyles, human rights, and gender equality; to promote a culture of peace, nonviolence, global citizenship, and appreciation for cultural diversity; and to encourage understanding of the contributions diverse cultures make to sustainable development all around the world, as stated in SDG Target 4.7. A central message of this book is that the SDG 4.7 themes are most effectively integrated into textbooks and other educational materials when the social and emotional learning (SEL) components of these topics are emphasized. The chapters are well written and interesting, most ideas are clearly presented, and readers will appreciate the practical advice. It is especially important for practitioners in all areas of development, and for those seeking to promote a sustainable future.

The book's contributors represent a wide range of educational backgrounds and fields, which makes this book relevant in diverse contexts. The 10 sections cover relevant areas of SDG 4.7, which calls on leaders from government, academia, civil society, and business to, by 2030, accelerate the implementation of Education for Sustainable Development around the world, and ensure that all learners acquire the knowledge and skills needed to promote sustainable development.

The first three sections address how to embed SDG 4.7 themes in textbooks, how to contextualize SEL in textbooks, and how to promote inclusion and social cohesion in textbooks. These sections suggest that, if the SDGs are to be reached by 2030, children

need different skills than previous generations. The general recommendations for changing curricula are neither provocative nor surprising, and they provide practical examples of how to develop and advance next-generation skills. Importantly, these sections provide researchers, policymakers, and practitioners with valuable materials for writing proposals, policy notes, and research papers.

Section four focuses on how interdisciplinary and holistic science can inform SEL, and on the mechanisms of cognitive development that underlie this relationship. Creating new textbooks and other SEL tools requires content, shape, and form, as sections five through nine demonstrate. While reading these sections, I imagined textbook authors around the world going from good intentions and principles to actually creating and implementing effective learning material for children.

This comprehensive volume offers examples from many countries, including Botswana and Somalia, India and Bangladesh, South Sudan and Rwanda, Afghanistan and New Zealand. These examples are both enjoyable and useful, and practitioners will find similarities with the kinds of challenges they face when developing materials across contexts, including how to create material that is engaging for all readers and relevant across gender, religion, reading skills, and various barriers. I was pleased that the contributors to this book have collaborated with both the biggest players in international development and a range of smaller NGOs. We Love Reading, one of the smaller NGOs represented, advocates that relevant learning material is crucial to empowering children and that it needs to tell stories that all children across the globe can identify with.

The last section raises the question of what actions NISSEM proposes to advance SDG 4.7 themes that are supported by SEL, and makes the following recommendations:

- Integrate SDG 4.7 themes into all education strategies.
- Measure and monitor SEL to ensure that all learners acquire the knowledge and skills needed to promote sustainable development of the relevant SDG 4.7 goals in terms of attitudes, behavior, skills, and content knowledge.
- Invest in teaching and learning materials and secure the donor commitments needed to achieve SDG 4.7.

This broad perspective gives an overview of the SEL field, which I deeply appreciate. However, for textbook authors working on SEL materials, there are some practical things I wish this book had provided, such as a list of where recommended

questionnaires and measurements can be found and how to access them. For example, the International Social and Emotional Learning Assessment is mentioned in several chapters and it would be useful to know how to find it. I also would have liked to see a list of existing SEL materials and how to access them in different languages.

From a more philosophical point of view, I found myself wondering if I am the only one running challenging SEL projects with data that are often messy and difficult to use to create a clear picture; the examples in the book make it seem that there are only neat, well-structured SEL experiences and evaluations going on. I searched the text for the words “failure” and “success,” and while the word “success” was used 89 times, “failure” was used only eight times, and never in reference to a project’s implementation or data collection. As a practitioner, I have seen a lot of data that are missing gender, age, and dates, where pre- and post-measures have been mixed up, or the work is not systematically evaluated, but this volume does not offer any examples of learning from failure. When compared with the startup and innovative environments I have been involved in, this book makes it appear that education’s SEL practitioners are not sharing their mistakes generously. Paradoxically, some of the stories and tools discussed in this volume are about students struggling to accept making mistakes. Learning from mistakes could be liberating for practitioners, and also more fun, empowering, and realistic.

I am working on a digitalized psychosocial tool to be used with Syrian refugees in Lebanon, and although this book may not make my everyday choices easier, it has definitely helped me base some of my decisions more solidly on research-based evidence and shared experiences from the field. For textbook authors, publishers, and SEL developers, *NISSEM Global Briefs: Educating for the Social, the Emotional and the Sustainable* is a must read.

SOLFRID RAKNES

*European Federation of Psychologists’ Associations (EFPA),
Board of Prevention and Promotion*

*The views expressed here are the author’s,
and do not represent the EFPA Board of Prevention and Promotion.*

For permission to reuse, please contact journal@inee.org.

HOW THE EDUCATION IN EMERGENCIES FIELD CAN HELP THE UNITED STATES RESPOND TO COVID-19

REBECCA WINTHROP AND HELEN SHWE HADANI

The field of education in emergencies (EiE) has developed rapidly over the last two decades. In the early 2000s, the Comparative and International Education Society held only a smattering of sessions that discussed education during or after crisis. The newly formed Education for All Fast Track Initiative—today the Global Partnership for Education (GPE)—by design did not channel its funding to fragile or conflict-affected states. The Inter-agency Network for Education in Emergencies (INEE), established in the early 2000s to help advance the Education for All movement in contexts affected by conflict and natural disaster, had a toolkit consisting of several blue cardboard boxes that held a selection of photocopied manuals and guides from various organizations, the best of what field-based experts had to offer at the time. Many of these materials focused on the nuts and bolts of delivering education services in refugee camps or other crisis contexts, and only a few drew direct links between child protection, children’s psychosocial wellbeing, and education continuity.

The argument that education was one of the best ways, if not the best, to support children’s protection and psychosocial wellbeing amid crisis has been central to catalyzing the growing attention to and development of the EiE field over the past two decades. Today, with half the GPE partner countries classified as fragile states, EiE draws from a robust and wide-ranging array of global minimum standards and guidance notes, copious digital resources, a multilingual network of thousands of practitioners and policymakers, multiple courses of study in institutions of higher learning, and many global mechanisms for providing financing to crisis contexts. After a great deal of work by hundreds and hundreds of practitioners, researchers, and advocates, EiE today is a well-established sector in the global humanitarian and development communities.

Journal on Education in Emergencies, Vol. 7, No. 2
Copyright © 2021 by the Inter-agency Network for Education in Emergencies (INEE).
ISSN 2518-6833

However, before the COVID-19 pandemic, the EiE sector and its know-how had been largely ignored in high-income countries. Relegated to the category of a developing-world problem, few schools in high-income countries had disaster preparedness plans that went beyond short-term school closings due to gun violence or other exceptional circumstances. School systems in the Global North were not familiar with the best practices for ensuring continuity of education in the midst of crisis, or with the many other lessons the EiE field can provide.

So what can the EiE field teach the developed world? Reflections on the US experience can shed light on this question. In spring 2020, when school closures began due to the COVID-19 pandemic, schools across the country were taken by surprise and largely left to fend for themselves. Standard EiE practices were frequently nonexistent, and school districts rarely delivered life-saving messages to their students, teachers, and personnel about the behaviors they needed to practice to stay safe. The principle of “do no harm”—namely, avoiding “exposing people to further harm” as a result of humanitarian interventions (Sphere Project 2011, 29)—was often overlooked in school-closure decisions. This resulted in the needs of the most vulnerable students, such as children at risk of domestic violence or those whose parents were frontline workers, to be disregarded, which left them exposed to greater risks during home-based schooling. District efforts frequently focused on pivoting the delivery of lessons to remote learning mechanisms, especially during the early response, and forged ahead with normal math, language, or science lessons, giving little attention to adapting the content of what children were studying to recognize that the world around them had changed, making space for their questions, or thinking through both young people’s and educators’ psychosocial needs. In all these respects, school districts would have greatly benefitted from EiE advice and guidance, such as that provided in the INEE technical note on the pandemic (INEE 2020).

In the United States, most of the public debate and attention given to children’s education amid the pandemic has been on the lost instructional time and so-called learning loss, which refers to children’s delayed academic progress. For example, students are estimated to be as much as one-third of an academic year behind grade level in reading (Education Analytics 2021; Spector 2021) and some nearly two-thirds of a year behind in math (Dorn et al. 2021); it is the country’s black and brown students who have been hardest hit by these losses. This rising inequality in academic outcomes is deeply concerning, but it should not eclipse attention to children’s psychosocial needs, which is a lesson the EiE community has learned repeatedly in numerous contexts.

As the pandemic has worn on, the US education community has begun to expand its focus from academic learning gaps to mental health concerns. This is in response to a nationally representative survey showing that, after just a few weeks of school closures, one-third of high school students reported increased feelings of anxiety, worry, and depression (America's Promise Alliance 2020). This has led to increasing calls for schools to hire more counselors to provide mental health services to students. While this is an important step, it by no means precludes other ways of addressing children's psychosocial needs.

Drawing from years of practitioner know-how and an increasing research base, largely in the developing world, the INEE Minimum Standards for Education can provide useful guidance, even for developed countries like the US, that are struggling to address children's education needs amid the pandemic. Several elements of good practice in the INEE standards are especially relevant for the US education community:

- The importance of integrating expressive activities, including play-based approaches, into children's daily routines as a way to support children's psychosocial wellbeing;
- Using the full range of spaces and assets available in communities (e.g., parks and school buildings) to help children learn and connect; and
- Harnessing the disruption caused by a crisis to find better ways of supporting children's long-term holistic development and addressing their cognitive, psychological, social, and physical needs (INEE 2010).

There are many creative ways to support children's psychosocial wellbeing, some of which were present prior to the pandemic, although only marginally in the US education system, that can be adopted and adapted for a range of contexts. One example is Playful Learning Landscapes (PLL), an initiative that uniquely blends the science of learning, placemaking, and community cohesion. PLL, which builds on multiple efforts and initiatives aimed at creating more child-friendly cities, transforms public and shared spaces into fun and enriching learning hubs for the development of healthy children, families, and communities (Bustamante et al. 2021; Hassinger-Das et al. 2021).

While children's play may look different during the pandemic, great effort should be made to ensure that they can still play safely at home or outdoors with their

caregivers and siblings. From the earliest age, humans crave social interaction. The classic “still-face experiment” clearly shows this. When a parent gazes at their infant with a blank expression (as they are instructed to do in the experiment), the interaction quickly goes downhill and the baby often gets upset (Adamson and Frick 2003). Research also tells us that “serve-and-return” interactions, in which children “serve” by showing an interest in something and their caregiver “returns” by responding in a supportive way, help to build connections in the brain that influence language and cognitive development (Romeo et al. 2018).

Experts argue that children need play now more than ever, since play enables them to express and process their feelings and experiences, especially those that are scary or confusing. An April 2020 article in *The Atlantic* shares parents’ anecdotes about children playing “CoronaBall,” a version of dodgeball, and “Social-distancing tag,” where children tag each other’s shadows (Cray 2020). It is clear that play permeates children’s lives, even during a pandemic.

PLL is one approach that can provide increased opportunities for children to play while schools are closed and families are the central locus of education. PLL locates its installations in spaces families frequent, such as bus stops, parks, supermarkets, and laundromats. While many of these spaces may look different during a pandemic, infusing public and shared spaces with designs informed by the latest science of learning can provide much-needed opportunities for quality caregiver-child interactions and the kind of playful learning children crave. Moreover, PLL’s flexible model can be adapted to key safety measures, including those related to high-touch surfaces and social distancing. Urban Thinkscape, a PLL installation that transformed an abandoned lot next to a bus stop in West Philadelphia into an interactive play space, is one of many that promote playful learning in an outdoor setting, which during the pandemic is generally thought to be safer than indoor spaces (Hassinger-Das, Palti et al. 2020). Transforming unexpected places such as city sidewalks and vacant lots into spaces where children can have playful learning opportunities can be a low-cost, COVID-friendly way to boost their learning during and after the pandemic.

Integrating interventions like PLL into public and shared spaces is one way the US education community can leverage EiE’s guidance to “build back better.” Data from pilot PLL installations in Philadelphia and Chicago, among other cities, show that PLL promotes the kind of caregiver-child communication that supports language learning and builds relationship; encourages children to talk about numbers, letters, and spatial relations; and increases caregivers’ understanding of the connection between play and learning (Bustamante et al. 2020; Hassinger-

Das, Palti et al. 2020; Hassinger-Das, Zosh et al. 2020; Ridge et al. 2015). At the same time, PLL engages communities in the revitalization of the public realm, which creates new opportunities for multigenerational social interaction, as well as more vibrant and livable cities.

The current crisis can be leveraged as a unique opportunity for educators, parents, researchers, and policymakers to reimagine more equitable education systems. As highlighted in a recent report on education reform (Hirsh-Pasek et al. 2020), children need to develop a breadth of skills beyond literacy and numeracy in order to engage in lifelong learning and to succeed in today's rapidly changing, globalized world. An overwhelming body of evidence points to play as the best way to equip children with a broad set of flexible competencies that enable them to tackle new and different challenges creatively. Play is ubiquitous in childhood and foundational to human development and learning, and it has a unique and important role in supporting children's socioemotional development. Pretend play, for example, enables children to practice emotional regulation, which ranges from infant self-soothing to becoming more conscious of one's emotions to developing strategies to manage those emotions (Thompson and Calkins 1996).

The education community will be drawing from lessons learned during the COVID-19 pandemic for years to come. These insights will be relevant to all countries across the globe. Who knows—a decade down the road, we may look back at the pandemic as the event that connected the EiE community to the developed world.

REFERENCES

- Adamson, Lauren B., and Janet E. Frick. 2003. "The Still Face: A History of a Shared Experimental Paradigm." *Infancy* 4 (4): 451-73. https://doi.org/10.1207/S15327078IN0404_01.
- America's Promise Alliance. 2020. "National Survey of High School Students During COVID-19 Finds Widespread Negative Impact on Learning Time, Emotional Health, Social Connection." Washington, DC: America's Promise Alliance. <https://www.americaspromise.org/press-release/national-survey-high-school-students-during-covid-19-finds-widespread-negative-impact>.

- Bustamante, Andres S., Helen S. Hadani, June Ahn, and Kathy Hirsh-Pasek. 2021. "Reimagining Urban Design for Play and Learning." *Dialogues: The Intersection of Emerging Research + Design for Learning* 5: 12-13. <https://pubs.royale.com/publication/?m=31173&i=724524&p=12&ver=html5>.
- Bustamante, Andres S., Molly Schlesinger, Kreshnik Nasi Begolli, Roberta M. Golinkoff, Nabil Shahidi, Shekufeh Zonji, Cristina Riesen, Natalie Evans, and Kathy Hirsh-Pasek. 2020. "More Than Just a Game: Transforming Social Interaction and STEM Play with Parkopolis." *Developmental Psychology* 56 (6): 1041-56. <https://doi.org/10.1037/dev0000923>.
- Cray, Kate. 2020. "How the Coronavirus Is Influencing Children's Play." *The Atlantic*, April 1, 2020. <https://www.theatlantic.com/family/archive/2020/04/coronavirus-tag-and-other-games-kids-play-during-a-pandemic/609253/>.
- Dorn, Emma, Bryan Hancock, Jimmy Sarakatsannis, and Ellen Viruleg. 2021. "COVID-19 and Education: The Lingering Effects of Unfinished Learning." *McKinsey & Company*, July 27, 2021. <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-education-the-lingering-effects-of-unfinished-learning>.
- Education Analytics. 2021. "COVID-19 Impacts on Learning and Well-Being." Madison, WI: Education Analytics and Policy Analysis for California Education. https://www.edanalytics.org/assets/resources/202106_covid_impacts_on_learning_and_wellbeing_overview.pdf.
- Hassinger-Das, Brenna, Itai Palti, Roberta Michnick Golinkoff, and Kathy Hirsh-Pasek. 2020. "Urban Thinkscape: Infusing Public Spaces with STEM Conversation and Interaction Opportunities." *Journal of Cognition and Development* 21 (1): 125-47. <https://doi.org/10.1080/15248372.2019.1673753>.
- Hassinger-Das, Brenna, Jennifer M. Zosh, Nicole Hansen, Meghan Talarowski, Kate Zmich, Roberta Michnick Golinkoff, and Kathy Hirsh-Pasek. 2020. "Play-and-Learn Spaces: Leveraging Library Spaces to Promote Caregiver and Child Interaction." *Library & Information Science Research* 42 (1): 101002. <https://doi.org/10.1016/j.lisr.2020.101002>.

- Hassinger-Das, Brenna, Jennifer M. Zosh, Andres S. Bustamante, Roberta M. Golinkoff, and Kathy Hirsh-Pasek. 2021. "Translating Cognitive Science in the Public Square." *Trends in Cognitive Sciences* 25 (10): 816-18. <https://doi.org/10.1016/j.tics.2021.07.001>.
- Hirsh-Pasek, Kathy, Helen S. Hadani, Elias Blinkoff, and Roberta Michnick Golinkoff. 2020. "A New Path to Education Reform: Playful Learning Promotes 21st-Century Skills in Schools and Beyond." *Brookings Institution, Policy* 2020, October 28, 2020. <https://www.brookings.edu/policy2020/bigideas/a-new-path-to-education-reform-playful-learning-promotes-21st-century-skills-in-schools-and-beyond/>.
- INEE. 2010. *INEE Minimum Standards for Education: Preparedness, Response, Recovery*. New York: INEE. https://inee.org/system/files/resources/INEE_Minimum_Standards_Handbook_2010%28HSP%29_EN.pdf.
- INEE. 2020. *Education During the COVID-19 Pandemic*. New York: INEE. <https://inee.org/system/files/resources/INEE%20Technical%20Note%20on%20COVID-19%20EN%202020-04-23.pdf>.
- Ridge, Katherine E., Deena Skolnick Weisberg, Hande Ilgaz, Kathy Hirsh-Pasek, and Roberta Michnick Golinkoff. 2015. "Supermarket Speak: Increasing Talk among Low-Socioeconomic Status Families." *Mind, Brain, and Education* 9 (3): 127-35. <https://doi.org/10.1111/mbe.12081>.
- Romeo, Rachel R., Julia A. Leonard, Sydney T. Robinson, Martin R. West, Allyson P. Mackey, Meredith L. Rowe, and John D. E. Gabrieli. 2018. "Beyond the 30-Million-Word Gap: Children's Conversational Exposure Is Associated with Language-Related Brain Function." *Psychological Science* 29 (5): 700-10. <https://doi.org/10.1177/0956797617742725>.
- Sphere Project. 2011. *Humanitarian Charter and Minimum Standards in Humanitarian Response*. Rugby, UK: Practical Action Publishing. <https://www.unhcr.org/uk/50b491b09.pdf>.
- Spector, Carrie. 2021. "New Stanford Study Finds Reading Skills among Young Students Stalled during the Pandemic." *Stanford News*, March 9, 2021. <https://news.stanford.edu/2021/03/09/reading-skills-young-students-stalled-pandemic/>.
- Thompson, Ross A., and Susan D. Calkins. 1996. "The Double-Edged Sword: Emotional Regulation for Children at Risk." *Development and Psychopathology* 8 (1): 163-82. <https://doi.org/10.1017/S0954579400007021>.

CONTRIBUTORS

OKSANA BASENKO (o.basenko@ukma.edu.ua) is a Research Associate at the Center for Mental Health and Psychosocial Support at the National University of Kyiv-Mohyla Academy. For the last six years she has led the implementation of several international mental health, psychosocial support, and education projects in Eastern Ukraine for schoolteachers, psychologists, and social workers. The projects address community-based education in crisis management and suicide prevention, and psychosocial support services for conflict-affected children and adolescents. Her current research interests include developmental psychology, community resilience, and mental health initiatives for children and families.

ARLENE BENITEZ (arlene.benitez@crs.org) serves as Technical Advisor in Education at Catholic Relief Services, where she provides strategic guidance and support to a broad range of country programs, particularly in education in emergencies. She has previously led research and programming efforts in education in emergencies, teacher education, curriculum design, higher education, and citizenship/human rights education.

SERGIY BOGDANOV (s.bogdanov@ukma.edu.ua) is a psychologist and psychotherapist. He earned his doctorate at Vienna Medical University before returning to Ukraine, where he established the Research Center for Mental Health at the National University of Kyiv-Mohyla Academy. The center addresses the country's increased need for efficient, evidence-based mental health interventions. His work focuses on developing, testing, and implementing interventions for conflict-affected populations at the community level. One such intervention is SAFE SPACE, a multilayer psychosocial school-based program that is being implemented in the eastern regions of Ukraine.

FELIPE BOLÍVAR RINCON (fbrincon@cmcenters.org) is a Colombian board-certified addiction medicine and internal medicine physician who trained at the Mount Sinai Health System in New York. His work has focused on the provision of care to and research on underserved and marginalized populations. He has been part of a study group led by Universidad de los Andes in Bogota and Harvard University, which studies the health of the imprisoned population. He also worked with populations affected by the armed conflict in Colombia. He currently practices in California with those impacted by the opioid epidemic and other substance use disorders.

CONTRIBUTORS

VOLODYMYR CHERNOBROVKIN (chernobrovkin@ukma.edu.ua), a doctor of psychology and professor, is Head of the Department of Psychology and Pedagogy of the National University of Kyiv-Mohyla Academy. His research interests include resilience and protective processes in the field of mental health and child development. He has long studied various aspects of teachers' interactions with students.

VIRA CHERNOBROVKINA (chernobrovkina.vira@gmail.com) is a Lecturer in the Department of Psychology and Pedagogy at the National University of Kiev-Mohyla Academy. Her research and teaching interests include emergency education, postgraduate training, and continuing education for psychosocial support and mental health service providers in crisis settings, especially in Eastern Ukraine.

MAMOUR CHOUL TURUK (mamourchoul@gmail.com) is an Associate Professor of Applied Linguistics and TESOL at the College of Education-Upper Nile University. His research and teaching interests involve thinking skills, discourse analysis, psycholinguistics, and sociolinguistics. He was a lead research associate on the UNICEF psychosocial support intervention impact evaluation project conducted through the Integrated Essential Emergency Education Services for Protection, Resilience, and Recovery project in South Sudan. His published works revolve around L2 classroom interactions and the infusion of critical thinking in L2 classrooms.

CHARLOTTE COLE (charlotte.cole@butterfly.org) is the Cofounder and Executive Director of Blue Butterfly Collaborative, a nonprofit organization that uses children's media to advance international development aims in education and health. Prior to founding Blue Butterfly, she was Senior Vice President of Global Education at Sesame Workshop, where she oversaw the education and research activities associated with international coproductions of *Sesame Street*.

JENNIFER DEBOER (deboerj@purdue.edu) is currently an Associate Professor of Engineering Education and Mechanical Engineering (courtesy) at Purdue University. She conducts education research and supports diverse students from around the world as they access, develop, and apply engineering skills in their own communities. In 2017, she received the American Society for Engineering Education Mara Washburn Women in Engineering Early Engineering Educator Award, and she has won multiple awards from the National Science Foundation, American Education Research Association, the Spencer Foundation, and the US Department of State.

CONTRIBUTORS

MARWAN DIAB (diabmarwan@gmail.com) is a Psychologist and Research Affiliate at Stellenbosch University in South Africa, where he specializes in trauma and mental health. His research and teaching interests involve resilience, psychosocial interventions among war-affected populations, children's mental health, violence, suicide and addiction, and trauma-related mental health problems, particularly in the Middle East and Palestine. He has coauthored articles on decolonizing a psychology curriculum in the context of siege and integrating a public health and human rights approach.

NIKHIT D'SA (ndsa@nd.edu) is an Assistant Professor and Senior Associate Director for Research at the University of Notre Dame's Global Center for the Development of the Whole Child. A developmental psychologist, he has worked as an education technical advisor and applied researcher in low-resource and fragile settings. His research focuses on how to incorporate the lived experience of children, caregivers, and teachers into the development of context-specific assessments. He also researches how settings can be leveraged to support children's learning and development.

JOSÉ M. FLORES (jose.flores@yale.edu) is an Addiction Psychiatry Fellow at Yale University. He is interested in the treatment of substance use disorders and addiction in youth, as well as statistical and epidemiological methods that can be applied to mental health.

KIM FOULDS (kim.foulds@sesame.org) is Vice President of Content Research and Evaluation at Sesame Workshop, the media and education nonprofit behind *Sesame Street*. She leads a team that conducts research projects across the globe. Using an array of methodologies to support projects from conception to completion and working with academic and research institutions to evaluate efficacy, she works to translate data into recommendations that maximize the impact of Sesame Workshop's content around the world.

PAUL FRISOLI (paul.frisoli@lego.com) is a Senior Program Specialist who supports the Playful Learning in Crisis Settings Initiative at the LEGO Foundation. His programmatic and research interests focus on social and emotional learning, wellbeing, and comprehensive teacher professional development in conflict and crisis settings. His doctoral research explored the connections between teachers' continuous professional learning communities and their social-emotional wellbeing in acute conflict settings.

CONTRIBUTORS

RENASHA GHIMIRE (renasha5@gmail.com) is a Research Coordinator at KOSHISH Nepal. Her research interests include cultural adaptation, psychological processes and outcomes, and acceptance and mindfulness-based approaches in the treatment of mental health problems. She leads research on identifying risk, protective, and preventive factors related to suicide, and on measuring the efficacy of various psychosocial interventions in the Nuwakot district of Nepal. She previously worked at the Transcultural Psychosocial Organization Nepal as a clinical supervisor.

ANDRIY GIRNYK (girnyk@ukma.edu.ua) is a Professor in the Department of Psychology and Pedagogy of the National University of Kyiv-Mohyla Academy and a Merited Education Worker of Ukraine. His research interests include the theory and psychology of conflict and the study of stress and resilience. His teaching focuses on simulation-based learning in higher education. He is also head of the Center for Conflict Research and Psychoanalysis of the National University and president of the Ukrainian Conflict Resolution Association.

KALINA GJICALI (kg1317@nyu.edu) is a Research Scientist in measurement and evaluation at NYU Global TIES for Children. She uses quantitative methods to investigate the validity of academic and social and emotional learning measures and to evaluate the effectiveness of remedial education interventions on child outcomes. In collaboration with external partner organizations, she has worked on research initiatives that span multiple contexts, including Lebanon, Jordan, Ghana, Kenya, Niger, and Colombia.

OSCAR GÓMEZ (ogomez@saldarriagaconcha.org) is a Research Assistant at the Health and Welfare Division of the Fundación Saldarriaga Concha and a Professor at the Physiological Sciences Department of the Pontificia Universidad Javeriana. His work focuses on the appraisal of evidence, methods, and data management for applied research in the fields of pharmacology, mental health, and health systems. He also has worked to strengthen policies and improve the environment of vulnerable communities in Colombia.

LINA MARÍA GONZÁLEZ BALLESTEROS (lgonzalez@saldarriagaconcha.org; lgonzalezb@javeriana.edu.co) is a Public Health Officer in Fundación Saldarriaga Concha and a Professor in the Department of Psychiatry at Pontificia Universidad Javeriana. She has developed multiple programs on resilience in underserved areas and areas affected by violence. Her research focuses on the elderly, mental health, stigma, and resilience. In her recent projects, she studies the effects of compassion and resilience in communities affected by armed conflict and care, intergenerationality, and wellbeing of the elderly.

CONTRIBUTORS

HELEN SHWE HADANI (hhadani@brookings.edu) is a Fellow at the Center for Universal Education and the Bass Center for Transformative Placemaking at the Brookings Institution where she leads Playful Learning Landscapes, an initiative that brings together the fields of developmental science and placemaking with the goal of improving child and community outcomes. Prior to joining Brookings, she spent 20 years in research and education settings working for toy and media companies, including Disney, Sesame Workshop, Apple, LEGO, and Fisher-Price. She holds a Bachelor of Arts in Cognitive Science from the University of Rochester and a Doctorate in Psychology from Stanford University.

KATERYNA HARBAR (h.kateryna@ukma.edu.ua) is a Junior Researcher in the Center for Mental Health and Psychosocial Support at the National University of Kyiv-Mohyla Academy. Her research interests include school climate, wellbeing, and implementation science. Her work involves studying mental health and psychosocial problems among children in Ukraine and the evaluation of psychosocial interventions.

SASCHA HEIN (sascha.hein@fu-berlin.de) is a Professor in the Department of Education and Psychology at the Free University of Berlin. His research focuses on the impact of formal and nonformal learning and parenting on child and adolescent development in diverse cultures. In his recent projects, he studies the development of children growing up in crisis-affected, humanitarian, and marginalized contexts.

KIMBERLY HOOK (kimberly.hook@bmc.org) is a Global Clinical Research Fellow and Clinical Psychologist affiliated with the Boston University Medical Campus-Massachusetts General Hospital. Her research interests span the fields of public health, behavioral health (specifically mood and substance use disorders), and global health. Her work focuses on increasing access to evidence-based mental health and substance use treatments by adapting psychological interventions to be implemented in underresourced settings.

IRINA IVANYUK (irinaivanyuk72@gmail.com) is a Senior Researcher in the Institute of Informational Technologies and Learning Tools of the National Academy of Pedagogical Sciences of Ukraine and a Project Coordinator at the Center for Mental Health and Psychosocial Support at the National University of Kyiv-Mohyla Academy. Her research interests include education policy development and the professional development of teachers of civic education. For the last six years, she has led the implementation of several education projects in Eastern Ukraine aimed at providing psychosocial support for teachers, school psychologists, and children.

CONTRIBUTORS

MARK J. D. JORDANS (mark.jordans@warchild.nl) is a Professor of Child and Adolescent Global Mental Health at the University of Amsterdam, and at the Center for Global Mental Health, King's College London. A child psychologist, he also works as Director of Research and Development for the Netherlands-based War Child project. His research interests are the development, implementation, and evaluation of psychosocial and mental health care systems in low- and middle-income countries, especially for children living in fragile states.

HA YEON KIM (hayeon@nyu.edu) is a Senior Research Scientist at NYU Global TIES for Children. She is a developmental and educational psychologist interested in measuring, examining, and supporting the academic and social-emotional development of children in low-resource settings globewide. Her current research focuses on understanding the experiences of children and teachers and supporting their academic, social-emotional, and professional learning and development in crisis-affected settings across the Middle East and North Africa, Turkey, sub-Saharan Africa, Latin America, and South and Southeast Asia.

SHANNA KOHN (shanna.kohn@sesame.org) is Senior Education Manager of Humanitarian Programs at Sesame Workshop. She leads the educational agenda for *Ahlan Simsim*, an Arabic-language adaptation of *Sesame Street*, and designs teaching and learning materials for children affected by displacement.

BRANDON A. KOHRT (bkohrt@gwu.edu) holds the Charles and Sonia Akman Professorship in Global Psychiatry at George Washington University. He is also a Professor of psychiatry and behavioral sciences, global health, and anthropology, and Director of the Global Mental Health Equity Lab. He has worked with populations affected by war and political violence, disasters, and other forms of adversity around the world. His work also addresses reducing stigma in healthcare settings to improve the quality of mental health services.

ALLYSON KRUPAR (akrupar@savechildren.org) is a Research, Evidence, and Learning Advisor at Save the Children US and an Adjunct Professor at the American University School of International Service and School of Professional and Extended Studies. Her research focuses on education in crisis and conflict contexts, with a particular focus on children, youth, and adults who have missed out on formal education. She also researches wellbeing, social-emotional learning, and community and home-based interventions in emergency contexts.

CONTRIBUTORS

POOJA LAKSHMIN (plakshmin@gmail.com) is a psychiatrist and author specializing in women's mental health, in particular on how broken systems impact women's emotional lives. She is a contributor to the *New York Times*, a medical advisor to Peloton, and the founder of Gemma, the first digital education platform dedicated exclusively to women's mental health. She is also an Assistant Professor at George Washington University.

AMALIA LONDOÑO TOBÓN (amalia.londonotobon@yale.edu) is a psychiatrist and researcher with expertise in the perinatal, childhood, family, and cultural aspects of mental health. Her clinical and research work aims to improve and disseminate mental health prevention and treatment interventions during pregnancy and the early childhood years, particularly for underserved and underresourced communities.

NAGENDRA P. LUITEL (luiteln@gmail.com) is a public health researcher from Nepal. He is the Head of Research at Transcultural Psychosocial Organization Nepal, a leading psychosocial and mental health organization in that country. His research interests include assessment of mental health problems and service needs among populations affected by humanitarian emergencies, the adaptation and validation of mental health instruments, and the development and evaluation of community-based interventions to increase the use of mental health services.

MACKENZIE MATTHEWS (mackenzie.matthews@rescue.org) is the Technical Advisor for Curriculum Development at the International Rescue Committee (IRC). She leads IRC's strategic effort to ensure that education curricula are tailored to the needs, values, and beliefs of clients and is reflective of each implementation context. Having begun her career as a classroom teacher, she brings to bear practical understanding of how theory meets practice. She works across IRC's education projects to support the innovative and meaningful development of content for children, families, and teachers.

NINA MENEZES CUNHA (ncunha@fhi360.org) has more than nine years of international development experience, with a focus on education. She currently is a Senior Research Associate and Monitoring and Evaluation Specialist at FHI 360. Her research supports the design, implementation, and publication of quantitative research on key issues in education. Her recent work includes developing and evaluating the psychometric properties of a teacher wellbeing assessment and a tool to measure soft skills among youth, and a research study to measure equity in education resource allocation across different countries.

CONTRIBUTORS

MOSES OLAYEMI (molayemi@purdue.edu) is a Graduate Research Assistant in the School of Engineering Education at Purdue University. His research revolves around the development of local competency in sub-Saharan Africa. He has been working within the sub-Saharan African STEM education landscape as a teacher trainer, teacher, and researcher for close to a decade. He aspires to leverage research-based empirical evidence to influence education policies in Africa.

ANA MARÍA ORTIZ HOYOS (aortiz@saldarriagaconcha.org) is a Coordinator of Health and Wellbeing in Fundación Saldarriaga Concha, a psychologist, and a master's candidate in public health. She has experience in the design, implementation, and management of social intervention programs, which have focused on promoting sustainability and improving the quality of life for different age groups in vulnerable populations, with an emphasis on mental health and integral development issues.

GLORIA A. PEDERSEN (gapedersen@gwu.edu) is a Research Associate in the Department of Psychiatry, Division of Global Mental Health, at the George Washington University School of Medicine and Health Sciences. She has been part of the Ensuring Quality and Psychological Support team since it launched in 2018, where she contributes to the review, development, and evaluation of performance-focused competencies for mental health and psychosocial support workforce training, and to the implementation of psychological and psychosocial manualized care.

KIRSI PELTONEN (kirsi.peltonen@tuni.fi) is a Senior Researcher at the University of Turku, Finland, and a Docent in Mental Health Psychology. Her research focuses on the mental health of children and adolescents exposed to violence and on interventions to help them. She has specifically studied preventive and targeted interventions for refugee children. She has trained professionals in different fields to use evidence-based methods to prevent and treat traumatization, and she conducts her research in close collaboration with hospitals, schools, and accommodation units.

CONTRIBUTORS

LILIANA ANGÉLICA PONGUTA (angelica.ponguta@yale.edu) is a member of the research faculty at the Yale Child Study Center. Her work is to advance early childhood development and positive youth development in low- and middle-income countries. She has worked in policy formulation and analysis and in program development and evaluation in more than 20 countries in Africa, Southeast Asia, Eastern and Central Europe, Latin America, and the Middle East. Her research also focuses on strategies to strengthen early childhood and youth development in humanitarian crisis settings.

RAIJA-LEENA PUNAMÄKI (raija-leena.punamaki-gitai@tuni.fi) is a Psychologist and Professor Emerita at Tampere University, Finland, who specializes in mental health, trauma, and child development. Her research focuses on therapeutic interventions in families with a risk of violence, substance abuse, and psychiatric disorders, and on infant and child socioemotional and cognitive development, early dyadic interaction, and family attachment. She also conducts developmental research on the medical and psychological determinants of socioemotional and psycho-physiological development and on the effectiveness of psychosocial school-based interventions among trauma-affected children and refugees.

TOM PUREKAL (tpureka1@nd.edu) currently manages a cross-section of programs in the Pulte Institute's Innovation and Practice Division. He previously spent nearly a decade working for Catholic Relief Services in India, South Sudan, and Myanmar. His experiences encompass peacebuilding and governance, education, disaster risk reduction, and water, sanitation, and hygiene. He is the Program Director of Supporting Holistic and Actionable Research in Education and an Assistant Teaching Professor at the Keough School of Global Affairs.

SAMIR R. QOUTA (samir.qouta@dohainstitute.edu.qa) is a Clinical Psychologist and Professor of Psychology at Doha Institute for Graduate Studies. His specialization is trauma and its social and mental health impact on children and families. His research projects include studies on mental health and child development in families exposed to traumatic events, therapy and evidence-based research among war-traumatized children and adults and torture survivors, and epidemiological research on psychiatric distress and psychological resources in the Palestinian community. He teaches science and therapy practices, and developmental, clinical, and health psychology.

CONTRIBUTORS

MANASWI SANGRAULA (sangraum@newschool.edu) is a postdoctoral researcher at The New School in New York City. Her research is focused on increasing access to mental health care using task-sharing interventions in New York City and with migrant communities in Latin America. She also spent several years working in Nepal with Group Problem Management Plus.

ALISON SCHAFER (aschafer@who.int) has worked in the humanitarian aid and development sector for more than 20 years. Trained in clinical psychology, she currently works for the World Health Organization leading the implementation of the Special Initiative for Mental Health. She previously coordinated the development of an online platform for assessing the competency of mental health service providers. She also spent 17 years working with World Vision Australia's International Humanitarian and Emergency Affairs teams in Asia, the Middle East, and Africa.

PRAGYA SHRESTHA (pragyaspecial@gmail.com) is a therapist and a Research Fellow with Transcultural Psychosocial Organization Nepal. She has training in many group-based therapies, such as group problem management, interpersonal psychotherapy, and Gestalt psychotherapy. She believes that a group is the best resource when dealing with an individual problem. This belief has enabled her to contribute to the development of basic group facilitation skills that can guide new group facilitators, supervisors, and trainers.

FERNANDA SOARES (fsoares@fhi360.org) is a Technical Advisor in the Research and Evaluation Department in the Global Education, Employment and Engagement Business Unit at FHI 360. She leads the research, monitoring, and evaluation of education and youth programs in low-resource, crisis, and conflict settings. She has been the principal investigator on studies of in-service and pre-service teacher professional development, teacher wellbeing, social and emotional learning, and youth development. She has employed a variety of qualitative, quantitative, and mixed methods approaches to conduct formative studies, impact assessments, implementation science research, and to develop measurement tools.

CONTRIBUTORS

CARLY TUBBS DOLAN (carly.tubbs@nyu.edu) is a Deputy Director at NYU Global TIES for Children, an international research center she helped found and launch. She leads a portfolio of research focused on education measurement, systems, and scaling in contexts of displacement. Her work involves collaborating with research-practice-policy partners in Lebanon and Peru to develop, adapt, test, and use data from holistic learning measures at a national scale.

MELISSA TUCKER (melissa.tucker@crs.org) is a Technical Advisor for psychosocial support and wellbeing in humanitarian response at Catholic Relief Services. She has a background in clinical psychology focused on trauma and complex humanitarian crises. She brings her experience in direct service provision and applied research to her support for country programs in mental health and psychosocial support activities, and to promoting the integration of these activities into other sectors, including education and peacebuilding.

ALEXANDER VINOGRADOV (a.g.vinogradov@knu.ua) is an Associate Professor in the Social Psychology Department at the Taras Shevchenko National University of Kyiv. He teaches courses in statistical data analysis, research methodology, and human resource management for psychology and sociology students. His research interests involve personal construct theory, interactionism, a social-cognitive approach to the assessment of personality, and multilevel modeling. His most recent research is with the International Situations Project and a comparative project with the University of Queensland that examines the effects of economic inequality on stereotypes.

MIKE WESSELLS (mikewessells@gmail.com) is a Professor in the Columbia University Program on Forced Migration and Health. A long-time psychosocial and child-protection practitioner, he has conducted extensive research on the holistic effects of war and political violence on children. He currently is lead researcher of an interagency, multicountry project on community-led child protection that features the importance of education. This work involves learning from children's lived experiences of violence and how they cope and remain resilient in a context of adversity. He regularly advises UN agencies, governments, and donors on issues of child protection and psychosocial support, including in communities and schools.

CONTRIBUTORS

WENDY WHEATON (wendylwheaton@gmail.com) has more than 20 years of experience in developing education standards, policies, guidance, programs, and technical response modalities, with particular expertise in conflict and crisis settings. Currently an education advisor at USAID, she has previously held positions at the World Bank, the United Nations, with international nonprofit organizations, and with academic technical groups focused on education, child protection, and mental health and psychosocial support. Her research interests include social and emotional learning for children and youth who have been displaced and exposed to violence, resilience through education, education in emergencies, and gender-based violence.

REBECCA WINTHROP (rwinthrop@brookings.edu) is a Senior Fellow and Codirector of the Center for Universal Education at the Brookings Institution. Her research focuses on global education, with special attention to the skills young people need to thrive in work, life, and as constructive citizens. She works to promote quality and relevant education, including exploring how education innovations can leapfrog progress, particularly for the most marginalized youth. She advises governments, international institutions, foundations, civil society organizations, and corporations on education issues. She currently serves as a board member and advisor for a number of global education organizations and lectures at Georgetown University.

ZEZHEN (MICHAEL) WU (zw1429@nyu.edu) is a doctoral student in the Psychology and Social Intervention program at New York University. His research examines how children in low-resource settings in different cultures develop strategies, norms, and mindsets to cope with stress. He also looks at how brief interventions can buffer learners against various situational and environmental threats, stigma, and discrimination, and how to design and implement these interventions to help children deal with developmental and contextual challenges.

JOURNAL ON EDUCATION IN EMERGENCIES

The *Journal on Education in Emergencies (JEiE)*, a scholarly, peer-reviewed journal, aims to fill gaps in education in emergencies (EiE) research and policy. Building on the tradition of collaboration between practitioners and academics in the EiE field, *JEiE*'s purpose is to improve learning in and across service-delivery, policymaking, and academic institutions by providing a space where scholars and practitioners can publish rigorous quantitative, qualitative, and mixed methods research articles, and robust and compelling field notes that both inform policy and practice and stir debate. *JEiE* provides access to the ideas and evidence needed to inform sound EiE programming, policymaking, funding decisions, academic program curricula, and future research.

JEiE specifically aims to:

1. **Publish rigorous scholarly and applied work** that sets the standard for evidence in the field
2. **Stimulate research and debate** to build evidence and collective knowledge about EiE
3. **Promote learning across service-delivery organizations, academic institutions, and policymakers** that is informed by evidence
4. **Define knowledge gaps and key trends** that will inform future research

To achieve these goals, *JEiE* seeks articles from scholars and practitioners who work across disciplines and sectors on a range of questions related to education in countries and regions affected by crisis and conflict. *JEiE* is part of and works closely with the Inter-agency Network for Education in Emergencies (INEE), today an open global network of more than 18,000 individual members affiliated with more than 4,000 organizations and institutions in 190 countries, to collect new research articles and field note submissions and to distribute high-

quality published work. This large global partnership of activists, academics, policymakers, and practitioners in education enables *JEiE* to make a unique and powerful contribution.

STRUCTURE OF *JEiE*

According to the INEE Minimum Standards (inee.org/standards), education in emergencies is defined as “quality learning opportunities for all ages in situations of crisis, including early childhood development, primary, secondary, non-formal, technical, vocational, higher and adult education.” *JEiE* publishes research related to education activities in the context of natural disasters and fragile or conflict-affected states, conflict-sensitive education, attacks on education, education for peacebuilding, peace education, education for resilience and disaster risk reduction, and forced migration and education.

ISSUES AND CONTENTS

Each issue features peer-reviewed articles written by researchers and practitioners in the field of EiE. The three sections of *JEiE* are:

EiE Research Articles (Section 1): Articles in this section have a clear research design; use an explicit, well-recognized theoretical or conceptual framework; employ rigorous research methods; and contribute evidence and advance knowledge on EiE. Articles that develop new EiE theoretical or conceptual frameworks or challenge existing ones are also welcome. Qualitative, quantitative, and mixed methods articles are appropriate.

EiE Field Notes (Section 2): Articles in this section address innovative approaches to EiE; progress and challenges in designing, implementing, and evaluating initiatives; and/or observations and commentary on research work. Articles in this section typically are authored by practitioners or practitioner-researcher teams.

EiE Book Reviews (Section 3): Articles in this section offer a critical review of a recently published or upcoming book, or of substantial studies, evaluations, meta-analyses, documentaries, or other media that focus on EiE.

Please see our website (inee.org/journal) for more information and detailed submission guidelines.

JOURNAL ON EDUCATION IN EMERGENCIES

DANA BURDE, EDITOR-IN-CHIEF

The *Journal on Education in Emergencies (JEiE)* publishes groundbreaking and outstanding scholarly and practitioner work on education in emergencies (EiE).

JEiE was established in response to the growing need for rigorous research to strengthen EiE evidence, support policy and practice, and improve learning in and across organizations, policy institutes, and academic institutions. *JEiE* facilitates EiE knowledge-generation and sharing, thus contributing to the professionalization of the EiE field.

<https://inee.org/journal>
journal@inee.org

ISSN 2518-6833