

Conducting longitudinal, process-oriented research with conflict-affected youth: Solving the inevitable challenges

ERIC F. DUBOW,^{a,b} J. LAWRENCE ABER,^c THERESA S. BETANCOURT,^d E. MARK CUMMINGS,^e AND L. ROWELL HUESMANN^b

^a*Bowling Green State University*; ^b*University of Michigan*; ^c*New York University*; ^d*Harvard T. H. Chan School of Public Health*; and ^e*University of Notre Dame*

Abstract

The reader might get the impression that the four projects described in this Special Section proceeded in a systematic and predictable way. Of course, those of us engaged in each research project encountered pitfalls and challenges along the way. A main goal of this Special Section is to provide pathways and encouragement for those who may be interested in advancing high-quality research on this topic. In this paper, we describe a set of practical and ethical challenges that we encountered in conducting our longitudinal, process-oriented, and translational research with conflict-affected youth, and we illustrate how problems can be solved with the goal of maintaining the internal and external validity of the research designs. We are hopeful that by describing the challenges of our work, and how we overcame them, which are seldom treated in this or any other literature on research on child development in high-risk contexts, we can offer a realistic and encouraging picture of conducting methodologically sound research in conflict-affected contexts.

In this paper, we describe examples of challenges that we judge to be illustrative of conducting longitudinal and translational research on youth exposed to political violence and armed conflict. First, we describe practical challenges of conceptualizing the research, developing collaborations with regional partners, and obtaining funding for the project. Second, we describe challenges encountered in planning aspects of the research design and implementing it in such high-risk settings. This section will be organized around threats to the projects' internal and external validity; we provide specific examples of design/implementation challenges and how the teams solved them to maintain the scientific integrity of the research. We end with a section on ethical challenges commonly faced in conducting this kind of research.

Practical Challenges: Conceptualization, Collaborations, and Funding Opportunities

The research teams' principal investigators are at universities in the United States, yet all have interests in studying conflicts in various regions around the world. In this section, using examples from the four research teams' experiences, we focus on the motivations guiding their interests in studying children exposed to political violence, how they developed regional

collaborations, and how they sought funding opportunities for their international research.

The two research teams of Huesmann and Cummings are similar in terms of the motivations for their interests in political violence and children (i.e., to expand their extant research programs on psychological process models accounting for exposure to conflict and violence effects on youth to a higher social-ecological context: the political/cultural context). These teams also are similar in the pathways toward developing regional collaborations, and in their funding sources.

Huesmann's team at the University of Michigan's Institute for Social Research, the Aggression Research Group, had conducted several projects focused on the effects of exposure to violence on children, having articulated a social-cognitive process mediating model explaining how exposure to violence across contexts (e.g., family, media, and neighborhood) affects children's aggressive behavior (Huesmann, 1988, 1998; Huesmann & Eron, 1984; Huesmann & Kirwil, 2007). The team was interested in applying their model to the context of children's exposure to political violence (see Dubow, Huesmann, & Boxer, 2009), and was excited when the opportunity arose at the NIH for funding this type of research. The NIH funding mechanism for the research conducted by both Huesmann's and Cummings' teams was overseen by Valerie Maholmes' Social and Affective Development/Child Maltreatment and Violence Program in the Child Development and Behavior Branch at the Eunice Kennedy Shriver National Institute for Child Health and Human Development. At the time, Huesmann was leading an internally funded Roots of Terrorism Research Initiative at the

This research was supported by Grant HD047814 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Address correspondence and reprint requests to: Eric F. Dubow, Research Center for Group Dynamics, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, MI 48106-1248; E-mail: edubow@bgsu.edu.

Institute for Social Research and invited as an initial speaker an Israeli psychology professor, Ariel Merari, who had interviewed a number of “failed suicide bombers.” The team discussed the possibility of conducting a research project with Israeli and Palestinian children with Professor Merari, who suggested an Israeli professor as a potential collaborator, who in turn suggested two potential collaborators: Simha Landau, a professor of criminology at the Hebrew University in Jerusalem, and Khalil Shikaki, the director of the Palestinian Center for Policy and Survey Research. Huesmann had previously collaborated with Landau, and other faculty at the University of Michigan had previously collaborated with Shikaki. A series of meetings with Landau and Shikaki led to an agreement on the research design and protocol and then the submission of the grant application to NIH that was eventually funded.

Cummings’ research career has focused on child development in families, particularly the effects of marital conflict and violence on children, with emotional security theory (Davies & Cummings, 1994) as a theoretical process model for accounting for the impact of conflict and violence in the family on child adjustment. Given his interests in conflict and conflict resolution, after he joined the faculty at the University of Notre Dame, he was appointed as a fellow at the Kroc Institute for International Peace Studies. Peace studies was inclusively defined at the Kroc, including fellows from political science, sociology, history, theology, and anthropology, among others; and Cummings was the only fellow actively engaged with the institute from a child development perspective. His view was that understanding psychological effects at the level of communities, families, and children was essential to the promise for lasting peace in any context of political violence. After a few unsuccessful attempts to connect with potential collaborating faculty members in Northern Ireland, Cummings received an invitation from the Kroc Institute to attend a talk and subsequent dinner for Ed Cairns, the most prominent scholar on political violence and children in Northern Ireland. As Cummings recounts,

Ed began his talk in the Hesburgh Center at Notre Dame by saying that there were three high-quality reviews of political violence and children in Northern Ireland, with one of the best written by Lovell and Cummings (2001) and published on-line by Kroc (this was based on a political science senior thesis of an undergraduate student that Cummings had chaired!). At that point, the AV system in the Hesburgh auditorium went down and the talk ended prematurely, but what had been said was enough to provide me with an entry to conversation with Cairns at dinner about possible research in this area.

Cummings soon after became aware of the same NIH funding mechanism noted above, and he, Cairns, and the Cummings research team developed a successful grant application, which was supplemented in later phases with major funding from the Northern Ireland government through a grant from the Office of the First Minister, Deputy First Minister.

Aber’s team at New York University was approached by the International Rescue Committee (IRC) to evaluate the

Learning to Read in a Healing Classroom (LRHC) intervention in the Democratic Republic of the Congo (DRC). LRHC is a universal school-based program developed by the IRC that uses a teacher professional development system to improve primary school-aged children’s academic skills and socioemotional development in conflict-affected countries. The IRC was motivated by a desire to generate rigorous evidence as to what might happen in schools and to children as a result of the LRHC program, versions of which they are currently using in more than 12 countries. Stakeholders from the IRC and New York University thus collaborated on developing a multilevel program theory of change, hypothesizing the mechanisms by which LRHC operated to change children’s outcomes. With funding from the US Agency for International Development and an anonymous donor and in collaboration with the DRC’s Ministry of Education, the IRC mounted a systematic effort to increase opportunities for equitable access to quality basic education for Congolese children and youth.

Research Design Pitfalls and Challenges

Conducting research in conflict-affected areas around the globe is fraught with foreseen and unforeseen circumstances that can threaten the integrity of the research design. Longitudinal, process-oriented, social-ecological research from a developmental psychopathology perspective places high demands on methodology, including measurement. Each of the research teams contributing to this issue encountered significant pitfalls and challenges in endeavoring to adequately measure key constructs in these multiple highly challenging contexts of armed conflict and political violence.

For example, fundamental requirements for measurement involve matters relating to internal and external validity. In their seminal textbooks on research methodology, Campbell and Stanley (1966) and Cook and Campbell (1979) described the concepts of internal and external validity and specific design/implementation threats to each. Briefly, internal validity focuses on the methodological integrity of the research design, for example, the fidelity of implementation to the planned design, the operational definitions of the theoretical constructs, whether historical and contextual variables other than the hypothesized key causal/predictor variables might be the actual variables affecting the outcome of the study, the quality of the measures, and in experimental and repeated measures designs, whether there is differential attrition across experimental groups or subsamples being compared. A research design is internally valid if we can reasonably infer that it is the experimental treatment, or in correlational research, the “predictor” variables, rather than some other variables, that predict/cause the hypothesized outcomes.

External validity refers to the degree to which the study’s results are generalizable to other samples or settings. Factors that affect external validity include how participants are selected (e.g., random sample or high-risk sample) and possible effects of multiple testing, among others.

Threats to internal validity

Fidelity of implementation of research design. Despite researchers' best efforts to develop internally valid research designs in high-risk contexts, implementation problems often cannot be avoided. Mobray, Holter, Teague, and Bybee (2003) noted, "without documentation and/or measurement of a program's adherence to an intended model, there is no way to determine whether unsuccessful outcomes reflect a failure of the model or failure to implement the model as intended" (p. 317). Of the four studies in this Special Section, Aber's research in the DRC and Betancourt's research in Sierra Leone faced significant implementation challenges to internal validity.

The DRC consistently ranks among the bottom five countries of the world on the Human Development Index, due to a long history of colonialism, political instability, civil war, government corruption, and exploitation of natural resources. Logistical challenges were encountered in training local field researchers, in difficulty accessing schools, in the lack of infrastructure to support technical solutions, in difficulty in transporting necessary materials, and the high cost associated with all of these. Although the DRC is often the focus of educational research related to topics such as school enrollment and child nutrition, Aber's team (Global TIES for Children, the international research center he codirects at New York University; and the IRC) were the first to implement a rigorous randomized control trial designed to estimate the impact of a school-based intervention on student social-emotional outcomes and student-perceived school climate in the DRC. However, pioneering such research from an outsider's perspective is difficult. In collaboration with the DRC's Ministry of Education, local Congolese field researchers and consultants had to be employed; however, this required extra commitment to training as it was difficult to find those with relevant experience in rigorous field research. These logistic problems threatened the research design, and thus the researchers' ability to conclude with confidence that their intervention, rather than other factors, was responsible for observed achievement and socioemotional outcomes. To cope with these threats to the fidelity of the implementation of the intervention, training involved developing materials and holding training sessions that met the needs of workers (e.g., providing transportation, meals, and lodging is standard practice, rather than a simple flat hourly payment that is common in the United States in the developmental and educational sciences). Ensuring the safety of frontline researchers in the field was particularly important as many schools are remote and require multiday trips on narrow, poorly maintained roads that may flood or experience landslides. The remote nature of schools and use of paper surveys (at the time) meant that researchers had to stay for 3 days at schools in order to interview all children, perform nightly data-quality checks to ensure individual surveys were complete, and train extra field researchers to allow for dropout or illness. Traveling in groups was paramount to safety in areas where travelers might

be accosted or where there were no phone networks. These conditions placed further constraints on transporting necessary materials for research, such as equipment that had to be hand delivered in the absence of any reliable postal system. All of the above generated excessive operational costs, well above the typical cost of conducting research in high-income countries. For example, Aber's team had to budget for unforeseen incidental expenses such as paying "tolls" for passage along a road, or printing surveys at a cost many times higher than in places where resources are readily available. These challenges demonstrate the necessity of an implementing partner who understands local conditions and how to anticipate them when planning a sufficient budget.

Betancourt's team in Sierra Leone faced similar problems and sought similar solutions. Sierra Leone faced 11 years of civil war prior to the implementation of the project. Deaths caused by the war extended beyond violent killings or injuries, including conditions such as economic insecurity, weakened community structures, insufficient social services, poor access to education, destruction of local economies, and declines in health infrastructure (Betancourt et al., 2008; Donnelly, 2011).

Although research on evidence-based mental health interventions is commonplace and accepted in Western countries, this is not the case in most war-torn regions. As Aber's team recommended, this necessitates knowledgeable and resourceful local partners. Further, Betancourt's concern was that the power differential that exists between researchers and the very vulnerable population would threaten implementation of the research design. Betancourt's team utilized concepts from community-based participatory research to ensure positive and engaged involvement of the community. Community-based participatory research relies on a collaborative partnership that equitably and actively involves community partners in all aspects of the research process (e.g., Israel, Schultz, Parker, & Becker, 1998; Krieger et al., 2002). For instance, research leadership engaged in capacity building among community partners on methodological issues, such as techniques for qualitative or quantitative data analysis, while partners' knowledgeable about the community of focus provided historical and cultural knowledge, thus ensuring two-way learning. Data collection using community-based participatory research methods meant the involvement of both community and academic partners in designing and implementing data collection. Once collected, such data were analyzed in a collaborative process and iteratively applied to the design, implementation, and evaluation of intervention models that were a good fit to the culture and context. For instance, in work with local community partners in Sierra Leone on the issue of intervention components, Betancourt's team learned about Sierra Leonean proverbs that captured the essence of key messages the intervention team was looking to impart to affected youth. Some examples include the proverb "Yu need for paddle yu one canoe," which means, "One has to set one's own direction in life." This proverb and related terms in the local language were then integrated into the manual for the Youth

Readiness Intervention, making it easy for participants to connect to the core intentions of the intervention.

Historical/contextual factors. Perhaps the ultimate challenge of implementing research in a conflict-affected country is the conflict itself. In the DRC, an outbreak of rebel activity in North Kivu in 2012 resulted in a temporary suspension of the intervention and forced the research and program team to make a decision about whether to continue surveying students and teachers in North Kivu. Halting surveys in North Kivu significantly impacted the original research design: it removed 13% of the total sample, and thus reduced the power of the team's impact estimates. Yet, at the time the conflict broke out, there was no way to tell how long it would last: what ended up being a temporary program suspension at the time could have been a permanent suspension. Responding productively to these obstacles requires the willingness and ability to adapt the research design. The collaborators carefully reviewed the design each year to come up with solutions that were methodologically sound and true to the original research priorities of the project. For example, the lack of a system to track individual students and teachers was addressed in part by controlling for prior-year performance at the school level, rather than the child level. In every case, whatever adaptations were made were always done to protect the internal validity of the study.

Huesmann's team faced a similar issue. For the Palestinian sample, the third wave of interviewing was briefly interrupted by the 3-week Gaza war (December 2008–January 2009), and the fourth wave of data collection in Israel was nearly at its end when a 1.5-month conflict erupted with Gaza, essentially ending data collection just short of the intended goal of the number of participants. These outbreaks of conflict afforded a "natural experiment" (an interrupted time series), allowing the research team to examine whether upsurges in conflict affected key adjustment outcomes, and whether these results were differential across the region (e.g., Israel, Gaza, and West Bank).

With regard to similar matters, Cummings' team had to time data collection to avoid the "marching season," which in Northern Ireland is a period of chronic sectarian parades and associated rioting or other forms of political violence. In addition to the dangers associated with being physically present in conflictual Belfast neighborhoods during this period, many residents typically leave the country during this period or are otherwise unavailable for participation in research. In addition, it was important to time data collection each year to be similarly proximate in time to the marching season, because of the possibility of carryover of hostilities or other residues of the marching season into subsequent months.

Specific measurement pitfalls and challenges. In this section, we address several measurement-related pitfalls and/or challenges that the teams encountered that could have affected the internal validity of their projects. As is the case with

much research in non-Western countries, one of the major challenges in conducting research with children in conflict-affected countries is the many technical challenges found in implementing valid and reliable measurement in understudied groups. For example, Huesmann's and Cummings' teams both spent the initial year of their projects conducting focus groups and pilot testing, primarily to develop several new measures and assessments. Both teams made efforts to include measures that had been used previously with samples in their regions. However, both teams needed to develop measures that clearly distinguished sectarian/ethnic-political violence from nonsectarian community violence, and these were a critical aspect of focus groups and pilot testing. Huesmann's team followed an intensive four-stage process for selection of measures. First, over the first several months of the project, the investigators met with regional collaborators to choose potential measures (with a focus on those used previously in the Middle East if possible) and to adapt items for the three ethnic subsamples (Israeli Jews, Israeli Arabs, and Palestinians). Second, the original English measures were translated and back-translated for accuracy by native-speaking research teams at the data collection sites. Third, regional collaborators conducted youth focus groups in each region for each age cohort (8-, 11-, and 14-year-olds) separately. Fourth, they conducted two rounds of pilot testing on the entire survey on nine parent-child dyads (three from each age group) in each region. The pilot testing included asking participants to discuss any items or response formatting that caused confusion. All interactions with participants were conducted with same-ethnicity interviewers, and the surveys were presented in appropriate native languages by region/ethnicity (i.e., Hebrew for Israeli Jews, and Arabic for Palestinians and Israeli Arabs; Israeli Arabs were able to select Hebrew or Arabic). Once the data were collected and prepared for analysis, the measures were first subjected to confirmatory factor analysis to determine whether there was measurement invariance across the three ethnic subgroups.

Similarly, the measurement development work for the Cummings group involved a multistage process. The first stage involved the development of focus group guidelines and prompts and subsequently the conduct of qualitative analyses by means of focus groups to derive ecologically valid information about everyday expressions of sectarian and nonsectarian community violence, and indicators of children's emotional insecurity in the community, from mothers living in conflictual, interfaced neighborhoods in Belfast. Interfaced neighborhoods were neighborhoods that were highly segregated by groups, typically among the most socially deprived in Northern Ireland, and were positioned adjacent to neighborhoods reflecting the other ethnic group. Next, members of the research team from both the United States and Northern Ireland met to derive initial items to reflect these constructs, including expressions of items in the manner employed by residents of Belfast, based on discussion among Northern Ireland and US group members and theoretical conceptualizations of the constructs based on past research and

theory on conflict, violence, and emotional security. The third step was to pilot test the preliminary versions of these scales in a two-wave longitudinal study including 130 mothers (104 participated in both phases) living in Derry–Londonderry, Northern Ireland. Fourth, psychometric tests were conducted to further refine the scales, including exploratory factor analyses, calculations of alpha levels, and tests of predictive validity. Aber's team noted the importance of a culture-specific understanding of key constructs. Attempts to evaluate concepts such as "cultural diversity" are difficult, given the complex and diverse history of the DRC. Without specific insider knowledge of regional, ethnic, or tribal histories, proxies for cultural diversity such as language or geography can be proffered as an educated guess, but remain only proxies until more in-depth validation is conducted. Local partners who are more familiar with both the history and the modern sociopolitical climate of the region are integral to helping researchers make informed decisions for measurement implementation and interpret unexpected findings. Smith's commentary on Aber's project noted a need to understand socioemotional constructs specific to the culture.

Research into children's own definitions of constructs such as victimization would provide a deeper understanding of their experiences and how those experiences connect to distal outcomes. For example, a child in DRC may consider it more "victimizing" to be labeled a particular ethnic group than to be called "a bad name." In other words, discerning the qualities that children themselves associate with a concept like victimization enhances the validity of the measure and provides a more complete picture of the experiences to be analyzed.

Aber's team also faced very specific measurement-related constraints based on donor requirements: the US Agency for International Development required all projects to use the Early Grade Reading Assessment to measure improvements in reading. However, after baseline testing, it was clear that many children were unable to perform most of activities in the measure, resulting in a very large proportion of zero scores (extreme floor effects). Aber's team engaged in fairly complex factor scoring methods to accommodate these data with the realization that this hinders communication with nontechnical audiences. These measurement challenges underscore the need to work closely with local partners who can work with researchers on the best ways to develop valid, reliable measures and to translate more complex results for nontechnical audiences.

Research Design Pitfalls and Challenges: Conducting Longitudinal Research

Achieving the goal of the most informative and reliable process-oriented, social–ecological research from a developmental psychopathology perspective ultimately requires the conduct of longitudinal research (Cole & Maxwell, 2003). The successful completion of longitudinal research is highly challenging and difficult under the most advantageous of circumstances. The pitfalls and challenges of advancing longi-

tudinal research designs in contexts of armed conflict and political violence involves all the problems one usually faces in this regard in any context. Further, our research teams needed to address various and context-specific dimensions of difficulty and/or challenge related to the particular social–ecological contexts being studied.

Participant attrition

Participant attrition is a concern in any research design in which repeated measures of participants are taken over time. Often, those participants who drop out of a study are different in systematic ways compared to those who remain in the study over time. Attrition affects not only external validity (the longitudinal sample is different from, and thus no longer representative of, the initial sample prior to attrition) but also internal validity (Miller & Hollist, 2007, p. 58)

... by altering the correlations among the variables in the study. This problem occurs in longitudinal research because the subsamples that are dropping out of the study at a higher rate are underrepresented in the longitudinal sample, which may lead to correlations between variables that are different from the true correlations in the original sample.

For example, if participants who were exposed to the highest levels of ethnic–political violence were more likely than those exposed to lower levels to drop out with each successive assessment wave, this would lower the relation between exposure to violence and potential negative outcomes. Because it is easy to see why more at-risk families might drop out of prospective studies over time, teams must build in mechanisms to prevent attrition from biasing their results.

First, careful thought needs to be given to appropriate, but noncoercive, participant incentives. Betancourt's team, for example, used small household gifts as incentives. Huesmann's team faced a 32% attrition rate among the Israeli Jewish sample from Wave 1 to Wave 2, mostly due to refusals: Israeli Jews indicated that they did not feel the monetary reimbursement was sufficient to justify their time, and due to significant exchange rate changes, the amount of money offered to each participant was significantly less than in Wave 1. Therefore, the team increased participant incentives for Israeli Jewish families.

Second, attrition is generally related to initial characteristics of the study participants. In prospective research, as in the studies reported in this Special Section, higher risk, lower income participants more often are lost to attrition. Therefore, anticipating that attrition would be related to measured study variables such as higher exposure to ethnic–political violence, and that attrition would be higher among the Israeli Jewish subsample, Huesmann's team oversampled Israeli youth in high-conflict areas (e.g., youth living in locations such as the northern part of Israel, in settlement locations, near Gaza).

Similarly, Cummings' team experienced a 36% attrition rate over 6 years. To cope with attrition, they added a supplemental set of families living in the original study areas at Wave 3. Cummings et al. (this Special Section) noted, "The

goal of the supplement was to have relatively similar distributions of families across neighborhoods at the outset of the study and to replenish the sample size lost from attrition, thus recruitment of new families oversampled relatively higher-risk neighborhoods characterized by greater attrition.”

Third, all of the studies in this Special Section utilized modern statistical techniques to deal with missing data as a result of attrition (see Little, this issue). As Little has noted,

The modern approaches supersede traditional approaches because, when properly implemented, they can correct for the bias that selective attrition and selective nonresponse introduce. They also have the ability to restore much of the power that is lost when data go missing. One important issue to emphasize here is the need for including the auxiliary variables that capture the reasons for the missing data in a given study.

Threats to external validity

To which populations do the studies' results generalize? As noted above, attrition bias might affect external validity if the longitudinal sample differs in systematic ways compared to the original sample, which had been chosen to be representative of the population. For example, if highly violence-exposed participants are more likely to drop out over time, the study may no longer represent the more diverse population from which the sample was drawn that included youth at all levels of exposure proportional to the population. Some of the solutions described above to reducing attrition will also reduce threats to external validity.

All studies in this Special Section faced the question “To what population do we expect our results to generalize?” Cummings’ team noted, “The original study areas were selected to obtain a representative sample of Catholics and Protestants and variation in levels of sectarian violence.” Huesmann’s team’s goal was to examine the effects of exposure to ethnic–political violence in Palestine and in Israel. Yet violence exposure is much higher among the general population in Palestine than in Israel; in Israel, the major population centers, Tel Aviv and Jerusalem, see less violence compared to other less populated parts of the country. Therefore, the team opted for assessing a representative sample of Palestinians but a sample of Israelis that was skewed toward high-conflict areas. Thus, it is not the case that the Israeli data from their study are representation of the Israeli population.

It may be somewhat more difficult to describe the degree to which the participants in Betancourt’s and Aber’s studies are representative of their populations. Betancourt’s research participants were

... war-affected youth comprising three groups: former child soldiers who had received services through Interim Care Centers (ICCs; $n = 264$), a community sample of youth not served by ICCs ($n = 137$), and a cohort of self-reintegrated former child soldiers recruited at Time 2 ($n = 128$).

Betancourt’s team addressed the concerns about the generalizability of their findings, “We were only able to sample 5 out of

14 Sierra Leonean districts as a result of logistical considerations.” Aber’s team collaborated with 153 schools in four educational subdivisions in Katanga (a province in eastern DRC).

The schools were organized in 40 school clusters of 2 to 6 schools, based on geographical proximity. The evaluation randomly selected a sample of 64 schools out of 153 to participate in data collection. Given unequal cluster sizes, 1 school was selected from clusters that contained 3 schools or fewer, and 2 schools were selected from clusters containing more than 3 schools. All the schools agreed to participate in the evaluation.

The multilevel statistical analysis addressed the question of “whether, and to what degree, school-level variability in students’ perceptions of their school ecologies accounts for the effect of LHRC on school-level student outcomes.” The issue of generalizability in this study relates to whether differences among the many LRHC schools in this study are representative of differences among all potential LRHC schools. Because of the large number of schools in the study, it is likely the results are fairly generalizable. Yet, as described above, an outbreak of rebel activity in 2012 suspended the project in one region, thus removing 13% of the total sample, which limited the generalizability of the results to less conflict-affected provinces of the DRC.

Sample selection. Best cultural practices need to be followed in developing participant samples. For Huesmann’s team, differing approaches were recommended by the Palestinian and Israeli local collaborators as best practice sampling approaches within their regions. For Palestinian families, random door-to-door recruitment, not by telephone (due to fewer land lines), was used; for Israeli families, random telephone and random door-to-door recruitment approaches were used, along with additional nonprobability sampling to achieve an overrepresentation of participants in higher likelihood exposure to ethnic–political conflict/violence areas. Cummings’ team followed a somewhat different approach. Given the close-knit nature of the Northern Ireland communities, community leaders were first identified by the demographer expert in these areas of Belfast, and letters were sent to these community leaders informing them that the team planned to contact families in their communities concerning participation in a research study of political violence, community, and family relationships and child development. Next, letters with the same description of the study were sent to families inviting them to participate. Families were then contacted by phone or engaged door-to-door, consistent with local custom in Belfast. Betancourt’s team collaborated with a major international nongovernmental organization and local community-based organizations in Sierra Leone to recruit from three groups: former child soldiers who had received services through ICCs, a cohort of self-reintegrated former child soldiers, and a community sample of youth not served by ICCs (door-to-door sampling). For the child soldier groups, the nongovernmental organizations and local organi-

zations shared lists; for the community sample, youth were recruited via door-to-door sampling.

Additional Pitfalls and Challenges: Ethical Issues

Study of social–ecological contexts of armed conflict and political violence from a developmental psychopathology perspective, because of the elevated dangers and threats associated with these contexts, also pose elevated pitfalls and challenges for addressing ethical issues and concerns. These matters are enhanced by additional levels of consideration in instances of intervention research.

Many ethical quandaries arise when conducting research with vulnerable populations, and this is amplified in regions of war, especially in low-income economies. Betancourt's project included a mental health services research component, so her team naturally anticipated coming into contact with a significant number of vulnerable youths and caregivers who might require appropriate responses to ensure participant safety, and referrals potentially to a higher level of care. Although participants were often asked to share very sensitive and personal information, they lived in settings where access to mental health services is poor. Betancourt found that referrals for services were on average about 5% of the study population, and involved serious cases that required serious responses. For example, in Sierra Leone, the team uncovered cases of young girls whose "caregivers" were refusing consent. Upon further investigation, these individuals turned out to be unrelated adults in the community who had the girls involved in prostitution and wanted to deny access to group counselling and services being offered in the intervention research as it threatened their exploitative relationship. In such cases, the research team worked closely with the biological parents to ensure that the girls were removed from these exploitative situations and that social workers conducted follow-up case management.

In order to anticipate risk of harm situations and ensure adequate follow-up in Sierra Leone, Betancourt's team developed risk of harm protocols that provided study staff with structured approaches to addressing high-risk situations, while ensuring safety and appropriate referral to available services. According to the protocol, those considered as being a risk to themselves or others or at risk of immediate harm due to abuse and neglect were referred for appropriate mental health care, social work services, and/or other appropriate authorities. All of the research teams in this Special Section developed similar protocols. A crucial part of the protocol is networking in advance with those who can provide no- or low-cost services to families who are identified as needing them. Unfortunately, a research team may not always be located within service-rich environments. In low-resource settings, few communities and even health systems may have awareness of the contributions of mental health to overall child and adolescent health and well-being, and few have the capacity to screen for or treat these conditions. Nonetheless, it is critical to anticipate and plan for referrals to other

forms of supports and to consider other resources that may exist outside of formal services. In Sierra Leone, local village leaders have been important in ensuring the safety and well-being of participants at risk of harm.

More specifically, Betancourt's team spent 7 months preparing mechanisms and identifying formal and nonformal supports to help anticipate and prepare to address risk of harm cases. They phoned and visited with youth and family service providers in each of the six districts where the study would be active. Upon meeting providers at each agency, they inquired about the nature of the services provided and associated costs, fee waivers for impoverished groups, how to make a referral, and whom at the agency to contact. As a result, in the first year following this preparation, the research team was able to respond successfully to nine risk of harm cases from untreated medical needs and hunger to young people who had a history of suicide attempts and active suicidal ideation. A larger group of participants demonstrating less severe risk of harm cases was also provided with information about resources in their community given their specific needs.

One of the research teams in this Special Section, Aber's, conducted an intervention, which requires consideration of additional ethical concerns. Given the cost and time required for such an intervention, is the intervention merely an imposition on an already-burdened system? Is the research project harmfully withholding intervention services from schools and children in the control group? For the latter question, Global TIES and IRC implemented a wait-list control design to meet the challenges associated with providing equitable services to participating schools, which was well received. It is also crucial that self-reflection continuously anticipates the needs of participants from theoretical conception to dissemination of results. The latter stage is particularly challenging as ethical dissemination demands focused delivery at all levels, ranging from peer-reviewed academic journals to ministerial policymakers to the teachers, administrators, and student participants who made the research possible. Social science researchers have historically exploited vulnerable populations for scientific gain, and it is of paramount importance that our work be *for* those who participate and not simply *about* those who participate. Finding ways to make research actionable is among the most substantive ways for research to serve its participants.

Conclusions

In terms of informing social policy and intervention designs for youth in conflict-affected regions around the world, the local commentaries associated with each of the research groups' reports suggest that the stakes could not be higher for implementing rigorous, methodologically sound process-oriented research. For example, from Sierra Leone, Reverend Bangura (this Special Section), suggests,

The findings from this study will also be critical in helping clinics and nongovernmental organizations redesign interventions to more effec-

tively target the coping strategies that have been successful in helping Sierra Leonean youth build resilience despite extreme trauma. More broadly, given the current momentum for a National Mental Health Plan in Sierra Leone, studies like this play an essential role in advocating with stakeholders and policymakers for the importance of community mental health interventions that promote youth psychosocial support.

Similarly, Smith et al. (this Special Section) pointed out the importance of Aber's team's work:

As practitioners, we attempt to balance our humanitarian commitment to act immediately with a commitment to learning through rigorous study to improve our actions in the future. . . . The study is valuable not just for its results but for its call to action for more evidence in conflict-affected settings, where there is disproportionate need and a significant evidence gap. . . . It is critical that studies move beyond a sole focus on outcomes to also examine the pathways and mechanisms that lead to change; implementation factors and costs for achieving outcomes; and contextual factors that affect both implementation and measurement.

In this paper, we have illustrated some of the practical, logistic, and technical challenges in conducting this type of

research. We hope that our experiences in developing fruitful collaborations and finding appropriate funding sources can be used as examples to encourage others to engage in research with youth exposed to ethnic-political violence and armed conflict. Threats to the integrity of research designs are present in almost any context, but our teams found that some threats were quantitatively and qualitatively different. We encountered all kinds of threats to the internal and external validity of our projects, including logistical obstacles in contexts where resources we often take for granted in the West were scarce, outbreaks of war during implementation, which often affect participant attrition, and concerns about the cross-cultural transportability of measurement instruments. Methodological training can certainly reduce these threats in part, but our teams emphasize the need to develop strong, indigenous collaborations to ensure sound culturally acceptable and methodologically sound design and implementation methods, including measurement of key constructs. Only in this way will our research be most ecologically valid when disseminated on a larger scale in the region, which is the main goal of this high-risk, but very rewarding research.

References

- Betancourt, T. S., Simmons, S., Borisova, I., Brewer, S. E., Iweala, U., & de la Soudiere, M. (2008). High hopes, grim reality: Reintegration and the education of former child soldiers in Sierra Leone. *Comparative Education Review*, *52*, 565–587.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*, 558.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.
- Davies, P. T., & Cummings, E. M. (1994). Marital conflict and child adjustment: An emotional security hypothesis. *Psychological Bulletin*, *116*, 387–411.
- Donnelly, J. (2011). How did Sierra Leone provide free health care? *Lancet*, *377*, 1393–1396.
- Dubow, E. F., Huesmann, L. R., & Boxer, P. (2009). A social-cognitive-ecological framework for understanding the impact of exposure to persistent ethnic-political violence on children's psychosocial adjustment. *Clinical Child and Family Psychology Review*, *12*, 113–126.
- Huesmann, L. R. (1988). An information processing model for the development of aggression. *Aggressive Behavior*, *14*, 13–24.
- Huesmann, L. R. (1998). The role of social information processing and cognitive schemas in the acquisition and maintenance of habitual aggressive behavior. In R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for policy* (pp. 73–109). New York: Academic Press.
- Huesmann, L. R., & Eron, L. D. (1984). Cognitive processes and the persistence of aggressive behavior. *Aggressive Behavior*, *10*, 243–251.
- Huesmann, L. R., & Kirwil, L. (2007). Why observing violence increases the risk of violent behavior by the observer. In D. J. Flannery, A. T. Vazsonyi, & I. D. Waldman (Eds.), *The Cambridge handbook of violent behavior and aggression* (pp. 545–570). Cambridge: Cambridge University Press.
- Israel, B. A., Schultz, A. J., Parker, E. A., & Becker, A. B. (1998). Review of community based research: Assessing partnership approaches to improve public health. *Annual Review of Public Health*, *19*, 173–202.
- Krieger, J., Allen, C., Cheadle, A., Ciske, S., Schier, J. K., Senturia, K., & Sullivan, M. (2002). Using community-based participatory research to address social determinants of health: Lessons learned from Seattle Partners for Healthy Communities. *Health Education & Behavior*, *29*, 361–382.
- Lovell, E. L., & Cummings, E. M. (2001). *Conflict, conflict resolution and the children of Northern Ireland: Towards understanding the impact on children and families* (Kroc Institute Occasional Paper Series 21: OP:1). South Bend, IN: University of Notre Dame, Joan B. Kroc Institute for International Peace Studies.
- Miller, R., & Hollist, C. (2007). Attrition bias. In N. Salkind (Ed.), *Encyclopedia of measurement and statistics* (Vol. 1, pp. 57–60). Thousand Oaks, CA: Sage.
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, *24*, 315–340.