

DOCTORAL THESIS

Assessing the impact of a sport for development and peace
intervention on the mental and physical health of adolescents
in Mullaitivu, Sri Lanka

Alexander Hamilton

St Cross College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Medical Sciences Division

Hilary 2016

Assessing the impact of a sport for development and peace intervention on the mental and physical health of adolescents in Mullaitivu, Sri Lanka.

Alexander Hamilton

Submitted for the Degree Doctor of Philosophy (Population Health) – Hilary term 2016

Abstract:

Introduction:

Mental disorders are a crucial contributor to the global disease burden. Exposure to conflict and conflict related stressors are risk factors for poor mental health in children and young adolescents. There is a growing body of evidence for the positive impact of sport and physical activity on the mental well-being of young people in developed, peaceful settings, but it is not clear whether these results can be replicated in post-conflict contexts. Despite this, sport and physical activity are increasingly incorporated into assistance for conflict-affected populations. This thesis assessed the impact of a Sport for Development and Peace (SDP) intervention on the mental and physical health of young adolescents in post-conflict Sri Lanka.

Methods:

I selected methods that balanced pragmatism and rigor and adapted them for use in a low-resource post-conflict setting. I assessed mental health using a locally devised measure, the Sri Lankan Index of Psychosocial Stress – Child Version (SLIPS-C), and physical health using the Multi Stage Fitness Test (MFT), BMI for age (BFA) and height for age (HFA). I tested the reliability and feasibility of the measurement protocols with two repeat measures studies (Kilinochchi study n = 38, Mullaitivu study n = 43). I conducted a cross sectional study (n = 367) in 8 schools which doubled up as the baseline for a pre post with control study (n = 214) in Mullaitivu.

Results:

Successive adaptations improved the intra-tester agreement for the SLIPS-C (Kilinochchi ICC = 0.701, Mullaitivu ICC = 0.878) and all variables showed a good level of reliability (ICC>0.75). There were no significant differences in the main outcome variables between the intervention [V] and control [C] groups at baseline. The target group significantly underperformed when compared to global MFT, BFA and HFA norms. 214 students (47.4% female) were included in the final impact analysis; loss to follow up was 34.2%. Mean SLIPS-C score improved significantly in both groups ([V] = $t(113) 2.46, p = 0.016$); [C] = $t(99) 5.04, p = <0.001$), but there was no significant difference between [V] and [C] groups at follow up. Adjusting for gender, age and ethnicity did not affect the significance of the results. Mean MFT score in the [V] cohort significantly deteriorated ($t(99) = 3.39, p = 0.004$), yet did not change significantly in the [C] group.

Discussion:

Successive rounds of field reliability tests improved the measurement protocols. I was able to draw limited conclusions about the impact of the intervention as it was implemented to a negligible degree, and the study was underpowered due to absenteeism. The results of the evaluation neither support nor disprove the supposition that SDP programmes can have an impact on mental health, as there was a community wide improvement in mental health. In evaluating the impact of sport in post-conflict countries, researchers face an array of theoretical, methodological and epistemological challenges. This thesis is a cautionary tale that highlights the complexity of evaluation in post-conflict zones and the relationships required to sustain it.

Acknowledgements:

There were many contributors to my PhD. The successful project in Sri Lanka was the result of hours of phone calls, conversations and rushed motorbike rides through Mullaitivu. It is difficult to properly detail the support I required at times during this process, and I am truly grateful for all who have blessed me with their assistance and patience.

Firstly, I would like to thank the BHF who have always provided a supportive, friendly environment for work: particularly Aiden and PK, who gave me my first project in Public Health research. I especially would like to thank Charlie Foster and Justin Richards for their continuing, at times perhaps ill-founded, faith in my ability to complete this thesis. It has been a difficult three years for Charlie, and he never failed to provide me with the support I required throughout. His patience, big thinking and selflessness will remain with me. Justin's amazing attention to detail and unending tenacity has taught me a lot. I thank you both for three years of mentorship, and here's to many more of friendship.

There are countless people that have made my PhD excellent fun. I want to thank OURLFC, who collectively subjected me to countless hours of fitness, and even more of hospital bed recovery time. You Gents have provided a much-needed stress vent over the course of my PhD. Stephanie, your love, humour and willingness to have your ear bent about peace was a constant source of support during my PhD, and I am lucky to have had your support.

I want to thank Generations For Peace for their sponsorship and mentoring. My experiences working with GFP have been extremely rewarding, and I want to thank GFP Delegates and Staff who have taught me countless lessons about peacebuilding around the world. Sairah, you have been a model politician and are an exemplary intermediary.

Without the support of the CDO staff and volunteers, this project would have been impossible. Sutharshan, thank you so much for your friendship and support in Mullaitivu. Your cheery chuckle was a constant source of mirth in tough times. Kirusan, without your logistical support I would have been dead in the water. Asela, Nirmani and Indika your support throughout the ethics process was crucial to getting all of this off the ground, so thank you.

I don't know quite how to thank Mum and Dad enough. You have always provided a stable base for me, and supported every twist and turn along the way. Whether consoling me over more doomed conversations with the University of Colombo Ethics Board, or feeding me when I couldn't get off the sofa, you have always been there with your ideas, comments and love. Thank you so much.

Declaration of Authorship

I have read the University regulation and declare that this thesis is my own work. Any material from other sources is clearly indicated.

The idea to assess the impact of a Sport for Development and Peace intervention was my own. I selected the themes for the research and selected the research methods. Generations For Peace supported me for several field trips to observe their programmes in situ. Sairah Yusuf and Dinesha Suppiah of Generations For Peace helped me identify Peacebuilding research priorities.

I spent a total of 8 months in Sri Lanka and received vital logistical support from several sources. The Community Development Organisation, led by Kirusan Ponnuchchamy, provided transport, data collectors and much-needed local knowledge.

I conducted all the statistical analyses myself, with support from colleagues at the Nuffield Department for Population Health and the British Heart Foundation Centre on Population Approaches for Non-Communicable Disease Prevention at the University of Oxford.

This thesis was written under the supervision of Dr Charlie Foster and Dr Justin Richards who have been involved throughout the project. Both Charlie and Justin contributed to initial discussions regarding the development of my research objectives and methods. Both provided feedback on the methods and edited drafts of my written work.

I am very grateful for Generations For Peace, the Nuffield Department of Population Health and St Cross College for funding my studies.

Word count: 47,990

Table of contents

CHAPTER 1.....	1
CHAPTER 2.....	3
2.1 Purpose of chapter	3
2.2 Global mental health	4
2.3 Conflict and mental health.....	6
2.4 Measurement of mental health:.....	7
2.5 Interventions for conflict-affected populations	11
2.6 What is SDP?	13
2.7 SDP Theory	16
2.8 The potential contribution of sport.....	20
2.9 Theoretical framework for thesis.....	22
2.10 Generations For Peace.....	24
CHAPTER 3.....	2
3.1 Purpose of chapter	2
3.2 Selection of case study.....	3
3.3 Local geography and demography	4
3.4 The Mental health landscape.....	8
3.5 Mental health problems: the epidemiological evidence.....	9
3.6 Sport in Mullaitivu.....	12
3.7 Conclusion:	15
CHAPTER 4.....	17
4.1 Purpose of chapter:.....	17
4.2 Research objective:	18
CHAPTER 5.....	22
5.1 Purpose of chapter:.....	22
5.2 Introduction	23
5.3 Methods: Pilot testing.....	27

5.4	Study design:	28
5.5	Outcome variables:	31
5.6	Data collection procedure:	33
5.7	Analysis and presentation:	35
5.8	Assumptions	35
5.9	Mental health:.....	35
5.10	Physical health:.....	36
5.11	Ethical clearance:.....	36
5.12	Results	37
5.13	Summary of studies:.....	37
5.14	Mental health:.....	39
5.15	Descriptive statistics	42
5.16	Intra-tester reliability	43
5.17	Physical health	43
5.18	Discussion	48
CHAPTER 6.....		54
6.1	Purpose of chapter	54
6.2	Introduction	55
6.3	Methods	59
6.4	Results	65
6.5	Sample Distribution.....	65
6.6	Descriptive statistics	68
6.7	Comparison to global norms.....	70
6.8	Comparison with adolescents in Kilinochchi	76
6.9	Association between outcome variables	80
6.10	Discussion	81
CHAPTER 7.....		89
7.1	Purpose of chapter	89
7.2	Introduction	90
7.3	Methods	97
7.4	Results	102
7.5	Sample distribution	103
7.6	Impact of the intervention – within group analysis	105

7.7	Impact of intervention – between group analysis.....	112
7.8	Discussion.....	116
CHAPTER 8.....		130
8.1	Purpose of chapter	130
8.2	Principal findings.....	131
8.3	Stage 3: Innovation testing.....	134
8.4	Implications	136
8.5	Strengths and limitations of the thesis.....	147
8.6	Recommendations: replication, dissemination and sustainability?.....	158
8.7	Conclusions	164
BIBLIOGRAPHY		167
APPENDIX A.....		181
Appendix A.1	Systematic review protocol and precis.....	181
Appendix A.2	Lederach’s Web-building theory.....	185
Appendix A.3	Sugden’s Ripple Effect Model	186
Appendix A.4	Coalter’s logic model for SDP programmes.....	187
Appendix A.5	A typical GFP programme	188
APPENDIX B.....		190
Appendix B.1	Map of Sri Lanka’s Northern Province	190
Appendix B.2	Mullaitivu district with each division highlighted.	191
APPENDIX C.....		192
Appendix C.1	SLIPS-C ORIGINAL VERSION (MH-1).....	192
Appendix C.2	SLIPS-C MH-2 version.....	200
Appendix C.3	SLIPS-GFP Final Version.....	203
Appendix C.4	BMI Measurement Protocol	207
Appendix C.5	MFT measurement protocol	210
Appendix C.6	Information sheet for parents.....	212
Appendix C.7	Consent forms	215
Appendix C.8	Normality tests for MH results in Mullaitivu and Kilinochchi.....	216

Appendix C.9	Ethics Approval for studies	223
APPENDIX D.....		227
Appendix D.1	Normality tests and comparison of sub-groups.....	227
Appendix D.2	Comparison to global normative values.....	232
Appendix D.3	Comparison with Kilinochchi.....	234
Appendix D.4	Association between MH and PH variables.....	237
APPENDIX E		241
Appendix E.1	Intervention outline	Error! Bookmark not defined.
Appendix E.2	Paired t test assumptions.....	241
Appendix E.3	ANOVA for M2 comparisons (between [V] and [C] groups)	245
Appendix E.4	Stratified M1 and M2 scores for all the major outcome variables in the Mullaitivu study	246
Appendix E.5	Intention to treat analysis.....	249
APPENDIX F		252
Appendix F.1	Bauman and Nutbeam (2014) 'Rocket Model'	252
Appendix F.2	Dissemination plan	253

List of Figures

FIGURE 2-1: Human development as a bridge to peace, adapted from Blagojevic 2012	18
FIGURE 2-2: SDP potential contribution to peacebuilding and psychosocial healing	19
FIGURE 2-3: Theoretical framework for the study	22
FIGURE 3-1: A crowd at a local football match descends into a fight.	13
FIGURE 3-2: A summary diagram of assumptions of this chapter.	15
FIGURE 4-1: theoretical framework for the proposed project. Adapted from Blagojevic 2012.	18
FIGURE 5-1: Kilinochchi study summary diagram	29
FIGURE 5-2: Mullaitivu study summary diagram	29
FIGURE 5-3: Study flow diagram for Kilinochchi and Mullaitivu repeat measures studies	37
FIGURE 5-4: Girls running the MFT in Mullaitivu.....	45
FIGURE 5-5: T2 MH assessment in Mullaitivu.	48
FIGURE 6-1: Map of Mullaitivu town and surrounding division with intervention and control schools marked on (scale 1cm = 2km)	61
FIGURE 6-2: Study summary diagram	65
FIGURE 6-3: Graph comparing the HFA Z proportions of 13 year olds in Mullaitivu at M1 compared to global normative values (n = 187)	72
FIGURE 6-4: Graph comparing HFA Z score proportions of 14 year olds in Mullaitivu at M1 and global normative values (n = 167).....	73
FIGURE 6-5: Graph comparing the BFA Z score proportions of 13 year olds in Mullaitivu at M1 with global normative values (n = 187)	73
FIGURE 6-6: Graph comparing the BFA Z proportions of 14 year olds in Mullaitivu at M1 with global normative values (n = 167).....	74
FIGURE 6-7: Graph of Mullaitivu boys MFT score at M1 compared to global averages (n 13 = 110, n 14 = 82)...	75
FIGURE 6-8: Graph of Mullaitivu girls MFT scores at M1 compared with global averages (n 13 = 77, n 14 = 85)	75
FIGURE 6-9: Graph comparing the BFA Z score proportions in 13 year olds Mullaitivu at M1 with Kilinochchi values (n= 631).....	77
FIGURE 6-10: Graph comparing the HFA Z score proportions in 13 years olds in Mullaitivu at M1 with Kilinochchi values (n = 631).....	78
FIGURE 6-11: Scatterplots plotting BMI, BFA and HFA against MFT score in Mullaitivu at M1 (n = 367)	80
FIGURE 6-12: Girls at Thanneeroottu School running on loose ground, barefoot and in full-length clothes. Photo taken by A. Hamilton in June 2015	81
FIGURE 7-1: Study flow diagram	97
FIGURE 7-2: Gantt chart of programme elements showing intended and actual delivery schedule.....	99
FIGURE 7-3: Study flow diagram	103
FIGURE 7-4: Crude mean and 95% CIs SLIPS-GFP scores at M1 and M2 for intervention and control groups (n = 214).....	106
FIGURE 7-5: Crude mean and 95% CIs of INT subscale score for intervention and control groups (n = 214)	106
FIGURE 7-6: Crude mean and 95% CIs of EXT subscale scores in intervention and control groups at M1 and M2	107
FIGURE 7-7: Crude mean and 95% CIs for WTH subscale score in intervention and control groups at M1 and M2	107
FIGURE 7-8: Crude mean and 95% CIs for MFT in intervention and control groups at M1 and M2	108
FIGURE 7-9: Crude mean and 95% CIs for BFA score in intervention and control group at M1 and M2	108
FIGURE 8-1: Diagram indicating the relationships between funders, evaluators and implementing partners. The area highlighted in grey represents the location of a project. Implementing partners and beneficiaries are both acutely affected by the programme location. Drawing on work of Richards and Tol et al. ^[171]	139

List of figures for Appendices

FIGURE A 1: Lederach’s Web-building theory for peacebuilding	185
FIGURE A 2: Sugden’s Ripple Effect Model	186
FIGURE A 3: Coalter’s logic model for SDP programmes.....	187
FIGURE A 4: Map of Northern Province	190
FIGURE A 5: Map of Mullaitivu district. Maritime Pattu is highlighted in red. The black arrow indicates the location of Mullaitivu town.....	191
FIGURE A 6: Height measurement diagram	208
FIGURE A 7: MFT test diagram (Richards 2011)	211
FIGURE A 8: Histograms of SLIPS-C (MH-2) in Kilinochchi, 2014	216
FIGURE A 9: Histograms of SLIPS-GFP score at T1 and T2	218
FIGURE A 10: Weight at T1 and T2 in Kilinochchi, 2014	219
FIGURE A 11: Height at T1 and T2 in Kilinochchi, 2014	219
FIGURE A 12: MFT score at T1 and T2 in Kilinochchi, 2014	219
FIGURE A 13: Histograms of MFT score in Mullaitivu, 2015.....	221
FIGURE A 14: Mean SLIPS-GFP scores stratified by school.....	230
FIGURE A 15: Mean and 95% confidence intervals for MFT score stratified by school	231
FIGURE A 16: Mean and 95% confidence intervals of BFA stratified by school	231
FIGURE A 17: Histogram of boys MFT scores in Mullaitivu at M1.....	232
FIGURE A 18: Histograms of girls MFT scores in Mullaitivu at M1	232
FIGURE A 19: Histograms of SLIPS-GFP score in Mullaitivu at M1 and Kilinochchi	234
FIGURE A 20: Histograms of MFT scores in Mullaitivu at M1 and Kilinochchi	235
FIGURE A 21: Scatterplots of PH outcome variables against SLIPS-GFP	237
FIGURE A 22: Scatterplots of PH outcome variables against INT score.....	238
FIGURE A 23: Scatterplots of PH outcome variables against EXT score in Mullaitivu at M1	238
FIGURE A 24: Scatterplots of PH outcome variables plotted against WTH score in Mullaitivu at M1	239
FIGURE A 25: Scatterplot of residuals against MFT score in Mullaitivu at M1	240
FIGURE A 26: PP plot of the regression-standardised residuals for HFA and MFT score	240
FIGURE A 27: Mean difference between BFA at M1 and M2 between intervention and controls.....	241
FIGURE A 28: Mean difference between MFT at M1 and M2 in intervention and control groups.....	242
FIGURE A 29: Mean difference between SLIPS-GFP score at M1 and M2 in intervention and control groups...	242
FIGURE A 30: Mean difference between Internalising subscale score at M1 and M2 in intervention and control groups.....	242
FIGURE A 31: Mean difference between externalising subscale score at M1 and M2 for intervention and control groups.....	243
FIGURE A 32: Mean difference between Withdrawal subscale scores at M1 and M2 in intervention and control groups.....	243
FIGURE A 33: Histograms of SLIPS-GFP score at M1 for participants with and without follow up data	249
FIGURE A 34: Histogram of MFT scores at M1 for participants with and without follow up data	249
FIGURE A 35: Histograms of BFA score at M1 for participants with and without follow up score	250

List of Tables

TABLE 2-1: Papadopolous' (2007) trauma grid	9
TABLE 2-2: Typical claims of SDP organisations, applying a socio-ecological framework.....	14
TABLE 3-1: Summary of risk factors for poor MH faced by residents of Mullaitivu	9
TABLE 5-1: Summary timeline of adaptations and test-retest analysis of SLIPS-C, Multi-stage Fitness Test (MFT) and Body Composition (BC).....	28
TABLE 5-2: List of abbreviations for study summary diagrams	30
TABLE 5-3: Table of cleaning rules for the SLIPS-C.....	36
TABLE 5-4: Table of abbreviations for summary study diagrams.....	38
TABLE 5-5: Crude mean and 95% confidence intervals for SLIPS-C score (MH-2 version) in Kilinochchi (T1 n = 38, T2 n = 31)	42
TABLE 5-6: Crude mean and 95% confidence intervals for SLIPS-C score (GFP version) in Mullaitivu (T1 n = 41, T2 n = 41).....	42
TABLE 5-7: Test-retest reliability of the adapted SLIPS-C, assessed with the ICC (MH-2 n =31, SLIPS-GFP n = 28)	43
TABLE 5-8: Crude mean and 95% confidence intervals for PH variables in Kilinochchi and Mullaitivu (KT1 n = 38; KT2 n = 31; MT1 n = 43; MT2 = 41)	46
TABLE 5-9: ICC statistic for the PH measures repeatability study in Kilinochchi and Mullaitivu (K n = 31; M n = 27).....	47
TABLE 6-1: Identifiers for map	61
TABLE 6-2: Number of participants in each school as an absolute figure and as percentages of the entire sample, stratified by sex (n = 367)	66
TABLE 6-3: Number of participants in [V] and [C] groups, stratified by gender (n = 367).....	67
TABLE 6-4: Number of participants in the [V] and [C] groups stratified by ethnic group (n = 367)	67
TABLE 6-5: M1 mean and 95% CIs for all major outcome variables (n = 367)	68
TABLE 6-6: Mean, 95% confidence intervals and standard deviation of all outcome variables at M1, stratified by intervention/control status (n = 367)	69
TABLE 6-7: Proportions and 95% confidence intervals of the Mullaitivu sample at M1 in each of the HFA categories according to the WHO global normative values, stratified by age and gender (n = 354)	70
TABLE 6-8: Proportions and 95% confidence intervals of the Mullaitivu sample in each of the BFA categories according to the WHO global normative values, stratified by age and gender (n = 354).....	71
TABLE 6-9: MFT scores in Mullaitivu at M1 compared with global normative values, stratified by gender and age (n = 354).....	74
TABLE 6-10: Proportions and 95% confidence intervals of 13-year-olds in Mullaitivu and Kilinochchi in each of the BFA categories according to the WHO global normative values, stratified gender (n = 631) ..	76
TABLE 6-11: Proportions and 95% confidence intervals of 13-year-olds in Mullaitivu and Kilinochchi in each of the HFA categories according to the WHO global normative values, stratified gender (n = 631) ..	77
TABLE 6-12: Mean and 95% confidence intervals for SLIPS-GFP and MFT in Mullaitivu at M1 and Kilinochchi (n = 828).....	78
TABLE 6-13: Levene's test for HOV and ANOVA results for Mullaitivu and Kilinochchi (df=1).....	79
TABLE 7-1: Legend for study diagram	97
TABLE 7-2: Legend for Gantt chart.....	99
TABLE 7-3: Table showing sample distribution across intervention and control groups (n = 214).....	104
TABLE 7-4: Crude mean and 95% confidence intervals for all outcome variables in the intervention group at M1 and M2 (n = 214)	105
TABLE 7-5: Results of paired t-tests all outcome variables comparing M1 and M2 crude means scores in intervention and control groups (n =214)	109

TABLE 7-6: Table of mean and 95% confidence interval MFT scores at M1 and M2, stratified by intervention status and gender (n = 214).....	110
TABLE 7-7: ANOVA comparing outcome variable scores at M2 for intervention and control groups (n = 214).	111
TABLE 7-8: Crude mean, 95% CIs and ANOVA for mean change between M1 and M2 for [V] and [C] groups (n = 214).....	112
TABLE 7-9: Mean and 95% confidence intervals for mean change between M1 and M2 in only Tamil participants (Intervention Tamils n = 50; Control Tamils n = 95).....	113
TABLE 7-10: ANOVA table adjusting for gender, age, ethnicity (n = 214).	113
TABLE 7-11: Table of results of ANOVA adjusting for baseline, gender, ethnicity and age (n =214).	114
TABLE 7-12: Mean and 95% confidence intervals at M1 for participants with and without M2 data.	115

List of APPENDIX Tables

FIGURE A 1: Lederach’s Web-building theory for peacebuilding	185
FIGURE A 2: Sugden’s Ripple Effect Model	186
FIGURE A 3: Coalter’s logic model for SDP programmes.....	187
FIGURE A 4: Map of Northern Province	190
FIGURE A 5: Map of Mullaitivu district. Maritime Pattu is highlighted in red. The black arrow indicates the location of Mullaitivu town.....	191
FIGURE A 6: Height measurement diagram	208
FIGURE A 7: MFT test diagram (Richards 2011)	211
FIGURE A 8: Histograms of SLIPS-C (MH-2) in Kilinochchi, 2014	216
FIGURE A 9: Histograms of SLIPS-GFP score at T1 and T2	218
FIGURE A 10: Weight at T1 and T2 in Kilinochchi, 2014	219
FIGURE A 11: Height at T1 and T2 in Kilinochchi, 2014	219
FIGURE A 12: MFT score at T1 and T2 in Kilinochchi, 2014	219
FIGURE A 13: Histograms of MFT score in Mullaitivu, 2015.....	221
FIGURE A 14: Mean SLIPS-GFP scores stratified by school.....	230
FIGURE A 15: Mean and 95% confidence intervals for MFT score stratified by school	231
FIGURE A 16: Mean and 95% confidence intervals of BFA stratified by school	231
FIGURE A 17: Histogram of boys MFT scores in Mullaitivu at M1.....	232
FIGURE A 18: Histograms of girls MFT scores in Mullaitivu at M1	232
FIGURE A 19: Histograms of SLIPS-GFP score in Mullaitivu at M1 and Kilinochchi	234
FIGURE A 20: Histograms of MFT scores in Mullaitivu at M1 and Kilinochchi	235
FIGURE A 21: Scatterplots of PH outcome variables against SLIPS-GFP	237
FIGURE A 22: Scatterplots of PH outcome variables against INT score.....	238
FIGURE A 23: Scatterplots of PH outcome variables against EXT score in Mullaitivu at M1.....	238
FIGURE A 24: Scatterplots of PH outcome variables plotted against WTH score in Mullaitivu at M1	239
FIGURE A 25: Scatterplot of residuals against MFT score in Mullaitivu at M1	240
FIGURE A 26: PP plot of the regression-standardised residuals for HFA and MFT score	240
FIGURE A 27: Mean difference between BFA at M1 and M2 between intervention and controls	241
FIGURE A 28: Mean difference between MFT at M1 and M2 in intervention and control groups.....	242
FIGURE A 29: Mean difference between SLIPS-GFP score at M1 and M2 in intervention and control groups...	242
FIGURE A 30: Mean difference between Internalising subscale score at M1 and M2 in intervention and control groups	242
FIGURE A 31: Mean difference between externalising subscale score at M1 and M2 for intervention and control groups	243
FIGURE A 32: Mean difference between Withdrawal subscale scores at M1 and M2 in intervention and control groups	243
FIGURE A 33: Histograms of SLIPS-GFP score at M1 for participants with and without follow up data	249
FIGURE A 34: Histogram of MFT scores at M1 for participants with and without follow up data	249
FIGURE A 35: Histograms of BFA score at M1 for participants with and without follow up score	250

List of acronyms:

ANOVA: Analysis of Variance

ASSIA: Applied Social Sciences Index and Abstracts

BC: Body Composition

BFA: Body Mass index for age

BMI: Body Mass Index

C: Control school

CAMD: Childhood and adolescent mental disorders

CDO: Community Development Organisation

CI: Confidence Interval

CF: Cardiorespiratory Fitness

ERIC: Education Resources Information Centre

EXT: Externalising sub-scale score

DF: Degrees of Freedom

HFA: Height for age

HIC: High-income country

HOV: Homogeneity of Variance

IDP: Internally displaced person

INGO: International non-Governmental Organisation

INT: Internalising sub-scale score

ICC: Inter-class Correlation Coefficient

KS: Kolmogorov-Smirnoff Test

LILACS: Latin American and Caribbean Health Sciences Literature

LMIC: Low- and middle-income country

LTTE: Liberation Tigers of Tamil Eelam

NCD: Non-Communicable Disease

M1: Baseline measurement, March 2015

M2: Follow up measurement, June 2015

MH: Mental Health

MH-1: Sri Lankan Index of Psychosocial Stress – Child Version

MH-2: Sri Lankan Index of Psychosocial Stress – Child Version 1st Adaptation

MH-3: Sri Lankan Index of Psychosocial Stress – GFP adapted version

MFT: Multi-Stage Fitness Test

NGO: Non-Governmental Organisation

OECD: Organisation for Economic Cooperation and Development

OxTREC: University of Oxford Tropical Research Ethics Committee

PA: Physical activity

PI: Principal Investigator

PP: Probability-Probability

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PTSD: Post-traumatic Stress Disorder

RA: Research Assistant

RCT: Randomised controlled trial

RTL: Research Team Leader

SD: Standard Deviation

SDQ: Strengths and Difficulties Questionnaire

SDP: Sport for Development and Peace

SDPIWG: Sport for Development and Peace International Working Group

S.E.K: Standard error of the kurtosis

S.E.S: Standard error of the skew

SfD: Sport for Development

SGBV: Sex and Gender Based Violence

SLIPS-C: Sri Lankan Index of Psychosocial Stress – Child Version

SLIPS-GFP: Sri Lankan Index of Psychosocial Stress – GFP adapted version

SPPY: Sports Program for Youth

SPSS: Statistical Package for the Social Sciences

UN: United Nations

UNDP: United Nations Development Program

UNICEF: United Nations International Children's Emergency Fund

UNOSDP: United Nations Office on Sport for Development and Peace

V: Intervention School

WHO: World Health Organisation

WTH: Withdrawal sub-scale score

CHAPTER 1

Background: A personal view

My undergraduate degree in Human Sciences opened my eyes to some of the broad influences on the human condition: environment, socio-political context and culture. During this time I became familiar with public health in developed contexts, but felt that the course was surely missing a huge portion of the globe, especially in light of what I had learnt studying South and Southern Africa.

As a graduate, I began working with Dr Charlie Foster, Dr Aiden Doherty and Dr Paul Kelly at the British Heart Foundation Health Promotion Research Group. It was at this point I learnt to distinguish between physical activity, sports and exercise, although I'm still somewhat shaky on the detail. While working at the BHF, I met Dr Justin Richards, who had recently returned from his PhD fieldwork in Gulu, Northern Uganda. I was charmed by stories of his fieldwork, and amused by his cynicism. Shortly after meeting Justin, I began working on an application for a scholarship from Generations For Peace, consuming Sport for Development literature clandestinely at my BHF desk.

In March 2012, I travelled with Justin to Gulu for the BHF; this was my first experience of working in a post-conflict zone. I am immensely grateful for this baptism into Public Health in conflict-affected regions. It has changed the way that I look at the world and affected my decision-making since. I enjoyed the practical challenge of working in Gulu; I was intrigued by the conflict legacy and the complex relationships that had unfolded between Non-Governmental Organisations, researchers and funders.

After hearing that my scholarship application was successful, I volunteered briefly for Generations For Peace. I wanted to understand the world of peacebuilding and to gain an insight into this huge, worldwide organisation. I worked briefly in Rwanda and Zimbabwe, both countries with deep psychological scarring from internal violence. In Rwanda, I was haunted by

a constant presence of unease; it seemed to hang over Kigali like a fog, permeating all daily actions and interactions. The idea that the lived environment and daily interactions are crucial components of post-conflict recovery, not just a focus on past traumatic exposures, began to take hold.

In 2013, I set about the task of evaluating the impact of Generations For Peace, an organisation that operates in over 50 countries, with thousands of volunteers and an assortment of project aims, objectives and capacities. I swiftly realised this was not a one-man job and I needed to reassess my objectives. I felt that my own, albeit limited, experience in public health evaluation provided a balance to the reflexive nature of many peacebuilding projects. How to deliver high quality design and measurement in an often transient and changing context? I began to try and marry peace and health, two sides of one coin in the development money bag.

I chose to focus on mental health: the silent burden. Global mental health is a fascinating area of study: an essential resource for individual and societal functioning, too often ignored, hidden and stigmatised. It is something that has affected my own life; a subject that I am passionate about and believe deserves more attention. I believe mental health is strategically situated, straddling the agendas of peacebuilding and public health.

CHAPTER 2

Peace, Conflict and Mental Health

2.1 Purpose of chapter

The purpose of this chapter is to outline the public health issues addressed in this thesis, and to describe and critique the activities of stakeholders in the field. I reviewed the current evidence base, identified gaps in the literature and established a rationale for my research objectives and fieldwork.

2.2 Global mental health

Mental health (MH) is a crucial component of health and well-being [1]. The World Health Organisation (WHO) define MH (or psychological well-being) as *“a state of well-being in which every individual realises his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community”* [2].

MH is not just the absence of mental disorder, but also the presence of positive mental states, and contains elements of functionality and positive feeling.

Mental disorders are a crucial contributor to disability and mortality, as well as to accidental and non-accidental injuries globally. As a result, mental illness is a major contributor to the global burden of disease [3, 4]. The common MH disorders account for 14% of the global burden of disease, about 7.4% of disability-adjusted life years worldwide and roughly a quarter of all health lost due to disability [1, 3, 5]. As well as causing immense human suffering, mental illness has economic and societal costs [6, 7].

Mental disorders affect between 10-20% of children worldwide [8, 9]. Common mental disorders in young people are: depression, anxiety-related disorders, post-traumatic stress disorder (PTSD) and emotional and behavioural disorders [10]. Results from longitudinal, clinical and population-based studies indicate that emotional and conduct disorders can impair development for age, and can influence subsequent quality of life for those affected and their families [11, 12]. Emotional and conduct disorders have been linked to:

- Adult psychiatric problems;
- Poor physical health (PH);
- Criminal convictions;
- Adverse behavioural outcomes and social impairment;
- Suicidal and high-risk behaviour;
- Low academic achievement and higher rates of school drop-out [8, 13-17];

The societal costs of adolescent mental disorders are less well documented than that of adult disorders [18]. Children with disorders have higher rates of healthcare utilisation and generate higher education, social service, criminal justice and informal care costs than their unaffected contemporaries [12, 19]. While it is clear that child and adolescent mental disorders (CAMD) impose a high individual burden, importantly, they are also risk factors for subsequent mental illness and many persist into adulthood [20].

Of the 2.2 billion adolescents and children worldwide, almost 90% of these live in low- and middle-income (LMIC) settings, where they constitute 50% of the population [21]. A systematic review of non-referral studies in LMICs revealed a prevalence of about 10-20% of CAMD, which is consistent with studies in high-income countries (HICs) [9, 22]. The reported prevalence of CAMD varies widely between countries (1.81% - 39.4%). Estimating the global burden of CAMD is complex and there are a number of factors that limit accurate analyses [22-24]. There is inadequate provision for assessing MH problems worldwide, on account of low awareness and priority of MH in many countries [9, 25, 26]. People suffering from mental illness are often socially excluded and are victims of neglect and discrimination [27]. Additionally, there is variable cross-cultural definition, expression and diagnoses of mental disorders. These factors are particularly pertinent in post-conflict settings, where epidemiological information is not always available, baseline information is lacking and data recording is disrupted [5].

Prevalence and social distribution of MH disorders has been well documented in HICs. There is a research gap regarding the magnitude of the problem, and appropriate strategies, policies and programmes to prevent MH disorders in LMIC and post-conflict settings [28]. CAMD is a crucial contributor to Non-Communicable Disease (NCD) related disability and death but, despite its prevalence, continues to draw insufficient funds, especially in areas of poverty and continuing violence [29].

2.3 Conflict and mental health

The impacts of violence are enduring and wide-ranging, both direct and indirect. Exposure to conflict and conflict-related stressors is a risk factor for elevated levels of poor MH in children and adolescents [30, 31].

Conflict not only exposes people to violent and traumatic events which are known to increase MH problems, but can also radically change the social and economic environment which can catalyse the emergence of psychiatric disorders [28, 32-34]. A lack of social support, prolonged displacement, breakdown of community structures and deprivation of basic needs are common in post-conflict zones [35]. Erosion of social support puts those already suffering from MH problems at increased risk [36]. These factors contribute to poor levels of MH among conflict-affected populations, and delay the peacebuilding process [37, 38]. Typically, affected populations suffer from elevated levels of PTSD, anxiety-type disorders, and increased substance and alcohol abuse [34, 37, 39]. Studies indicate levels of mental distress as high as 50% in post-conflict zones [34].

People affected by conflict and natural disasters are a unique population; the post-conflict context has complex disease distribution with multiple, crosscutting health challenges including many health-related barriers to development [40]. Post conflict states are increasingly affected by NCDs, including towering rates of mental illness, and some face the dual burden of disease [41].

There is no universally accepted definition of a post-conflict setting and in reality each context is unique. The term 'post-conflict' describes many countries in different states of development and unrest [42]. Despite differences, many countries face similar crosscutting social, economic and political challenges, a broad definition is:

“a situation in which open warfare has come to an end. Such situations remain tense for many years or decades and can easily relapse into large-scale conflict [43].”

It can be a period of rapid change and often characterised by a resettling of the civilian population, disarmament and reconstruction of livelihoods. As a result, research in the post-conflict context is beset with difficulties [44].

2.4 Measurement of mental health:

MH is a complex and multifaceted construct that has many potential measurable elements, both positive and negative. Measurement of MH suffers as a result of this complexity and lack of a mutual definition. MH can be examined as a whole, as individual components that contribute to a whole, or through a combination of a number of individual constructs [45]. Each approach has merits and shortcomings; for example, collating constructs into single measure provides a convenient, broad assessment that might account for many of the important influences on MH. However such an approach may be imprecise and, particularly in a diagnostic setting, limit the capacity to distinguish meaningfully between cases. Conversely, focusing on a specific MH construct will provide a more accurate assessment of the construct of interest, but may be at risk of being reductionist and fail to identify important, related aspects of MH. Likewise, only measuring mental illness and omitting mental well-being may miss crucial aspects of MH. In reality, different measurement approaches will be appropriate for different lines of enquiry.

The presence of certain patterns or clusters of thoughts, feelings and behaviours (referred to as symptoms) is used to diagnose mental disorders, or assess positive MH constructs. Mental illness is typically when a group of symptoms are present for a specified period of time and impact mood or capacity to function [46]. Symptoms are typically assessed by a medical professional through semi-structured interview or discussion, or through self-completed diagnostic screening tools [47]. Clinician administered semi-structured interview is the most commonly used 'gold standard' approach in psychiatric research, typically applying the Diagnostic and Statistical Manual of Mental Disorders (DSM) diagnostic criteria [48]. However, the validity of the DSM criteria has been questioned, as it is based on a consensus about clusters of 'clinical' symptoms, rather than any truly objective markers. The line between abnormality and disorder is socially determined, and not universally agreed upon. The search for psychiatric gold-standard measures is on-going [47].

MH assessment in post-conflict zones is beset with methodological and theoretical complications. The medical (or trauma-focused) and psychosocial approaches are the two core

paradigms for the assessment of and assistance for conflict-affected people [49]. For trauma-focused advocates, exposure to the violence and destruction of war is the critical MH determining factor. Conversely, psychosocial advocates focus on the material and social conditions that are depleted as a result of war, rather than on the direct effect of conflict on MH [50]. The differing theoretical approaches have implications for proposed measurement, diagnosis and treatment methods.

The majority of epidemiological studies in post-conflict zones have focused on the prevalence rates of depression, PTSD and anxiety disorders [51]. People exposed to traumatic events and conflict display an array of psychological responses, highlighted by Papadopolous' (2007) 'trauma grid' (TABLE 2.1) [52]. One of the key issues in post-conflict MH is that most of the population will have been exposed to multiple trauma events, and as a result will show acute stress reactions [36]. Acute stress reactions are common in the aftermath of violence, but these are not necessarily indicative of PTSD. Cultural variation in the expression of psychological distress and the use of Western screening tools may lead to incorrect identification of symptoms, and subsequent treatment allocation [36, 49, 53].

As in the trauma grid many individuals will show resilience, and some post-traumatic growth, when exposed to stressors. The trauma grid serves as a reminder for practitioners that individual, negative responses to trauma are only one of many possible outcomes and a framework to discuss broader issues surrounding trauma-related distress. As trauma occurs on multiple levels, advocates of the psychosocial paradigm claim that the over-riding focus on PTSD assessment may fail to identify suffering and potential intervention areas. Only assessing MH outcomes at the individual level fails to reflect collective and community trauma, which may be more salient in different cultures [54].

TABLE 2-1: Papadopolous' (2007) trauma grid

Level	Negative effects			Neutral effects	Positive effects
	Injury			Resilience	Adversity-activated development
	Psychiatric disorders	Distressing psychological reaction	Ordinary Human Suffering		
Individual					
Family					
Community					
Society/culture					

Practical constraints affect the use of optimum methods in post-conflict contexts. Post-conflict MH assessment methods are limited by an over reliance on symptom checklists that have not been validated in the research context; these tend to overestimate disorder prevalence and obscure the ability to distinguish between normal psychological stress and severe psychiatric disorder [9, 55]. The broad range of responses to trauma and non-optimal measurement methods contribute to the complications of assessing MH in post-conflict zones.

Despite methodological and theoretical complications there is extensive evidence that children exposed to traumatic events are more likely to exhibit signs of emotional distress, mental and behavioural disorders [56]. Best practice advocates the use of MH assessment instruments that have been adapted to account for local culture, to ensure validity of diagnoses [57]. The use of mixed methods to triangulate assessments and, particularly when working with children and adolescents, the use of third party reporting can improve validity [44].

MH and psychosocial problems in post-conflict settings are highly interrelated, and difficult to disentangle. This thesis adopts a psychosocial measurement approach for several reasons. The psychosocial agenda resonates better with the aims of peacebuilding, as trauma-centred approaches tend to focus on more specific groups screened for mental illness. Secondly, a trauma centred approach may fail to identify many causes and elements of suffering in the target

population. The target community in Sri Lanka have been described as 'collectivist', and psychosocial measures of MH will better reflect their experience [51]. A strictly MH approach fails to adequately take into account the inter-relationship between the individual and their wider social environment.

Coordinating an appropriate response to the challenges that mental disorders present requires an understanding of determinants, resources and effective interventions that influence MH.

Despite burgeoning research efforts, convincing evidence regarding the impact of interventions to address conflict-related mental disorders remains noticeably absent [30, 39, 58, 59].

2.5 Interventions for conflict-affected populations

There is now a wide range of actors and approaches engaged in the assistance of conflict-affected populations. Generally, psychosocial support methods aim to develop resilience in communities and individuals through community-based social programmes [57]. The psychosocial paradigm encourages an appreciation of the dynamic relationship between MH and the socio-ecological environment [50]. In contrast to this, trauma-focused practitioners are characterised as using drug treatments or intensive therapy methods to treat diagnosed psychological problems.

There is now a broad array of psychosocial programmes delivered in post-conflict settings that encompass many different intervention modalities. Interventions in young people typically aim to bolster protective support in families, schools, communities and societal levels [58].

Peacebuilding is another, related field of post-conflict assistance. Similar to psychosocial interventions, peacebuilding has a social dimension at its core [57, 60, 61]. Post-conflict peacebuilding is a complex and multi-faceted challenge that contemporary nations face as they attempt to rebuild society disrupted by conflict, on human and structural levels.

Peacebuilding can be loosely defined as intentional actions that contribute to the establishment of 'positive' peace, tantamount to social justice and equality, not just the absence of violence [62]. Definitions of peacebuilding vary among multilateral organisations, e.g. the UN, Non-Governmental Organisations (NGOs), funders and policy makers. It is a complex and nebulous entity, and the multiplicity and ambiguity that surrounds definitions of peacebuilding represents a challenge for coordination between stakeholders [61, 63].

However, the Utstein Pallette defines four key components of peacebuilding:

- The provision of security (physical and human);
- The establishment for the socio-economic foundations for long term, sustainable peace (social and economic);
- The establishment of the political framework for sustainable peace (political);

- Reconciliation and healing the wounds of war (social and psychological) [64].

Reconciliation and healing the wounds of war, a cornerstone of peacebuilding, refers to 'healing' on both a social and psychological level: addressing the burden of mental illness and promoting new, positive relationships between previously warring parties. There are clear similarities between psychosocial healing and peacebuilding, the amelioration of mental distress through the development of the community and broader social context.

In reality, the psychosocial/trauma dichotomy is false; the population is affected to various degrees, and practitioners support the provision of multi-level care for conflict-affected populations, based on need [58, 65]. Similarly, singular peacebuilding actions are unlikely to be effective unless they are strategically implemented and join up with multi-sectorial stakeholders [63].

Several publications review the field of psychosocial assistance for young people in conflict-affected settings [66-69]. Jordans et al. (2016) conducted a systematic review of the impact of MH and psychosocial support for children affected by armed conflict [70]. Some intervention modalities reported moderate positive effect sizes, albeit in specific sub-groups. Reviews conducted by Betancourt et al. (2013) and Jordans et al. (2009) reach similar moderate positive effect conclusions, but cite a shortage of evidence, small effect sizes and a lack of treatment modality research [58, 71].

Structured activities and sporting elements are often included in psychosocial programmes, but the sport or physical activity (PA) elements are often improperly described. Jordans et al. (2016) call for further exploration of PA interventions in this context, as reviews from HICs report an association between PA and self-esteem, social interaction and depression/anxiety [72, 73]. The use of sport for psychosocial healing is a subcategory of a wider movement, known as Sport for Development and Peace (SDP). The use of sport to build peace and to achieve development objectives is now widespread.

2.6 What is SDP?

SDP is the use of sport, PA and play as a mechanism to achieve specific social development or peace objectives. The use of SDP is widespread and involves many multilateral organisations (e.g. United Nations (UN) agencies) international Non-Governmental Organisations (iNGOs) (e.g. Right to Play) sporting bodies (e.g. FIFA), government bodies (e.g. UK Sport, Australian Sports Commission), private companies and academic institutions [74].

SDP covers a broad array of actors implementing a spectrum of programme models, who claim manifold crosscutting positive outcomes. The development and history of SDP has been documented [75]. Growth in the sector has remained unchecked over the last half decade, despite the sensitivity of the context and nature of the work. There are 585 organisations and 213 projects registered on the International Platform for Development and Peace¹, a number that has almost tripled since 2008 [76]. As many SDP organisations either do not register, or are not English speaking, this number is likely to be an underrepresentation of SDPs current status [77]. Sport is often heralded as a panacea for the issues faced by youth in post-conflict zones [78]. In reality, it is a social construct, which in some settings can contribute to well-being and harmony, whilst in others it feeds hostilities and fuels violence [79]. The role of sport should not be essentialised in modern society [80].

Sport is often portrayed as a tool that can contribute to health and social development; provided it is used responsibly, in culturally appropriate ways, with local ownership and based on thorough needs assessment [81]. In fact, it may be that the most important aspect of SDP programmes is the quality of the coach or implementer, who if positive, skilled and understanding can facilitate the transfer of skills and values learned in programmes to real-world behavioural changes [82-84].

Diverse practices are incorporated under the SDP headline, and SDP agencies and practitioners make a wide-ranging array of claims [75]. According to the Sport for Development and Peace International Working Group (SDPIWG), sport can contribute to all of the Millennium

¹ Search conducted on 23rd of August 2014.

Development Goals [85]. TABLE 2.2 outlines some of the claims made by SDP organisations applying a socio-ecological framework. Despite the claims, there is little evidence in peer-reviewed literature, particularly in the sport for peace sector [81, 86, 87].

TABLE 2-2: Typical claims of SDP organisations, applying a socio-ecological framework.

Socio-ecological level of outcome	Claims
Wider social	Socio-economic foundations for peace Social inclusion Reduced inequality
Community	Build civil society, Capacity building, Reduce crime, Empowerment Gender norms, Cross-cultural relationships, Engagement, Knowledge base, Develop social capital
Inter-personal	Develop trust relationships, Normalize life and environment, Provide space for dialogue, Develop 'citizenship', Develop networks, Develop social capital
Intra-personal	MH, PH, Life skills, Conflict Management, Education, Control, Tolerance, Respect, Attitude change, Values, Respect, Empathy

Current evidence for many of these outcomes is typically grey literature describing programme reach, or anecdotal success stories [88, 89]. Research is carried out with the assumption that sport will have positive health benefits, to prove concepts, rather than objectively [86]. The international enthusiasm, and UN championing, has blurred the gap between assumed benefits and the evidence based knowledge of sports participation. Put simply, the claims over-reach the current evidence base [88].

Of particular interest are the MH related claims of SDP organisations, the majority of which remain without strong evidence. A review coordinated by the SDPIWG assessed the evidence base for SDP programmes on the MH of adolescents [90]. The review leans heavily on studies

conducted in high-income, peaceful settings and it is not clear whether the results of these programmes are reproducible in post-conflict countries.

I conducted a systematic review of the MH impacts of programmes for adolescents in post-conflict settings that had clearly defined sport or PA elements to address this gap in the literature (systematic review protocol and precis: APPENDIX A.1) [91]. My review built on previous reviews of evidence conducted by Richards et al. (2011), Cronin (2011) and Zakus (2007) [90, 92, 93].

The findings of my review contradicted the broad claims of the positive effect of sport and PA on MH. Two randomised controlled trials (RCTs) conducted in northern Uganda reported no impact on anxiety and depression like symptoms in war-affected adolescents [94, 95]. A 'plus sport' intervention, a competitive sport for development league, had an adverse effect on the MH of participating boys and a negligible impact on participating girls [93]. While a 'sport plus' programme, a creative play intervention for depressed adolescents, had no impact on the MH of adolescents aged 14-17 [94]. The third publication, an evaluation of a Dance Movement Therapy study, described significant reductions in all MH outcomes measured, but the quality of the reporting is too poor to critically assess these claims [96]. This particular study illustrates a weakness in the SDP field at large, in which the majority of evaluations are of poor design, and typically implement pre-post methodologies without control populations [76, 97].

Several authors comment on the shortcomings of anecdotal evidence and the need for better monitoring and evaluation in SDP [81, 87, 98, 99]. My review identified several evaluations conducted by SDP agencies of programmes that claim to improve aspects of MH in post-conflict contexts [91]. Despite this, few organisations used measures that have been validated and are reliability tested in the target population. The majority of identified articles were descriptive and talk broadly about the potential of sport, or identified case studies, without detailing any specific MH outcomes. Studies that did cover MH lack control groups, or relied on the observations of coaches and facilitators to determine significant change [82, 100]. Additionally, studies were typically short term, despite best practice guidance calling for longer studies to allow for intervention maintenance effects to be assessed.

2.7 SDP Theory

The shortage of evidence supporting the use of SDP to achieve MH and peace objectives is coupled with a lack of theory and mechanistic research; how do proposed actions link up with the intended outcomes? Theories of change and logic models are often unarticulated and untested, and programmes have unclear or non-specific goals [75, 101, 102].

A paucity of theory might limit the potential for rigorous evaluation and effective programme design in the SDP field. Practitioners are characterised as outlining broad, unsupported benefits for participation in programmes in order to attract investment [77]. These claims are often moulded to the expectations of their funding bodies, rather than based on a proper needs assessment [75]. Too often, how individual level changes link up with broader peace objectives is largely uncovered and unspecified [103]. There is a need to understand the conditions, structures and process in which sport can be used to build peace, improve MH and contribute to social change.

Theoretical and practical complications mean that peacebuilding programmes often lack rigorous process or impact evaluation, programme theory and modelling. Fast and Neufeldt (2005) identify key areas of difficulty in peacebuilding evaluation:

1. The outcomes of peacebuilding programmes are long term – complicating short-term analyses.
2. Stakeholders have different expectations regarding outcomes, measurement and timescales.
3. Peacebuilding lacks a universally accepted definition.
4. The outcomes of programmes are not necessarily tangible and quantifiable: many interventions aim to change attitudes, opinions and beliefs.

Mixed-methods and participatory approaches go some way to bridging differences in donor-recipient expectations and accounting for the intangible nature of peacebuilding outcomes [104]. Likewise, participatory evaluations can contribute to community capacity and provide outcomes

that are relevant to local stakeholders, rather than importing foreign notions of peace and associated indicators [105]. The challenge of developing reliable sensitive indicators on the pathway to peace and the subsequent use of these in planning and implementation is critical to improving practice. MH research has successfully operationalised 'intangible' outcomes, such as attitudes and feelings, which can assist with peacebuilding evaluation.

Current theoretical models, such as Lederach's Web Theory and Sugden's Ripple Effect Model are rudimentary, lack explanatory power and do not address the challenges highlighted by Fast and Neufeldt (2005) (the models are attached in APPENDIX A.2 and A.3) [106]. The models are highly abstract, and have limited potential for the design and implementation of SDP evaluations, although they highlight important theoretical considerations in the design of programmes.

Coalter (2008), Lyras (2011) and Richards (2011) outline taxonomic models for the design and evaluation of SDP interventions. Coalter's (2008) model is linear, unidirectional and, in parts, non sequitur (APPENDIX A.4)^[87, 107]. It poses direct links between 'sporting outcomes' and 'intermediate outcomes', and implies directional causality without any explanation of how outcomes are achieved. It suggests that individual level changes will facilitate broader societal change, which are ubiquitous and unproven assumptions in the use of sport to promote social development [88]. Lyras' (2011) model, while comprehensive, fails to take into account the mutually implicative nature of the relationships and outcomes that it details [87].

Richards' (2011) outline a comprehensive program planning theory, which provides a clear analysis of the actions of SDP stakeholders and their intended programme outcomes. It highlights the complex multi-directional interaction between outputs and impacts well, and provides a good basis on which to analyse the role of different stakeholders in planning and implementing an SDP intervention. As a taxonomic model it is broad, and does not detail how programme elements might lead to specific SDP outcomes of interest. In accounting for both broad scale change, and micro level change, it fails to identify any theoretical information on the impact of sport on intended SDP outcomes.

Blagojevic (2012) outlines a theoretical model for the contribution of sport to peace, via human development [60]. Any sustainable advance in human development is dependent on the other core areas of peacebuilding (political, economic and physical spheres) supporting this change, no impact occurs in isolation. According to Blagojevic, individual level changes are unlikely to be sustained unless broader social and economic conditions support new behaviours and ideas. If we accept that social development is a core component of peacebuilding, then Blagojevic (2012) states that sport can have two fundamental levels of impact:

- **Empowering individuals:** by promoting physical, emotional and mental well-being
- **Rebuilding communities:** through reconciliation and transformation of relationships

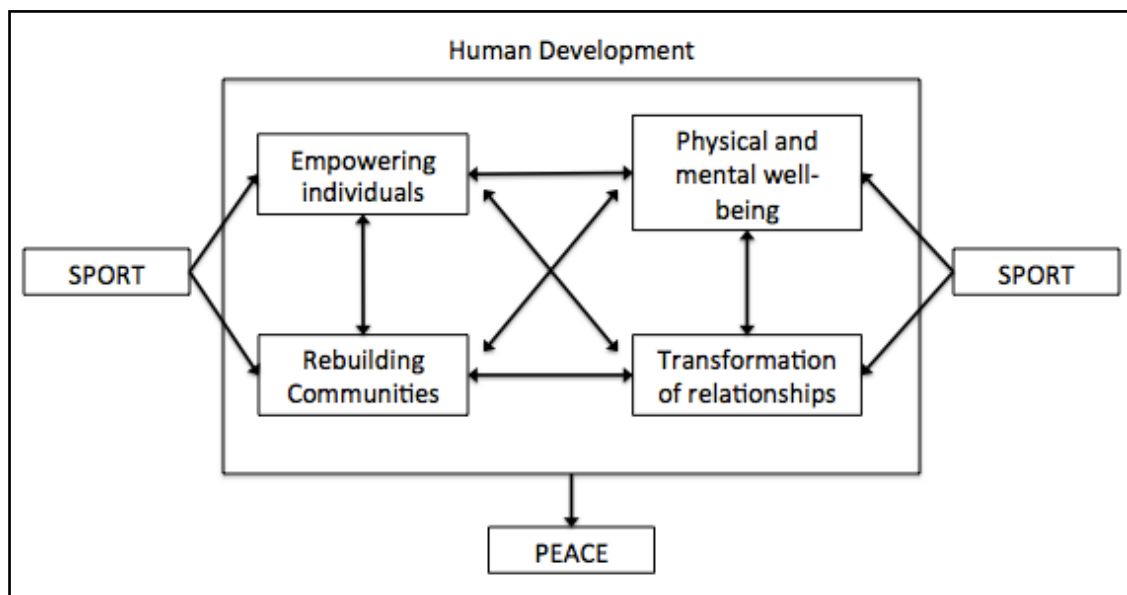


FIGURE 2-1: Human development as a bridge to peace, adapted from Blagojevic 2012

Many of the pathways in FIGURE 2.1 outlined in the theory are poorly evidenced, or are unclear and uncertain. However, according to the theory, individual, community and social level interventions in post-conflict settings have the potential to improve MH and well-being, and in this way contribute to human development, a crucial aspect of peacebuilding. I argue that the objectives of peacebuilding and psychosocial healing are interconnected (FIGURE 2.2), and that evidence from evaluations of SDP programmes can contribute to both fields of knowledge. In fact, SDP may be a broad mechanism by which psychosocial and peacebuilding interventionists can achieve common goals.

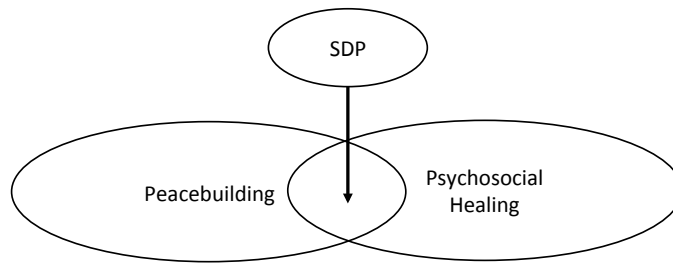


FIGURE 2-2: SDP potential contribution to peacebuilding and psychosocial healing

The complex nature of peacebuilding evaluation, and the fact that many SDP organisations fail to properly elucidate their theory of change, inhibits the development of appropriate, valid and reliable indicators. MH, as a component of human development, may operate as a midway indicator in the peacebuilding process. Given the international championing and growing investment in SDP, it is necessary to critically assess the potential role of sport and sport for peace activities in the amelioration of mental illness and promotion of positive aspects of MH.

2.8 The potential contribution of sport

Several authors state that sport and PA-based programmes have the potential to address MH in post-conflict contexts [49, 70, 90, 108, 109]. SDP interventions are heterogeneous in their methodological approach and programme elements, but there are core features that might reinforce protective factors, or reduce risk factors, for child and adolescent MH.

PA alone may play a role in addressing MH issues. PA impacts MH through a variety of mechanisms including psychological, psychosocial, and physiological [110]. There is a growing body of evidence for various positive impacts of sport and PA on the physical, mental and social wellbeing of young people, including those 'at-risk' in developed, peaceful settings [72, 73, 111-114]. Meta-analytic evidence from HICs suggests a positive impact of interventions on self-esteem in young people [112]. Ahn and Fedewa (2011) conducted a meta-analysis of the relationship between PA and positive and negative aspects of MH. RCT evidence indicated that PA significantly reduced depression, anxiety and increased self-esteem in young people [115].

Good PH is crucial for promoting MH in LMIC contexts [25, 116]. Physical fitness (PF) is a good predictor of health-related outcomes in adolescents, despite its genetic component [117].

Increased cardiorespiratory fitness (CF) is an outcome of increased PA and may result in improved overall PH [118]. RCT and quasi-experimental evidence supports the proposition that SDP programmes may directly improve CF in post-conflict and LMIC settings [95, 119].

CF may be associated with several MH outcomes. There is growing evidence from developed settings of the protective role of CF against depression and stress responses [120-122]. The relationship between CF and MH remains unclear in developing settings; but studies in HIC suggest CF is positively correlated with improved MH [123]. I anticipate that increased PA as a result of SDP programmes, provided it is of sufficient intensity and duration, can contribute to improved PF and MH.

SDP programmes can contribute to MH in more ways than just via the improvement of PH. Team sports are associated with improved health outcomes compared to individual exercise in peaceful settings [111]. Assumed outcomes of sports participation are:

- Increased skills and mastery ^[124];
- Increased social interaction ^[125];
- Increased self-esteem, efficacy and confidence ^[126].

All of these factors can potentially contribute to improve positive elements of MH. In fact, programme elements beyond the typical 'sport' component may be the crucial components of interventions ^[81, 84]. The establishment of good role models, positive relationships and skills development are crucial to the promotion of MH in LMIC and post-conflict contexts ^[20, 23-25]. The Psychosocial Working Group state that developing skills and enhancing effective social engagement is crucial in promoting MH in young people ^[50].

The individual, their behaviour and the environment are in 'reciprocal determinism' ^[127]; they are constantly interacting and influencing one another ^[128]. Appreciating this interaction can contribute to how we understand the potential impact of SDP interventions on MH. MH supportive behaviours, or well-being outcomes, might result from the interaction between individual and environmental determinants. Providing a safe, fun space in which to play and be active may well provide a good foundation for the promotion of MH in post-conflict contexts.

2.9 Theoretical framework for thesis

There is considerable crossover between the practices of psychosocial healing and peacebuilding. Human development, 'healing the wounds of war' and relationship building are key components of the peacebuilding process (section 2.4). SDP organisations claim to stimulate MH benefits alongside a myriad of new, positive relationships across multiple boundaries, social inclusion and psychosocial healing, among other outcomes. There is a shortage of evidence that SDP organisations are actually achieving these outcomes.

FIGURE 2.3 outlines the theoretical framework for this study.

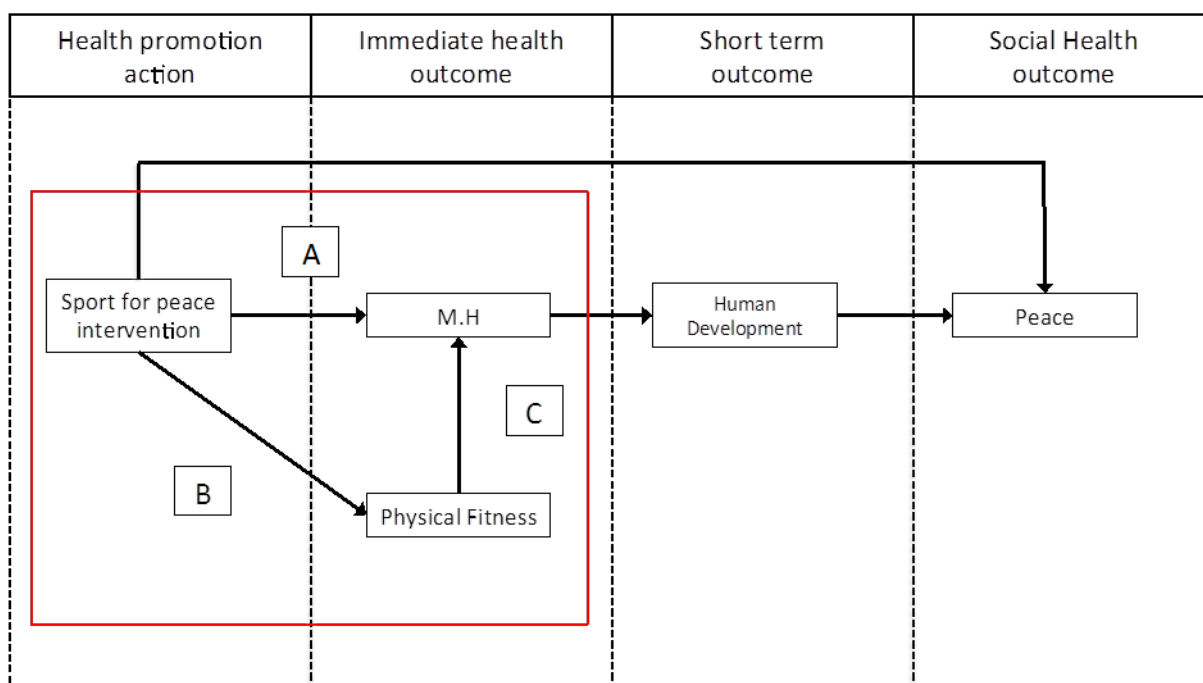


FIGURE 2-3: Theoretical framework for the study

A simple, clear theoretical framework assists with systematic data collection and a coherent evaluation [129]. This study will assess the assumptions identified in FIGURE 2.3. Firstly, that SDP programmes can contribute to MH improvement [assumption A]. Secondly, SDP interventions can improve CF as a result of increased PA [assumption B]. Lastly, that increased CF may partially explain variation in MH [assumption C]. It is likely that complex social and environmental factors contribute to peacebuilding and MH in SDP programmes, however it is

beyond the scope of this project to assess these. The variables highlighted in the red box are quantifiable, measureable and theoretically associated with one another. The identification of intermediate variables on the path to peace can help with the evaluation of SDP programmes and inform future intervention design. There are potential feedback loops between peace, human development and MH improvements. These are not marked on the simplified theoretical framework, but are certainly important potential relationships.

The theoretical framework, while simple, addresses some of the challenges identified by Fast and Neufeldt (2005) ^[130]. In this thesis, I will contribute to the evidence for sport for peace programmes by conducting a case study assessing the impact of a Generations For Peace (GFP) programme on the MH and PH of adolescents in northern Sri Lanka.

2.10 Generations For Peace – The intervention

I chose to evaluate the impact of GFP, an SDP NGO based in Amman, Jordan. GFP is a leading global non-profit peacebuilding organisation founded by HRH Feisal Al Hussein and Sarah Kabbani in 2007. GFP trains young people to promote active tolerance and responsible citizenship in communities experiencing conflict. GFP remains a sport-oriented peacebuilding organisation, but promotes the use of locally relevant peacebuilding methods. GFP has trained over 8,800 volunteers from 50 different countries, and claims to impact the lives of more than 227,000 children. The basic premise is to identify and train youth leaders to develop, advocate and use sport to promote peace around the world.

GFP recruit youth leaders in target countries as volunteers and spokespeople for the NGO. Volunteers are encouraged to apply to attend GFP Basic Training in Amman, Jordan. This is an intensive 10-day course in which participants are trained in the GFP Curriculum using sport to build sustainable peace.

This research concerns GFP Sports for Peace Programmes for Youth (SPPY). 'SPPY' encompasses a broad array of programmes with different target groups, methods and objectives. SPPY involve the use of sports based games to develop certain skills, attributes and values that are identified as crucial to the development of sustainable peace in the target community. Local partners are expected to select and adapt games and methods that suit their context. Several programmes involve re-establishing connections between divided groups and provide the basis for new, positive relationships between opposing communities. I selected SPPY because they are the core of GFP activities. A generic GFP SPPY programme is outlined in APPENDIX A.5.

SUMMARY TABLE

- CAMH is a crucial contributor to the global burden of disease;
- Exposure to conflict and disaster-related stressors are risk factors for poor MH in young people;
- Despite the prevalence and impact of mental disorders in young people, they continues to draw insufficient funding, including in areas affected by protracted violence and disasters;
- Peacebuilding is a complex task, that suffers from a paucity of well-conducted evaluation;
- There is an increasing evidence base regarding the positive effects of sport and PA on young people in developed, peaceful settings, although the transferability of these outcomes to a post-conflict context is unknown;
- Despite this, physical activity and exercise-based interventions are incorporated into assistance for conflict-affected populations;
- A lack of viable programme theory exacerbates the lack of good quality evidence for SDP;
- I have outlined a simple theoretical framework with specific, measurable elements that links the agendas of peacebuilding, SDP and public health.

CHAPTER 3

Mullaitivu district

3.1 Purpose of chapter

The purpose of this chapter is to outline my rationale for choosing Northern Province in Sri Lanka as the case study site. The chapter also provides a brief history of the region and some relevant demographic and geographical information. More specifically, it describes the current living conditions and how these might influence MH. It outlines the MH services available in the region and the challenges faced by agents operating there. I draw on literature and my own personal experiences and conversations to introduce the research context in Northern Province of Sri Lanka.

3.2 Selection of case study

I initially considered multiple sites for case studies; but it was not within the scope of this project to assess multiple conflict zones, which would be beneficial as each has unique disease burden.

I selected Northern Sri Lanka as it fits a number of important criteria for the research. Crucially, it fitted my criteria for a 'post-war' setting. At the time of planning, previous Government restrictions on the humanitarian community and international researchers meant that limited research had been conducted in the conflict-affected regions. Research in Northern Province and other Sri Lankan districts has identified a considerable MH burden in adults and adolescents [131-134]. In addition, there are culturally specific psychosocial well-being assessment tools that are validated, reliable and sensitive, designed for use with the target population [51]. From a practical perspective, GFP began a cycle of sport for peace programmes in Mullaitivu in 2014-16. I undertook an exploratory fieldtrip in August 2013 and concluded that a pragmatic evaluation of a SPPY would be feasible in this area.

During the planning phase of the project, foreign nationals no longer needed permission from the Sri Lankan Ministry of Defence to enter the Northern regions of Sri Lanka. The relaxation of border controls, and the lack of good quality research in the area, represented a pragmatic research opportunity.

3.3 Local geography and demography

Sri Lanka is an island located at the South Eastern tip of India, home to a multi-ethnic society. The majority of Sri Lankans are Sinhalese (74%), 18% are Tamil and 7% Muslim (often referred to as 'Moors' in Sri Lanka) ^[135]. Sinhalese are almost exclusively Buddhist, whilst Tamils are both Christian and Hindu. Literacy rates are high and life expectancy is 76.3 years ^[136]. Sri Lanka is divided into 9 provinces. Northern Province comprises five districts: Jaffna, Mullaitivu, Kilinochchi, Mannar and Vavuniya (APPENDIX B.1). Maritime Pattu division in Mullaitivu district was the location of my extended fieldwork (APPENDIX B.2).

Mullaitivu district lies on the north-eastern coast of Sri Lanka. Tamils are the most populous ethnic group and account for 84% of the population, 8% are Sinhalese and 7% are Muslim ^[135]. GFP intervention sites are distributed along the coast of Maritime Pattu and around Mullaitivu town. Maritime Pattu was severely affected by the 2004 Tsunami and the protracted civil war. Despite limited data, it is clear that the history and demography of the region has had a deep epidemiological effect on the region.

3.3.1 Civil War and the aftermath

In 2009 Sri Lanka emerged from 26 years of intermittent civil war. In 1976 a separatist movement, the Liberation Tigers of Tamil Eelam (LTTE or Tamil Tigers), claimed an independent Tamil state in North East Sri Lanka. The underlying causes of the war are complex, yet popular opinion is that the unfair distribution of economic growth and perceived injustice in political representation for Tamils were crucial factors.

The civil war and the ensuing socio-environmental conditions are not conducive to health and well-being. During the conflict thousands of people died, were injured, or went missing; many more were displaced, often multiple times. In the direct aftermath, almost all the residents of Mullaitivu were housed in internment camps. Even today, many people are still yet to find a long-term solution and face state led violence and oppression. As a result of the war Sri Lanka is divided along many lines: linguistically, religiously, ethnically, politically and geographically ^[137].

3.3.2 During the conflict

Throughout the conflict period, Sri Lanka as a whole continued to develop, and the Human Development Index (HDI) continued to rise ^[105]. Despite these indicators, conditions in the Northern and Eastern Provinces plummeted, especially in 2008-2009, during the final stages of the war. Education, health and economic development were all affected ^[138]. The conflict produced over 500,000 internally displaced people (IDPs), many who suffered multiple displacements and whose homes were destroyed ^[132]. This disrupted traditional livelihoods and left thousands unable to cope without international foreign aid. In addition, the majority of large-scale economic infrastructure was destroyed or badly damaged, which crippled the local economy ^[138]. The civil war also triggered mass emigration of Tamils, which drained the region of key human and financial resources in both the health and education sectors ^[139].

In May 2009 the Sri Lankan Military effectively wiped out the LTTE leadership ^[140]. The final stages of the conflict took place in Maritime Pattu and Puthukkudiyiruppu divisions in the north-eastern region of the country. These were arguably the worst affected regions of Sri Lanka. The actions of the Sri Lankan Military in Mullaitivu have been internationally condemned. They are criticised for the indiscriminate shelling of civilians, hospitals and humanitarian objects; the denial of humanitarian assistance, and for human rights violations both in and outside the conflict zone ^[141].

At the end of the conflict, Sri Lanka faced a massive humanitarian disaster yet effectively shut the gates to international assistance, bar a few actors. Relationships between the Government of Sri Lanka and the international community have remained poor, and they continue to come under fire from human rights organisations ^[142].

3.3.3 After the civil war

In the immediate post-war period, almost the entire population of Mullaitivu was housed in IDP camps, the nature and conditions of which have been internationally criticised. Hygiene was extremely limited and most had no access to safe drinking water or functional latrines ^[143].

Residents had restricted movement, contrary to international law; NGOs and International NGOs

(INGOs) were not permitted access, and all camps were over-crowded, under-resourced and ill prepared [144]. Camp residents were screened for affiliation with the LTTE and there are multiple allegations of torture and rape against Sri Lankan Army soldiers during this period [145]. During this period, the population was extremely susceptible to the psychological traumas of war. These people had almost ubiquitously witnessed human rights abuses, had been incarcerated in degrading camps and had their livelihoods destroyed.

3.3.4 After the camps

In the current post-war period, the majority of IDPs have returned to their previous homes, but many have failed to reach a durable solution and continue to face considerable difficulty in accessing basic necessities [144]. Relocation in Mullaitivu is still on-going, 6 years after the end of the civil war [144]. The previous government failed to make achievement of sustainable solutions a priority, favouring large-scale development projects and continued heavy militarisation. In 2014, only 6.4% of rural households in Mullaitivu had access to safe drinking water [146]. Those left destitute by the conflict have seen little improvement in their lives, despite millions of dollars of international aid, as the Military continues to dictate development in the region [147].

3.3.5 State-led violence

The elimination of the LTTE may have ended open warfare, but the cessation of violence is not tantamount to peaceful co-existence and equality. Non-violence does not inherently readjust society and create equal political systems that serve the needs of all the population. The pre-existing social problems and the root causes of violence endure. In the post-conflict era, the population of Northern Province continues to suffer from socio-political isolation and are the recipients of state-led direct, structural and cultural violence.

The Military continue to impinge upon reconstruction and reconciliation efforts in the area. They forcibly occupy large swathes of former residential and agricultural land for military camps and plantations, and continue to limit free movement in the region [144, 148]. The heavy militarisation of the Northern Province remains one of the main obstacles to positive readjustment [133, 147].

Fresh allegations of violence against civilians by soldiers, police and the intelligence service continue to emerge. A recent report stated that torture, abductions and sex and gender based violence (SGBV) are at higher levels in the post-war period than ever before [145]. Torture is still used to intimidate and oppress people in northern Sri Lanka, on the premise of national security. Those with a perceived link to previous Tamil Tiger activity are at an increased risk of torture across Sri Lanka [149].

Women, particularly widows, face a desperate lack of security and support. There has been a rise in prostitution, coercive sexual relationships and trafficking [150]. Reports suggest that violence, gender-based and general, is pervasive and growing in Northern Province. This is not just at the hands of the military. Academic researchers and focus groups conducted by INGOs have raised concerns about an increase in the abuse of alcohol, child abuse and levels of SGBV inside the home [150-153]. According to local residents, the struggling economy, unemployment and rapid socio-cultural change are all contributing to rates of alcohol and drug abuse, anti-social behaviour and school drop outs [151].

The breakdown of traditional family networks, livelihoods and community structures leaves a fractured community, cut off from the rest of Sri Lanka. Restrictions on the humanitarian community, which faces chronic shortages in funding, exacerbate the current challenges faced in the region [154]. Governmental security concerns have hindered formal assessments throughout much of the post-war period. There is a growing body of literature documenting the suffering of the residents of Northern Province [54, 155].

Put simply, the majority of people in Northern Province have been exposed to highly traumatic circumstances and the current lived environment lacks sufficient stability and support to enable everyone to make a proper recovery. The Military continue to impinge upon normal life and may provide a physical reminder of the suffering endured during the end stages of the civil war [133]. Of particular interest are the health-related impacts of the conflict and the ensuing post-conflict environment.

3.4 The mental health landscape

Three decades of war has had both direct and indirect health effects on individuals, families, communities and institutions [54]. Residents of Mullaitivu have been exposed to a broad array of risk factors for poor MH and PH as a result of the civil conflict. All of this has taken place against the backdrop of a reduction in MH support, both formal and informal [155].

The environment is a well-recognised determinant of health and MH [28]. Residents of Mullaitivu district experience many risk factors associated with poor psychological health that are common in LMICs [33]. To my knowledge, no systematic review documenting risk factors for poor MH in children in low and middle-income or post-conflict countries exists. Despite this, evidence suggests that child and adolescent MH is determined by individual characteristics, social relationships and the prevailing social environment [156].

Mullaitivu has complex health epidemiology, with multiple cross cutting challenges. The measurement and reporting of health-related outcomes remains poor in northern Sri Lanka, and the data that is available is dubious [157]. MH has been recognised as key component of primary care in Sri Lanka, but services are primarily based around institutional care, and published research on adolescent MH remains scarce [134, 155, 158]. This is despite compelling evidence that, in reality, there is a considerable treatment gap between needs and services, particularly for children and youth [59]. In the north, services remain fragmented and psychosocial assistance in the region is limited [159].

The previous section identified the presence of many risk factors for poor MH in Mullaitivu for adults and for young people [See Section 3.3] and research indicates high levels of exposure to war, tsunami and everyday stressors among young people [160, 161]. It is not possible to infer levels of MH and well-being based on the presence or absence of risk and protective factors alone; two different people exposed to identical environmental conditions can display different MH profiles [162]. MH determines, and is determined by, a broad range of social, economic, and environmental factors on an individual, community and social level, which work in manifold interacting processes [25].

TABLE 3.1 summarises the risk factors for poor MH faced by Mullaitivu residents, according to a socio-ecological framework. I chose a socio-ecological framework as other models of health promotion fail to capture to the multi-level, complex associations between MH and the environment. I feel the socio-ecological framework is the most appropriate in this case, as the purpose of TABLE 3.1 is to list the associations between multi-level exposures and MH outcomes, rather than explain them. I argue that the presence of large numbers of risk factors, and limited epidemiological data, in Mullaitivu is sufficient to warrant further investigation.

TABLE 3-1: Summary of risk factors for poor MH faced by residents of Mullaitivu

Level	Risk Factor	Level	Risk factor
Individual	Exposure to 'trauma' [33, 51, 163] Physical injury and ill health ² Internal displacement [33] Alcohol or drug abuse [28, 164] Child soldiering [165] Economic difficulties [28] Physical abuse and intimidation [145] Gender factors [33]	Community	Unemployment Destruction of social capital and social networks [54] Insecurity [33] Isolation Rural setting
Family	Sex and gender based violence [145] Child abuse [151, 153] Female-headed families Low income and assets [33] Loss of family Pathological family dynamics [161] MH problems in family member, including substance abuse [153]	Social	Rapid socio-cultural change Lack of access to economic opportunity [33, 166] Lack of social support Non-urban setting [146] Lack of formal MH services Institutional damage Insecurity [33, 166] Poor living conditions [132, 167] Political and social exclusion

3.5 Mental health problems: the epidemiological evidence.

Epidemiological research into the effect of Sri Lanka's conflict, tsunami and post-conflict environment on child and adolescent MH is limited [159]. Publications in adult populations have identified varying levels of common and severe mental disorders dependent, in part, on exposure to traumatic events and current life situation [51, 54, 132, 133, 155, 167-170]. Which stressors

² Including infectious diseases, NCDs and malnutrition

'matter most' remains a divisive issue in post-conflict MH research [171]. Studies on MH in Sri Lanka are primarily cross-sectional in nature, and limited by their sample size and assessment of highly localised populations.

Studies in young people in Sri Lanka indicate high levels of mental disorder. Elbert et al. (2009) conducted a cross-sectional study in 5th grade students in Jaffna. Only 8% of the sample reported having witnessed no traumatic experiences, and 79% reported having witnessed combat. Elbert et al. (2009) reported a PTSD rate of 23% in their sample; importantly a PTSD diagnosis impaired memory and language capabilities [172]. Sriskandarajah et al. (2015) conducted a multiple site cross-sectional study in Jaffna. Sampling locations were selected based on their exposure to the conflict and tsunami: PTSD prevalence varied by district according to exposure (Least-: most-affected, 1.7%: 33.6%). Several studies indicate high levels of suicidal ideation among young people, which is correlated with major depressive disorder and other substance abuse disorders [173]. Sri Lanka's suicide rate is now fourth highest in the world, although this figure is currently contested [134, 174].

Several authors support the 'Building Block Hypothesis': multiple exposures to traumatic events leads to poorer subsequent MH outcomes [175]. Much of the CAMH research conducted in Sri Lanka is concerned with the relationship between trauma exposure, child abuse and rates of mental disorder. Somasundaram (2010) provides an excellent summary of the common experiences faced by young people in northern Sri Lanka, which include 'shelling, shooting, helicopter attack, deaths, injury, detentions and many more' [54]. Exposure to mass trauma and experiences of family violence are independently associated with MH problems in children [161, 175]. Catani et al (2008) identified an association between war-related violence, family violence and PTSD in young children in Sri Lanka [161].

Family situation, partner violence and child abuse were strongly correlated with poor psychological outcomes in multiple studies in northern Sri Lanka. Abuse of alcohol in fathers' is associated with violence against children, and rates of harmful drinking were present in 1/5 households in a dyad and triad study conducted by Sriskandarajah et al. (2015) [153]. Exposure to traumatic events, and the psychological sequelae, impact upon daily functioning and impair

school performance [132, 134]. Fernando et al. (2010) indicated that everyday stressors are an important mediating factor in the relationship between war-related exposures and MH outcomes [163]. This suggests that changing elements of the current lived environment for young people might have an impact on MH in the region. The results of the study conducted by Jayasuriya et al. (2016) support this inference [133]. Alongside the lack of published CAMH intervention research and associated evidence, there is a lack of formal services available to those suffering from MH complications [176].

3.5.1 Mental health interventions

The role of international organisations has waxed and waned in Northern Province. Three key instances incited intervention in the last 15 years: the ceasefire in 2002, the Tsunami of 2004 and the official end to the civil war in 2009. Shifting government policy has created an uncertain terrain for potential interventionists, especially in the current post-war era. Currently, the majority of direct assistance to the conflict-affected population includes housing, livelihood development and grants schemes. There has been a recent decline in funding for development, health and psychosocial support programmes [167]. At the time of writing, the WHO is conducting an MH training intervention for primary care practitioners in Mannar and Puttalam; the results of this study are yet to be published [177].

3.5.2 Mental health assistance

MH has been recognised as a key component of health in Sri Lanka and has formed part of the primary care plan since the 1980s [176]. Despite this important recognition, the population of Mullaitivu presents a considerable challenge as it is very low-density, poor and almost entirely rural. This means that people have considerable difficulty in accessing the appropriate services and care. Psychosocial services have improved in the last five years with the gradual input of several international NGOs, but access and availability of services remains inadequate given the history of the region. UNICEF has played a role in the training of teacher counsellors and community counsellors who are trained to admit serious cases to hospitals. However, Mullaitivu has only one psychiatrist, stationed at the district hospital in Mullaitivu town [146].

There are very few child and adolescent-focused MH services or interventions in this region ^[155]. Psychosocial support has previously been informal, but with relocation and breakdown of community structures, current networks are unlikely to be sufficient ^[168]. International organisations have attempted to initiate psychosocial programmes for young people, only to have them forcibly shut down by the Presidential Task Force (a paramilitary force which is charged with the regulation of NGOs in Sri Lanka) in both Kilinochchi and Mullaitivu [Personal Interview. August. 2013] ^[154]. The liability of domestic politics and international relations has produced an unstable base for international organisations. This has led to short-term, insubstantial interventions.

Several publications call for the development of sports-based interventions for young people who have been affected by conflict ^[155]. To my knowledge, there are no organisations implementing sports-based programmes in Maritime Pattu division. A few organisations are implementing sport for development programmes elsewhere in Northern Province, including Handicap International and Terre des Hommes (Informal conversation Handicap International. August 2013). Previously there have been schemes to provide equipment for sport, but no lasting investment in the use of sport for development or psychosocial healing in the majority of Northern Province. This is despite the widespread popularity of sport, particularly among young men and boys.

3.6 Sport in Mullaitivu

Sport enjoys a prominent role in Mullaitivu but remains a male dominated sphere. The rural North is a conservative society in which traditional Hindu gender roles prevail. Groups of young men or adolescents often play sport in the evening, most often cricket, volleyball and football. Attitudes are often intensely competitive. On more than one occasion I have witnessed fights between players and even supporters during local 'friendly' matches (see FIGURE 3.1).



FIGURE 3-1: A crowd at a local football match descends into a fight.

The image below shows a local match on one of the best pitches in the district during the dry season. Images taken by A Hamilton in August 2013.

The environment is variable throughout the year. During the rainy season many pitches and fields remain totally waterlogged, whereas at the tail end of the dry season they can become unplayable dustbowls. Those facilities that remain playable year-round are usually military camp pitches and are not available to local residents.

School sports are not well developed in Mullaitivu. The majority of schools have a play area, but most remain unkempt; equipment is poorly maintained or non-existent. Sports competitions are held annually at district, provincial and national level. Other than during these periods, sport and PA are not a priority for schools. Competition for higher education is fierce and teaching methods are often draconian. Children are worked hard, and put under considerable pressure to achieve in regular examinations ^[134]. Almost all students take part in extra tuition after school

and on weekends, regardless of family income. As a result, sport and play activity in and around school time suffers.

Sport and play are potential intervention methods and have been employed in many different settings to achieve a variety of social and health objectives ^[178]. The widespread popularity and underdevelopment of sport in northern Sri Lanka means it may well be a popular intervention approach in the region, despite the limited expertise and equipment.

3.7 Conclusion:

Sri Lanka experienced a drawn out and violent civil war which has had a perceptible impact on its population, particularly those living in the northern regions of the country. Many people lost everything because of the conflict: family, home and livelihood. For many, the conditions in the post-conflict environment are improving steadily. However, there are those are yet to reach a long-term stable condition.

The growing evidence base suggests that the war and post-war stressors faced by adults and children will have had a considerable psychological impact. Without robust epidemiological information it is not possible to identify the conditions, structures and interventions that might improve MH and well-being in Mullaitivu. War affects all aspects of life, which, in turn, determine well-being, as summarised by FIGURE 3.2. The exact nature of this relationship is unclear at this moment in time, but it is evident that daily stressors partially mediate the relationship between exposure to conflict and subsequent MH problems.

Sport and PA are playing a more prominent role in the assistance provided to post-conflict populations in other areas of the world. The under-development of sport and PA for psychosocial well-being provides a research and intervention opportunity in Northern Sri Lanka. Initial field visits confirmed that a pragmatic evaluation would be possible in this context.

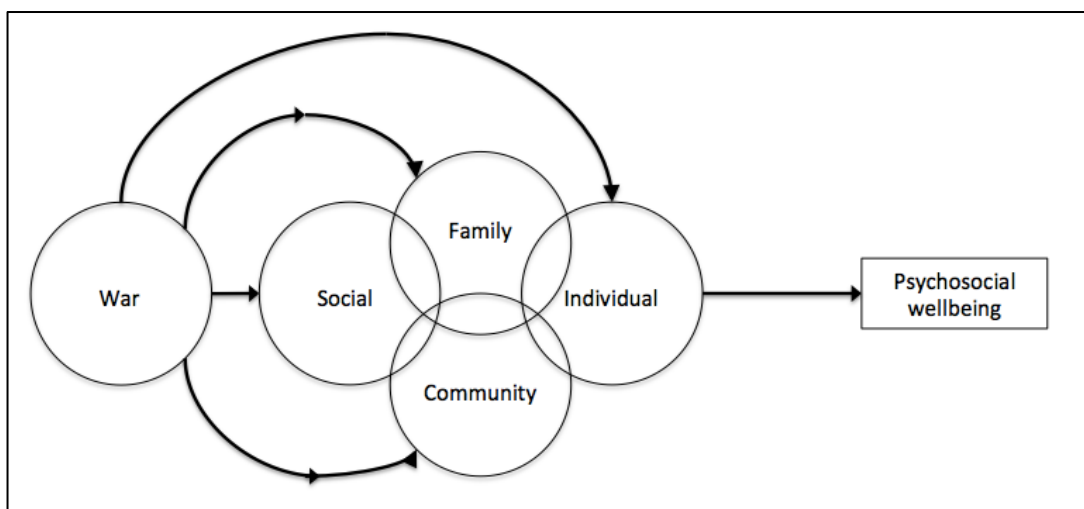


FIGURE 3-2: A summary diagram of assumptions of this chapter.

SUMMARY TABLE

- Sri Lanka is a multi-ethnic country off the coast of India that endured a destructive civil war between 1983-2015;
- The end stages of the civil war displaced the majority of Northern Province; combatants from both sides are accused of war crimes against the civilian population;
- The arms of the state meant to protect civilians continue to deny residents their human rights and freedoms;
- The population of Mullaitivu has been exposed to multiple risk factors for poor MH, during the conflict and in the post-conflict environment.
- There is little epidemiological information available for the region, but that which exists implies a considerable MH burden among children and young people;
- MH services are underdeveloped given the history of the region and international assistance remains limited;
- Sport and play might offer a novel intervention method for this region, given its prominence and popularity.

CHAPTER 4

Research questions, objectives and road map

4.1 Purpose of chapter:

The section will outline the over-arching objectives of my research and the research questions that I intend to answer. The specific methods will be covered in the relevant chapters. This section aims to bring together the work in previous chapters to define clearly identifiable and achievable research questions.

4.2 Research objective:

The primary objective of this research is:

“To contribute to the evidence base for the impact of sport for peace and development programmes on the mental and physical health of adolescents living in a post-conflict region.”

To achieve this aim I conducted a case study in Mullaitivu, which collected valid and reliable data using methods that aimed to optimize the level of evidence obtained, given the practical and logistical constraints of the research context. A literature review and previous chapters have identified gaps and assumptions to which the research questions relate.

- Young adolescents in Mullaitivu district will have high levels of psychosocial stress on account of their exposure to war- and non war-related stressors.
- PH, including fitness, is associated with MH in Mullaitivu (Assumption [C]).
- A sport for development and peace intervention may have an impact on the MH and PH of young adolescents in this context (Assumptions [A] and [B]).

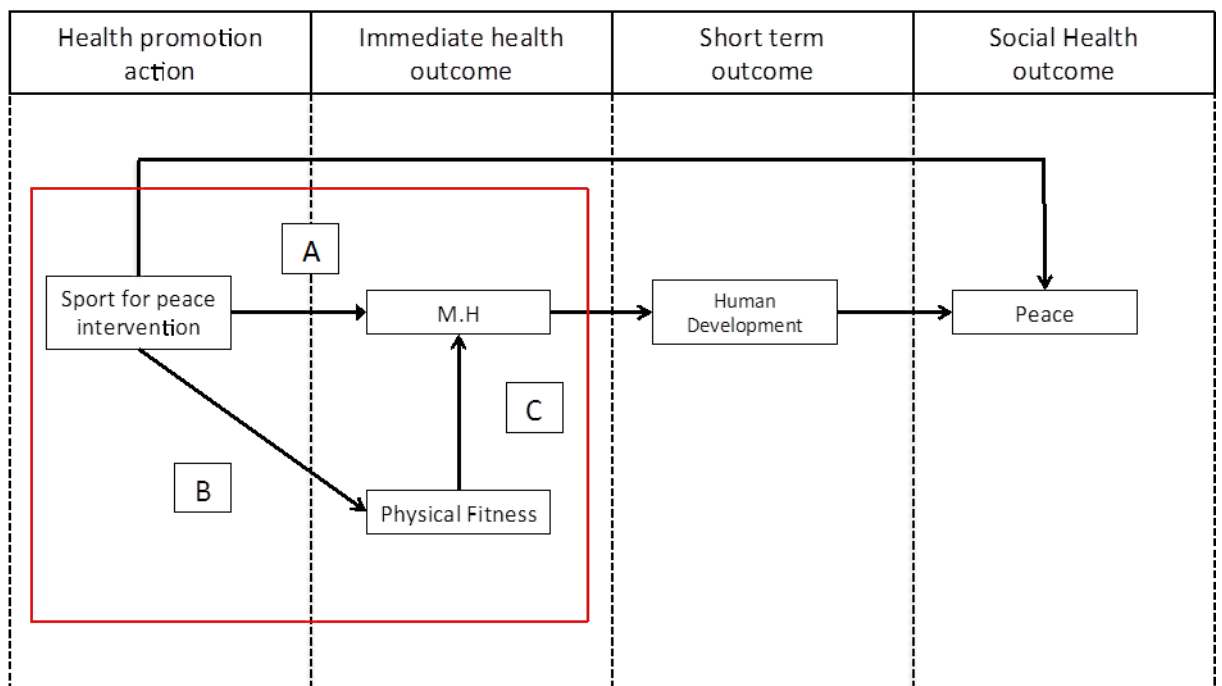


FIGURE 4-1: theoretical framework for the proposed project. Adapted from Blagojevic 2012.

Assumptions are noted in the diagram as A, B and C. The area highlighted in red is the aspect of the theoretical framework that the current project is concerned with.

The related overarching research questions are:

1. What are the baseline levels of, and association between, MH and PH in adolescents in Mullaitivu?
2. Was the SDP intervention effective?

(1) What are the baseline levels of, and association between, MH and PH in adolescents in Mullaitivu?

Prior to assessing the impact of an intervention it is crucial to identify measureable and reliable impact indicators and associated measurement protocols through pre-testing. Having assessed the quality of the methods it is imperative to establish the baseline levels of the outcomes variables prior to the intervention start date.

CHAPTER 5 identifies the methods and metrics used to measure MH and PH in Mullaitivu, Sri Lanka. The primary objective was to assess the feasibility and reliability of the adapted methods in the context, without compromising their accuracy.

Specific research questions:

- What measures are appropriate for use in northern Sri Lanka?
- How do the adapted measures perform?

CHAPTER 6 is a baseline study of the MH and PH variables. The results of the pre-testing informed the data collection methods for a cross-sectional study on the MH and PH of adolescents in Mullaitivu. The objective was to identify the baseline levels of MH and PH in the target population. This chapter compares the levels of MH and PH with regional and international norms and is also a needs analysis, prior to the initiation of an intervention. I also assess the relationship between PH and MH in the target population.

Specific research questions:

- How does the MH score performance of adolescents in Mullaitivu compare with regional norms?
- What are the baseline levels of selected PH variables in Mullaitivu and how do they compare with international and regional norms?
- Is there an association between fitness and MH? (Assumption [C] in FIGURE 4.1)?

(2) Was the SDP intervention effective?

The previous chapters form the basis for an evaluation of the impact of a voluntary, school-based sport for peace and development intervention on the MH and PH of adolescents in Mullaitivu. The theoretical framework for the study is outlined in FIGURE 4.1. The results of the study are discussed in light of the evidence base and the observed delivery of the programme.

CHAPTER 7 assess the impact of an SDP intervention on the MH and PH of adolescents in Mullaitivu. It addresses the assumptions highlighted in the theoretical framework.

Specific research questions:

- What is the impact of an SDP intervention on MH and PH (Assumptions [A] and [B] in FIGURE 4.1)?

CHAPTER 8 discusses and explains the results of the intervention where possible. It highlights the strengths and limitations of the evaluation in light of the current evidence base, theory and previous chapters.

- What are the implications for policy, practice and research in the context of current SDP theory and the available evidence base?
- What lessons can be learned for future sport for peace and development evaluation and programmes?
- How could this study have been improved and what work should be conducted in the future?

CHAPTER 5

Selection, adaptation and pre-testing of methods

5.1 Purpose of chapter:

The purpose of this chapter is to identify, explain and justify the selection of the methods for the measurement of the primary and secondary outcome variables. It describes any adaptations made to previously validated protocols. Lastly, it involves testing the feasibility and reliability of the adapted research methods and comment on their implementation in the main study.

5.2 Introduction

A key task in a formative evaluation is to select, adapt and pre-test methods and metrics and ensure that they are feasible for the context [179]. The methods selected aim to collect valid and reliable data that optimize the level of evidence obtained, considering the practical and logistical constraints of the research context. I consulted local stakeholders with previous research experience in the area and drew on my own field experience before finalizing the methods. This was to ensure the relevance and maximise the sustainability of the methods in a low resource setting.

Mullaitivu is an almost exclusively rural district with limited road access in many areas [146]. Access to testing sites is limited by road and weather conditions; vehicular access can be impossible in sustained bouts of poor weather. Reliable public transport is not widely available in the region. Motorcycles are an effective transport method as they are able to negotiate potholed and sandy roads. I decided to use three-wheeler motorcycles and motorbikes to access testing sites. This is because they are able to access hard to reach sites, considerably cheaper than cars and are the most common form of transport in the region. As a result, the equipment needed to be lightweight and easily transportable, preferably in a backpack.

After considering other options for sampling sites, I decided to collect data in schools. School participation rates are high (91.4% for ages 10-14), which maximizes project reach in a difficult context [180]. School based research also facilitates the tracking of participants over time. Schools in northern Sri Lanka are, in general, poorly equipped and have limited available support staff. Class size varies from 15-100+ children. Further, children and staff are under immense pressure to perform well in examinations, which are held termly. As a result, the methods needed to be suitable for sampling large numbers of children in a short space of time, with limited disruption of school time and no high tech equipment that requires a reliable electricity supply.

Conducting research in schools in northern Sri Lanka requires a full understanding of the complex bureaucratic processes that are, at first, opaque. There are several levels of clearance required before one can enter a school and begin testing. At the time of the research, any non Sri

Lankan required permission from the Ministry of Defence to access the intended research area and access was granted for a limited time only. This meant the planning, development and testing of methods had to be adapted accordingly. Before approaching schools, researchers must obtain permission letters from the Provincial, Zonal and District Officers of Education. This is a complex process that is difficult without considerable local support, as officials are almost always unavailable. At the time of the research, any methods that attempted to dig up the post-war experiences were strictly controlled and, according to local researchers, unlikely to obtain the relevant clearance from Government Officials or an ethical board. As a result, the methods had to be minimally invasive and remain as politically neutral as possible. This approval process highlights some of the idiosyncrasies of conducting research in this context.

The heavy militarization of the region and the mistrust of international researchers, particularly of MH researchers, was another important consideration in the planning and delivery of the methods. Personal and staff security was high priority, as many NGOs have reported intimidation of staff and participants in Kilinochchi and Mullaitivu. I took several steps to minimize the risk to participants. In addition, I made sure to explain the purpose of my work to all officials and obtain the relevant clearance letters. These were kept on my person at all times.

During the feasibility study I attempted to record some interviews with local sports club members. At the time of the feasibility study, this was extremely unpopular and drew a lot of negative attention from soldiers, who insisted on monitoring many of the interviews. This clearly altered responses and was a source of bias. As a result, I tried to minimize the amount of audio-visual equipment used in the study so as to draw as little attention as possible.

The importing of expensive technology is a security concern in a resource poor environment. Although theft is uncommon, it can draw attention to researchers and their work. This highlights another important concern, the importing of foreign equipment and methods that are not sustainable in the absence of international financial and technical support. From a practical perspective, facilities in schools are limited and could not support the use of lots of electronic equipment.

Local stakeholders indicated that there was a lack of research experience and capacity in the district, especially in the field of MH. This, coupled with the lack of English language skills and questionnaire experience among young people³, meant adjusting methods accordingly. In 2011 there were 400 children to each English teacher in Northern Province. As a result, levels of English are poor and so all research had to be conducted in either Tamil or Sinhalese. This meant that using tools that had already been validated in this language and context was the most appropriate approach given the time frame of the study. Additionally, methods had to be easily reproducible and require limited formal training to implement.

I believe that as a researcher you have several ethical and moral obligations, and the protection of project participants and staff should be of primary concern. This is particularly pertinent in a context like Mullaitivu, where residents may face intimidation and harassment after the researcher departs. Researchers should consider the ethics of their selected methods; appropriateness and sustainability are key concerns. For example, in Muslim culture, it is not acceptable for male researchers to touch female participants when taking measurements. The selection of methods involved balancing cultural, political and environmental concerns against the requirement for methodological rigor. I drew on my own experiences, stakeholder advice and a systematic literature review to inform my initial selections [91].

Although the selected methods have been validity tested in their original form, adaptations to the protocols were made to ensure their suitability for the research context. Gold standard validity tests are not feasible in this context due to the lack of resources and research experience. Reliability tests are possible and provide sufficient information to assess the suitability of the methods for assessing the impact of a programme; provided they are stable and sensitive to genuine change. The context limits the complexity of the validity and reliability testing that could be completed.

The Sri Lankan Index for Psychosocial Stress (SLIPS-C) has only previously been used as a cross sectional measure, and as such the test-retest reliability has never been examined. The original 49-item- Child version (SLIPS-C) had poor test-retest reliability in this context. As a result, I

³ Informal conversation, G. Fernando 2013

reduced the number of items on the questionnaire. The original version would not have been a useful programme indicator on account of its instability over time, despite having good construct validity ^[181].

The discussion section of this chapter follows the guidelines for writing a structured discussion outlined by Docherty and Smith (1999) ^[182].

5.2.1 Summary of pilot testing:

The planning and piloting of methods had three distinct stages. The first was a preliminary field visit to Mullaitivu in 2013. This was to familiarize myself with the context and plan an evaluation in consultation with GFPs local stakeholders. The second stage was a repeat measures study in Kilinochchi district in 2014. As a result of the initial repeatability study, I adapted the MH tool in an attempt to improve its test-retest reliability. The final methods testing took place in March 2015 in Mullaitivu before the implementation of the GFP intervention.

5.3 Methods: Pilot testing

5.3.1 Summary

I used a repeat measures design in both Kilinochchi (2014) and Mullaitivu (2015). I used a convenience sample at an easily accessible and compliant school. Participation was voluntary. Participants in Year 8 and 9 were included in the study. TABLE 5.1 is a summary of the adaptations and a timeline of the test-retest assessments.

There were some differences between the outcome variables in the two studies. The Kilinochchi (2014) repeat measures study compared the original version of the SLIPS-C (MH-1) (APPENDIX C.1) with an adapted version (MH-2) (APPENDIX C.2). The study aimed to assess the feasibility and repeatability of an adapted SLIPS-C. The study also assessed which tool was simpler to complete through informal participant and assessor discussions. The Mullaitivu study (2015) assessed the repeatability of a further adapted version of the SLIPS-C, the SLIPS-GFP (APPENDIX C.3).

The Kilinochchi (2014) study confirmed the inter-tester reliability of the anthropometric measurement protocol; these were not repeated in the second study. Measurement of the Multi Stage Fitness Test (MFT) was included in the second study as adaptations to the protocol were made after the first inter-tester reliability study.

I assessed the intra-tester reliability of the MH assessment protocol because I intended to keep the same Research Team Leader (RTL) for the duration of the studies. I assessed the inter-rater agreement for the PH assessment protocol because I was aware that I was likely to experience some shifts in availability of staff, and as such could not ensure the same data collectors would be available on each data collection day.

TABLE 5-1: Summary timeline of adaptations and test-retest analysis of SLIPS-C, Multi-stage Fitness Test (MFT) and Body Composition (BC)

Date	September 2013	December 2013	November 2014 (Kilinochchi)	January 2015	March 2015 (Mullaitivu)
MH	Permission to use SLIPS-C obtained	Initial adaptations to SLIPS-C. <u>MH-2</u> created	MH-2 test-retest analysis Feasibility of MH-1 compared with MH-2	Second wave of adaptation. <u>SLIPS-GFP</u> created (MH - 3)	Test-retest analysis of SLIPS-GFP
MFT	-	-	Inter-tester reliability of MFT protocol	Protocol refined	Repeat inter-tester reliability of MFT
BC	-	-	Inter-tester reliability of Body composition protocol	-	-

5.4 Study design:

5.4.1 Kilinochchi study:

The Kilinochchi study was a series of measurements over the course of 5 days. MH and physical PH were assessed on two occasions with a 3-day gap. The RTL assessed MH at T1 and T2. The Principal Investigator (PI) and an additional Research Assistant (RA1) conducted the MFT at T1. The remaining RA (RA2) measured body composition (BC) at T1. At T2, the RAs switched responsibilities. The study was conducted in Barathy Maha Vidyalayam. A study summary is below.

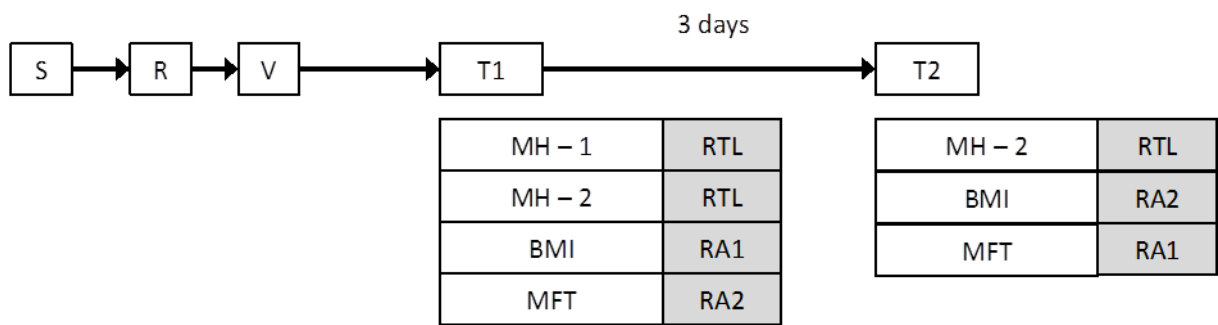


FIGURE 5-1: Kilinochchi study summary diagram

5.4.2 Mullaitivu study:

The Mullaitivu study was a series of repeat measures conducted over the course of a week. T1 and T2 were three days apart. The RTL sampled MH at T1 and T2. The study followed the same protocol as in Kilinochchi, but was conducted in Chemmali Government Mixed Tamil School in Mullaitivu.

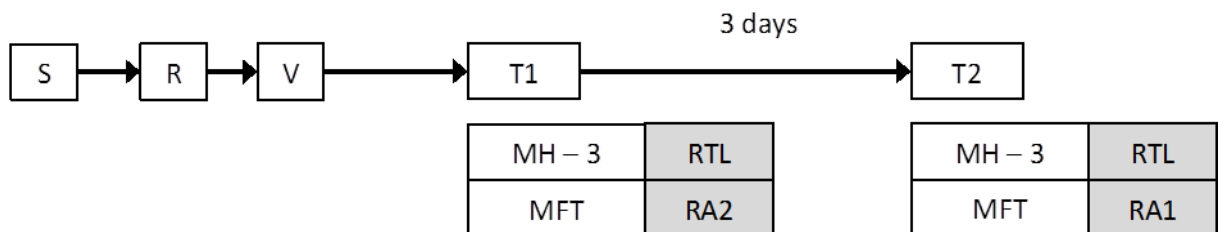


FIGURE 5-2: Mullaitivu study summary diagram

TABLE 5-2: List of abbreviations for study summary diagrams

Abbreviation	Item
S	Selection of school for study
R	Recruitment of participants
V	Volunteering for participation
T1	First measurement
T2	Second measurement
MH-1	Original SLIPS-C
MH-2	2 nd version of SLIPS-C
MH-3 (SLIPS-GFP)	3 rd version of SLIPS-C
BMI	Body mass index
MFT	Multi Stage Fitness Test
RTL	Research Team Leader
RA	Research Assistant

5.4.3 Sample Size:

Both studies used a convenience sample that aimed to recruit the maximum number of participants per school. The schools were selected based on their geographic location and ease of access. My relationship with the school principal was also an important factor in selecting schools.

5.4.4 Recruitment:

Participants were recruited through their school. Participation was voluntary and subjects were free to drop out at any time. All students in attendance were invited to take part in the study. Recruitment through a school was the only way to ensure participants would turn up for measurement on repeated days.

5.4.5 Tracking:

The study collected demographic data on each participant. This was used to track participants over the course of the study and throughout the analyses.

5.5 Outcome variables:

5.5.1 The Sri Lankan Index for Psychosocial Stress – Child Version (SLIPS-C).

There are several concerns when assessing MH in a post-conflict context [183]. The nature, experience and culture of the sample population are key elements in the selection of methods and data collection. Another important consideration is the utility of the information for local and international stakeholders [105]. I initially considered using the Strengths and Difficulties Questionnaire (SDQ), as this is available and validated in Tamil and Sinhalese [131, 184]. The SDQ is difficult to implement because it requires information from multiple sources (parents or guardians, teachers and test subject).

I chose the SLIPS-C because I felt that it reflected the Sri Lankan context well, and would provide valuable information for GFP and the Community Development Organisation (CDO), GFP's local implementing partner, in the design of future programmes. I identified the SLIPS-C through literature review and obtained copies of the tool through personal communication with the primary author. I assessed three different versions of the SLIPS-C and compared their reliability in the research context: the original SLIPS-C, (MH-1, APPENDIX C.1) the second version (MH-2, APPENDIX C.2) and a final reduced item version (SLIPS-GFP, APPENDIX C.3).

Children's distress and well-being in post-conflict zones have typically been conceptualised using Western diagnostic criteria, psychological constructs and self-report measures [163, 185]. A locally devised psychosocial assessment tool may provide better insight into MH than Western-derived measures in this context [186, 187]. Contextualised tools should provide relevant information for intervention design and development.

Bolton and Tang outlined a culturally grounded methodology for the design of context-sensitive MH assessment measures [188]. The method has shown success in other post-conflict settings [95].

¹⁸⁷¹. Fernando et al designed the SLIPS-C using a similar method. The SLIPS-C has been validated in both Tamil and Sinhalese populations in Northern Sri Lanka and shows good levels of internal consistency (Cronbach's alpha = 0.85)^[163]. Through personal contact with the primary author I was able to obtain advice on the implementation and interpretation of the tool.

The original SLIPS-C is a 49-item emic psychosocial well-being assessment tool. Participants indicate on a Likert scale (ranging from 0 = never to 4 = almost always) how often they experienced a certain feeling or emotion in the last two weeks. It has three sub-scales: internalizing behaviour, externalizing behaviour and withdrawal. A high SLIPS-C score has been associated with PTSD symptoms in previous studies ^[163].

Paraprofessionals have assessed large numbers of participants using the SLIPS-C in a similar context ^[163]. It requires no audio or video equipment and produces data that can be statistically analysed. Importantly, it is designed to reflect the Sri Lanka experience of MH.

5.5.2 Body composition: BMI for age (BFA) and Height for age (HFA)

I initially considered using alternative measures including body fat ratio, skinfold thickness tests and waist to hip ratio. Body composition was measured using BFA and HFA. BFA and HFA are crude measures of acute and chronic nutritional status respectively. Large-scale trials have been conducted in multiple countries and there are standard cut off points for weight categories. This allows for comparison between the target population and global and regional populations ^[189].

Calculation of BFA and HFA require the measurement of height, weight and age.

The protocols for the measurement of height and weight have been standardized, are easily reproducible and require minimal training ^[190]. I adapted the methods to ensure they were suitable for use Mullaitivu. In the standardised protocol height is measured using a Stadiometer. These are not widely available in northern Sri Lanka, nor are they suitable for transport on motorcycles. The adapted protocol used a tailor's tape, attached to a flat wall with masking tape. Height is then measured by placing a ruler on the top of the participant's head to identify height to the nearest 0.5cm. In the standardised protocol, weight is measured using electrical scales. In the adapted method, weight was measured using standard bathroom scales and then rounded to

the nearest 0.5cm. Electronic scales are not widely available in Mullaitivu, and I did not want to 'carry in' any external equipment where possible. The reliance upon unreliable batteries would have also been a concern in Mullaitivu.

Global standard growth references and thresholds have been published [189]. A reference population indicates the average BMI given the age and sex of an individual, as well as the distribution of individuals around the mean. The adapted BMI protocol is available in APPENDIX C.4. To ensure I had not reduced the compromised the reliability of the protocol, I tested the inter-rater agreement of the height and weight measurement protocol in the field.

5.5.3 Multi-stage fitness test (MFT):

CF was measured using the Multi-stage fitness test. The standardized MFT has been found to be a reliable and valid method to assess CF in school-age populations [191, 192]. Richards et al: (2011) tested the inter and intra rater reliability of the MFT protocol in a post conflict context [95, 193]. Improved CF is associated with improved health outcomes [123]. Low levels of CF in childhood and adolescence are associated with premature mortality, especially from cardiovascular disease [194]. The relationship between CF and MH is unclear, and remains understudied in LMIC settings [121]. The adapted MFT protocol is available in APPENDIX C.5.

5.6 Data collection procedure:

Prior to contacting any school all the relevant permissions were sought. The PI visited the principals to obtain permission and to explain the purpose of the study. Class teachers were given information sheets (APPENDIX C.6) and consent forms (APPENDIX C.7) in the relevant local language. Participants were also given information sheets at this time.

The research team convened at the target school each morning. The research team consisted of the following:

- (1) Principal Investigator
- (2) Research Team Leader

(3) 2 x Research Assistant

All researchers were provided with a notepad and recorded any problems encountered. At the end of the day all team members debriefed and discussed their day notes. Any adaptations made to the original protocol were noted accordingly.

On arrival at the school site the PI met with the Principal and discussed the protocol. After this was completed, the schoolteachers were briefed by the RTL.

Participants were briefed by the RTL and given the choice to opt out at this stage. Participants gave informed consent at this point. Participants were given a test sheet to complete and were guided through completion of the demographic information by schoolteachers and the RTL. The RTL led the questionnaire in each class, guiding participants through each question. The PI located a suitable place for the MFT with an RA and began set up.

5.6.1 Summary of data collection

STAGE 1	<i>Sri Lankan Index for Psychosocial Stress</i> : administered by the RTL and Class Teachers.
---------	---

STAGE 2	Station A	<i>Anthropometric measurements</i> : administered by RA.
	Station B	<i>MFT</i> : administered by PI and RA.

Once participants had completed Stage 1 (MH testing) they moved to Stage 2 (PH testing). Upon completion of both stages participants handed in their sheets and returned to class. The PI and RTL double-checked that all sheets were completed. If an item on the MH questionnaire was not complete, the individual was located and encouraged to complete the item in private. Injured participants did not take part in the MFT, although this only affected one participant. At the end of the data collection, the RTL led a class discussion to identify any difficulties with the questionnaire.

5.7 Analysis and presentation:

Statistical package for the Social Sciences 7 (SPSS) was used to enter and analyse data.

Feasibility results are presented in narrative form. All data were cleaned and significant outliers were removed prior to analysis. Descriptive statistics (mean, SD and 95% CIs) were calculated for all outcome measures. All assumptions underlying statistical tests were verified prior to commencing analysis.

5.8 Assumptions

The statistical methods used in this study assume that the samples were normally distributed and had homogeneity of variance (HOV). Normality was initially assessed using histograms for the crude results of each of the MH and PF variables. Initial assessment warranted further tests of normality and statistical tests were also completed. HOV was assessed using the Levene's Test. All assumptions underlying the statistical analyses are in APPENDIX C.8.

5.9 Mental health:

There are three crucial components of an evaluative measure: reliability, validity and responsiveness^[181]. The validity and internal consistency ($\alpha=0.85$) of the SLIPS-C has been established in previous studies^[195, 196]. One element of reliability is the extent to which a measure produces the same result in independent repeated trials, under the same test conditions (intra-rater reliability)^[197]. This can be calculated because the RTL conducted MH assessment at T1 and T2 in both studies. Intra-rater reliability was assessed using the Inter-Class Correlation Coefficient (ICC). The ICC gives an indication of agreement between scores, taking into account within and between item variations. A two-way random effects model for absolute agreement was used. Intra-reliability was calculated for both MH-2 and SLIPS-GFP. The responsiveness of the SLIPS-C cannot be estimated at this point because it has not been implemented in conjunction with an intervention^[181].

Data were cleaned according to the protocol outlined by G. Fernando, the corresponding author for the original papers, in personal communication. The data cleaning rules are outlined below:

TABLE 5-3: Table of cleaning rules for the SLIPS-C

Item	Event	Action	Justification
I saw birds	If answer = 0 (never)	Drop questionnaire	Sri Lankan children see birds daily
I told lies	If answer = 0 (never)	Drop questionnaire	Participant unlikely to have insight into own actions
I saw my friends (repeat item)	If two answers more than 2 scores apart	Drop questionnaire	Participant not paying attention to questions

5.10 Physical health:

Inter-rater agreement was calculated for height, weight and MFT to ensure the reliability of the adapted protocols in northern Sri Lanka.

The MFT is scored in levels and stages (e.g.: Level 4: Stage 1) [198]. There are several ways to report the results of the MFT. Some researchers use the MFT to predict VO_2 max, and form a transformation equation to apply at an individual level [199]. Additionally, researchers have converted MFT score into the speed (kmph) at the last completed level [191]. For the calculation of the ICC I have chosen to report the results as Level: stage. Many participants failed to score above level 1; they effectively had a speed last completed level of 0kmph, which causes problems when testing for normality.

5.11 Ethical clearance:

Ethical clearance for all the studies was obtained from the University of Oxford Tropical Research Ethics Committee (OxTREC). Local ethical clearance for all the studies was obtained from the University of Colombo Ethics Review Committee. The repeatability study was included as a subsidiary project within the main study. Letters are attached in APPENDIX C.9.

5.12 Results

The adaptations are reported in this section as results, as they were part of the iterative development of the final measurement protocols. In doing so I hope to recreate the logical steps that I took to creating the final MH assessment tool.

5.13 Summary of studies:

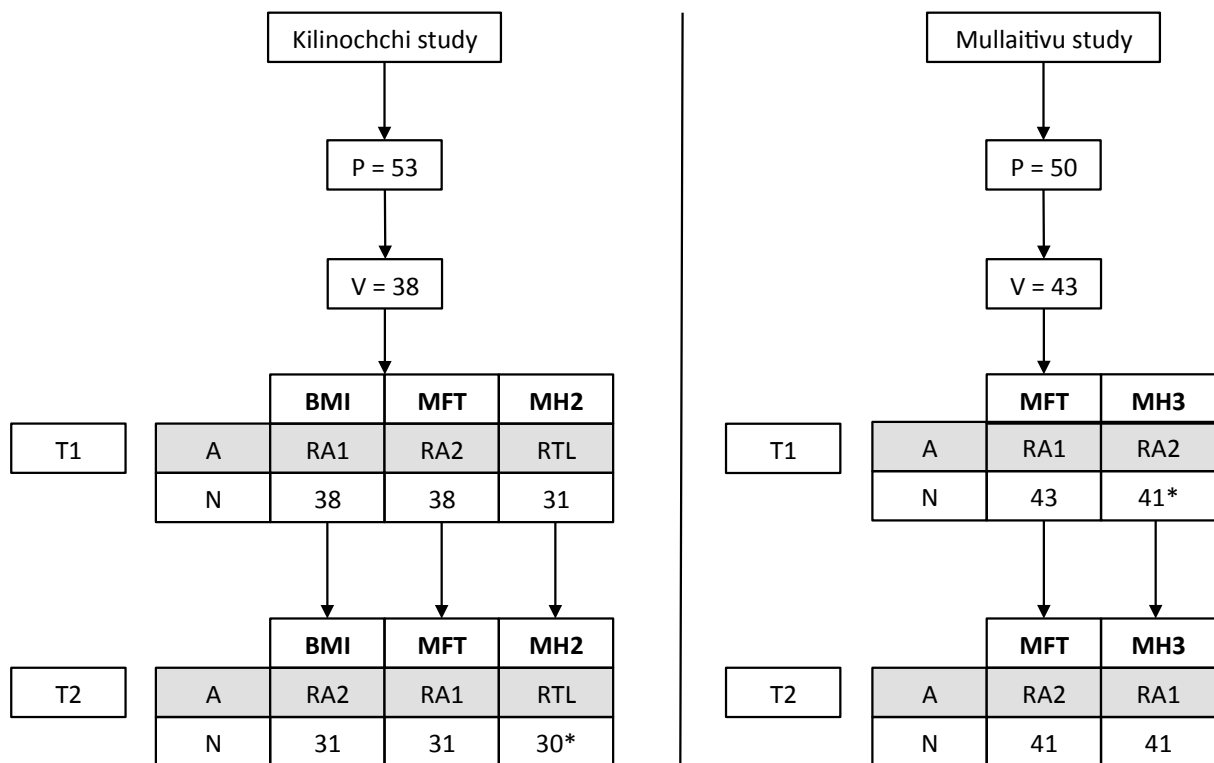


FIGURE 5-3: Study flow diagram for Kilinochchi and Mullaitivu repeat measures studies

- * Denotes where data was lost due to an incomplete questionnaire.
- NB: The number of participants in the subsequent reliability analysis is lower than the total number measured. Different individuals turned up at T1 and T2 and could not be paired for analysis.

TABLE 5-4: Table of abbreviations for summary study diagrams

Abbreviation	Item
P	Total number of students in the class
V	Participants present on the day of data collection
T1	First measurement
T2	Second measurement
T3	Third measurement
MH-1	Original SLIPS-C
MH-2	2 nd version of SLIPS-C
MH-3	3 rd version of SLIPS-C (SLIPS-GFP)
BMI	Body mass index
MFT	Multi Stage Fitness Test
RTL	Research Team Leader
RA	Research Assistant
A	Assessor
X	Unknown individual

5.14 Mental health:

5.14.1 Feasibility:

The SLIPS-C has been implemented previously in northern Sri Lanka in both a Tamil and Sinhalese population [163, 195]. The measure was validated in this setting and had high internal consistency ($\alpha = 0.85$). Through two waves of adaptation I improved the reliability of the SLIPS-C and the implementation protocol.

There were no major problems reported by the RTL or schoolteachers regarding the implementation of the questionnaire. As well as protocol changes, I made updates to the structure of the questionnaire and, eventually, to the number of items.

5.14.2 Pre-Kilinochchi changes:

The previous experiences of Dr G. Fernando were useful in the implementation of the MH test. Prior to the first round of data collection I made some adaptations to the original protocol and the structure of the questionnaire. I separated children from their friends and into mixed groups. This decision was based on a formative evaluation carried out by Dr Richards et al. in Uganda [193]. They reported that this reduced the amount of answer copying, which was also a concern in my initial data collection. Additionally, I encouraged the RTL to read through each item twice to ensure it had been understood before moving on. After questionnaires were handed in, each RT member checked each sheet at least once. This was to ensure there were as few missed answers as possible and improve internal validity.

I initially changed questionnaire to a 'tick box' format so it fitted on a single page of A4 paper. This is because printing is unreliable and costly in northern Sri Lanka. I also felt that it would reduce the error rate of children missing pages and forgetting to circle answers, as had previously been a problem (anecdotal evidence, Fernando: 2013). Examples of each method are below.

Original 'circle' answer method:

(1) I saw birds				
	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

'Tick box' method:

		Never	Rarely	Sometimes	Often	Almost Always
			x	xxx	xxxxx	xxxxxxx
1	I saw birds					

5.14.3 Results of Kilinochchi study:

According to class discussions, the MH-2 tick format for answering questions was more difficult to read and interpret than the original MH-1 circle format. Participants mentioned that they often lost their place in the questionnaire and had to locate the correct point to resume. This became evident when the MH-2 sheets were handed in and checked, as there were lots of gaps where students had lost their place and then recommenced answering in the wrong place.

The error rate in the tick box format questionnaires was higher than in the circle format. The error rate in the tick box format was 3/10 sheets and 0.5/10 sheets in the circle format in the test-retest sample. Participants often missed lines, double-ticked lines or marked an answer and crossed it out without returning. This increased the rate of random error. This was corroborated by the RTL who said that it was necessary to stop more often in the tick box format to allow a student to locate the relevant position than in the circle method.

An additional source of error was that participants began to tick down one side for extended sections of the questionnaire. This was noticeable when entering the questionnaire into a data spread sheet. There are reverse coded items in the SLIPS-C; any individual simply ticking down

one side will fail to identify these reverse coded items. The score of the individual will then mirror the questionnaire effectively giving them scores of only 0 or 4 throughout the test.

The cause of this is difficult to identify. It is unclear whether this is a result of the length of the questionnaire, the layout or any other factor. However, teachers and the RTL felt the questionnaire was long and participants had trouble focusing for the duration of the test. This is especially pertinent as, according to Fernando, Northern Sri Lankans have little experience with questionnaires. In response to these initial observations, and the poor test-retest reliability of the measure, I decided to reduce the number of items on the questionnaire and to revert to the circle answer method before a second round of testing in Mullaitivu.

5.14.4 Reducing the number of items on the questionnaire

When reducing the number of items on the questionnaire, I consulted a senior statistician and a qualitative researcher. Using an item correlation matrix I identified items that performed well and were associated with one another. I began to remove items that were redundant or poor. For example, several items had an average score that was well below one, I concluded that these were poor items in this context as the symptoms were not present in the target population. In addition, if the inclusion of a poor-scoring item lowered the overall internal consistency of the score, I felt it was justifiable to remove the item from the SLIPS-GFP. I maintained a balance of the three sub-scales whilst reducing the items on the questionnaire. I made sure to maintain the internal consistency of the questionnaire by using the cross-sectional data from Kilinochchi in 2014 to support my adaptations. The reduced version of the questionnaire (SLIPS-GFP) is available in APPENDIX C.3.

In addition to removing items from the main SLIPS-C, I also removed an additional item in which participants were asked to rate their life from 0-10 at the time of the questionnaire. During the initial wave of testing this was often left incomplete and, from my observation, participants appeared to be copying one another on this particular question.

5.14.5 Mullaitivu

There were no problems reported with the implementation of the SLIPS-GFP protocol. Error rate using the SLIPS-GFP version was lower than the MH-2 version. The SLIPS-GFP was quicker to implement and, as seen below, had better intra-rater reliability than previous versions.

5.15 Descriptive statistics

A small number of participants were sampled using MH-1 version to provide feasibility information, these participants were not included in the intra-rater reliability analysis. The crude mean and 95% confidence intervals of the SLIPS-C scores are presented in tabular form below. Data were cleaned according to the protocol outlined by Dr G. Fernando (email communication, 2014) and outliers were removed.

TABLE 5-5: Crude mean and 95% confidence intervals for SLIPS-C score (MH-2 version) in Kilinochchi (T1 n = 38, T2 n = 31)

Outcome variable	Crude mean and 95% confidence limits	
	T1	T2
n	38	31
% girls	50%	58%
SLIPS-C (MH-2) score (95% CI)	55.21 (50.20 – 60.22)	53.30 (48.11 – 58.46)

TABLE 5-6: Crude mean and 95% confidence intervals for SLIPS-C score (GFP version) in Mullaitivu (T1 n = 41, T2 n = 41)

Outcome variable	T1	T2
n	41	41
% girls	34%	46%
SLIPS-GFP score (95% CI)	20.60 (17.57 – 23.70)	19.22 (16.35 – 22.09)

5.16 Intra-tester reliability

The underlying assumptions for the intra-tester reliability analysis were satisfied after transformation (APPENDIX C.8). Intra-tester reliability for the MH-2 version of the SLIPS-C was good ($r = 0.701$). Reduction and refinement of the questionnaire improved the intra-tester reliability of the tool from good ($r = 0.701$) to excellent ($r = 0.878$) [200]. The second wave of adaptation also improved the internal consistency of the measure in the sample population, as measured by Cronbach's Alpha ($\alpha = 0.899$). Results are tabulated below:

TABLE 5-7: Test-retest reliability of the adapted SLIPS-C, assessed with the ICC (MH-2 n =31, SLIPS-GFP n = 28)

SLIPS-C version	N	Cronbach's Alpha	ICC
MH-2	31	0.701	0.701
SLIPS-GFP	28	0.899	0.878

5.17 Physical health

PH was assessed using measures of BC and CF. The adapted protocols were repeatable in this context and reliable. As a result of the formative evaluation, I made minor adaptations to the original protocols. The selection process and adaptations are reported in this section as results, as they were part of the iterative development of the final measurement protocols.

5.17.1 BMI

I initially considered other methods to analyse body shape and fat distribution, but alternative methods were not suitable in this context because of taboos in Hindu and Muslim culture concerning contact between males and females. For example, waist circumference measurement requires physical contact between the participant and the researcher. Further, these measures often require special equipment not available locally or are harder to record consistently or accurately, especially for larger samples [201].

In the standardized protocol for measuring height and weight, an electronic scale and a stadiometer are used [190]. These are not widely available in northern Sri Lanka, nor are they suitable for transportation on a three-wheel motorcycle. Sri Lankan schools are encouraged to measure and monitor the height and weight of pupils. As a result, the equipment required for the measurement of BMI is available locally. Typically spring scales are used to measure weight and a ruler against the wall is used to measure height in Sri Lankan schools. I adapted this protocol in the following way.

For the measurement of height, a tailors tape was used and fixed to an appropriate wall with electrical or parcel tape. A ruler was then placed on top of the head to ensure an accurate reading. Height was measured to the nearest 0.5cm. Weight was measured using a standard bathroom scale to the nearest 0.5 kg.

5.17.2 MFT:

I initially considered measuring PF using a sequence of tests. Many standard PH tests were not appropriate for this context. My own personal experience using an adapted MFT in a post-conflict setting was important in the selection of the method. Evidence indicates that physical abilities and motor skills are strongly associated with one another; suggesting one test can reflect overall fitness in youth when additional fitness tests are not practical [202].

The multi stage fitness test has a standardized protocol using a 20-meter measurement tape, an audio file and cones [198]. The equipment is available locally. The audio recording for the test is freely available for download. However, I purchased a speaker that operated on a widely available mobile phone battery. This meant that I could purchase several replacement batteries and recharge them at my next convenience. This avoided purchasing multiple sets of single-use batteries. The speaker was left with the local organisation at the end of the data collection period for use in further testing. Sports cones were available locally, but expensive. Small plastic food bowls were a suitable low cost alternative.

The MFT is susceptible to systematic and random error. A list of important factors for the test is listed below:

- Weather
- Ground conditions
- Footwear and Clothing
- Attitude (group and individual)
- Cultural norms
- Explanation
- Pacing strategy
- Motivation

Many of these factors are difficult to control. It is crucial that assessors create positive group mentality during the test. The attitude of the participants and assessors was particularly salient in this context. It was clear that assessors were not comfortable vocally and physically encouraging MFT participants to continue the test and try to reach maximal effort. Cultural traits such as this can cause systematic error; the implications and steps taken are discussed in lieu. Likewise, since the data collection was conducted in several schools, there was considerable variation in the quality of the surface on which the MFT test was held. FIGURE 5.4 highlights the difference in school field quality in Mullaitivu.



FIGURE 5-4: Girls running the MFT in Mullaitivu

Photograph taken by A.Hamilton 2015

5.17.3 Descriptive statistics:

The crude mean scores for weight, height and MFT at T1 and T2 in Kilinochchi and Mullaitivu are listed below. Height and weight measurements were not taken in the Mullaitivu study as the results in Kilinochchi indicated that the protocol was reliable and repeatable.

TABLE 5-8: Crude mean and 95% confidence intervals for PH variables in Kilinochchi and Mullaitivu (KT1 n = 38; KT2 n = 31; MT1 n = 43; MT2 = 41)

Location	Kilinochchi		Mullaitivu	
	N			
N	38	31	43	41
Time	T1	T2	T1	T2
Weight (kg)	37.92 (35.04 - 40.44)	37.42 (34.67 - 40.14)		
Height (cm)	150.64 (148.72 - 152.57)	150.13 (147.87 - 152.39)		
MFT score (Stage: Level)	4.20 (3.74 - 4.67)	5.48 (4.88 - 6.08)	5.71 (5.14 - 6.29)	6.17 (5.62 - 6.37)
MFT score (kmph)	9.29 (9.00 - 9.57)	9.96 (9.62 - 10.31)	10.07 (9.77 - 10.37)	10.32 (10.02 - 10.64)

5.17.4 Inter-tester reliability

The assumptions for the use of the ICC in an inter-tester reliability analysis were satisfied (APPENDIX C.8). Inter-rater agreement for height, weight and MFT are listed below. Height and weight measurement showed excellent levels of inter-rater agreement (height $r = 0.992$, weight $r = 0.993$). The inter-rater agreement for the MFT improved between the two studies (Kilinochchi $r = 0.777$, Mullaitivu $r = 0.830$), although agreement was already very good [200].

TABLE 5-9: ICC statistic for the PH measures repeatability study in Kilinochchi and Mullaitivu (K n = 31; M n = 27)

Outcome	ICC	
	Kilinochchi n = 31	Mullaitivu n = 27
Weight	0.992 (0.983 - 0.996)	
Height	0.993 (0.985 - 0.997)	
MFT (Stage: level)	0.759 (0.558- 0.876)	0.830 (0.632 - 0.922)

5.18 Discussion

5.18.1 Principal findings:

This study developed and tested adapted protocols for the assessment of MH and PH in a low-resource setting. The adapted protocols built on previously validated methods, were feasible and show field reliability.

Successive adaptations to the SLIPS-C ensured a final measurement tool (SLIPS-GFP) that was reliable, as demonstrated by the excellent intra-rater reliability, especially considering the research context. For example, FIGURE 5.5 is a photograph of the MH reliability testing in Mullaitivu, in which we were not allocated a classroom for our measurements at T2. Discussions with teachers and children ensured the tool was user-friendly; this should limit random error as a result of participant miscomprehension.



FIGURE 5-5: T2 MH assessment in Mullaitivu.

Photograph taken by A. Hamilton in March 2015.

The adapted protocols for the measurement of height, weight and MFT were reliable and consistent. Inter-rater reliability was excellent for all PH outcome variables; only the MFT was

below an agreement level of 0.95 or above. The MFT was susceptible to between-rater variation because it can be quite challenging to deliver at times. Likewise, clear explanation of the protocol was essential to ensuring participant compliance. The ability to clearly explain the test is a source of inter-rater discrepancy.

5.18.2 Weaknesses:

The principal weakness in this study was the lack of a comparison between the adapted protocols and 'gold standard' measures ^[203]. This was true for both the PH and MH variables. There were some broad methodological challenges that meant I was unable to fully establish the psychometric properties of the adapted MH protocol. The adaptations of the original protocols and the nature of the research context may have introduced additional bias.

Changes to the number of items on the SLIPS-C might have affected its validity. The initial adaptation to a tick box format reduced internal validity, by increasing the error rate considerably. Reducing the number of items in a measure may affect content and construct validity; a reduced item measure may not assess the construct of interest satisfactorily, and may fail to capture important components of psychosocial wellbeing ^[204].

Reducing the number of items may have also reduced the sensitivity of the tool and fail to accurately distinguish between participants. An associated point is that the improved reliability of the reduced SLIPS-C may in fact be an artefact of reduced sensitivity. By reducing the number of items on the scale, I may have effectively reduced the capacity of the measure to identify change, an important reflection considering the intended use of the SLIPS-C as a programme indicator. In addition, we could not assess the concurrent or predictive validity of the adapted MH measure in this study, which would certainly have been beneficial given the intended use of the SLIPS-C ^[205].

The same was true for the PH variables. In only measuring to the nearest 0.5cm/0.5kg I reduced the sensitivity of the anthropometric measurement methods, which might have reduced the capacity to identify change.

The lack of relevant research experience in the community influenced the planning, design and implementation of the protocols. The selection of a local school was double-edged sword; the relationship was necessary for access, but may have encouraged a laid back, informal approach to data collection.

The choice to hold the study in a school may have reduced the external validity of the protocols for the wider population, as a result of selection bias. School attendance was high in the district, but there were still many children who did not attend school. These children are more likely to have poorer literacy skills, and potentially have poorer psychosocial well-being ^[158]. These examples are indicative of the research context and highlight the importance of training and ***maintaining*** a research team in this context, where there is very limited capacity to conduct research.

The limited participant numbers in the study meant that boys and girls data was combined. Not stratifying by gender or age affected the distribution of the data and meant that we could not assess age or gender specific differences. This created normality concerns and data had to be transformed for statistical analyses in some instances. In addition, no sample size calculations were performed for either study.

I only assessed the inter-rater reliability for the PH assessment methods. The intra-rater reliability of the PH assessment methods is unknown in this particular context. The inter-rater reliability is an assessment of the training of the data collection team; the intra-rater reliability assesses the degree of agreement between multiple admissions of the same test by the same person in the same experimental conditions. The lack of this information for the PH variables was a weakness in this study. Likewise, the lack of inter-rater agreement for the MH assessment protocol was a weakness. In the event that the RTL was absent, a different assessor took their place, which may have been a potential source of bias in the latter studies in this thesis.

5.18.3 Strengths:

Despite the weaknesses highlighted, there are some considerable strengths of this study.

Successive waves of pre-testing allowed continued adaptation and incorporation of ideas and

changes in the build up to the intervention. The opportunity to do multiple rounds of formative testing ensured that the MH measure was repeatable, which is crucial. The pretesting of methods and materials is uncommon in this setting, where stakeholder priorities are often at odds [105, 107]. Funding bodies are typically interested in the impact of programmes rather than extensive pre-testing, and many implementers see methods testing as an unnecessary drain on resources [206]. Despite this, pre-testing is a crucial component in the planning and delivery of an effective intervention and associated evaluation [179].

The selection of locally devised MH assessment methods contributes to local ownership and sustainability, and minimises risk to participants [207]. Researchers should be confident in the suitability of their methods, as the presence of the military is a constant threat to sustainability and security of any project in Mullaitivu.

Mullaitivu is a very low-income district; the materials and methods were deliberately selected so that they could be implemented without specialist training or equipment to ensure they were as sustainable as possible. It is of note that any formative testing exceeds previous efforts in this setting and the final results were positive.

5.18.4 Implications:

The pre-testing of methods highlighted important practical and theoretical challenges for MH and PH measurement. There are considerable methodological challenges, but perhaps the most relevant is that lack of experience in research in the area. The reliance upon volunteer work is a problem; it is difficult to develop a skilled research team when members are constantly unavailable due to other commitments.

The poorly collected data at T3 indicates that the same assessor should be used for data collection in the main study. This was not possible in all the schools in the main study as data collection takes place in Tamil, Sinhala and Muslim schools. I believe it was paramount to have a native speaker conduct the questionnaire, so that they can assist the participants properly. In the subsequent intervention study, it was not appropriate for a man to conduct the

questionnaire with female participants in Muslim schools. I attempted to maximise the continuity of the research team throughout the data collection process.

To address the concerns about the validity of the adapted SLIPS-C I re-introduced a validated self-rated happiness scale [208]. Provided the happiness scale is valid in this context, I could assess concurrent validity. In addition this potentially contextualised the adapted SLIPS-C, which has not been used in conjunction with an intervention or any other MH assessment tool.

MFT scores were reported as kmph in the baseline analysis, rather than in the stage: level format. This allowed for comparisons with global and regional norms. Additionally, to limit measurement error in the MFT I encouraged standardization in the explanation of the protocol. This was difficult given the language barrier. I ensured to provide a Tamil and Sinhala script that was read at the beginning of each MFT test. In this way, I attempted to reduce the potential for between observer variations. The nature and attitude of the MFT assessors was paramount for maintaining validity in this test. For the MFT measurement, I attempted to recruit individuals who had an interest in sports and physical education and, where possible, had experience in assessing PH in adolescents.

During the pre-testing and feasibility study, I became aware of several cultural idiosyncrasies that helped with access to schools and participants. Developing a relationship with school principals was a crucial factor for access and should be considered before any attempt to collect data. In the main study, the intervention sample size was predetermined; there was little that could be done about loss to follow up from the cohort. However, by ensuring a positive relationship with the school, I endeavoured to sample in each site on two days. This ensured that any missing data were picked up in a second round of data collection. Perhaps one of the most important elements of the pre-testing of methods was to expose the inexperienced research team to the requirements of conducting scientific research; ensuring implementation in the subsequent study could be conducted efficiently and efficaciously. Certainly, it was essential for me to begin to understand the limitations and broad methodological challenges of operating in this context.

SUMMARY BOX

- Selection, adaptation and pre-testing of methods is a crucial and informative stage of evaluation design;
- Data collection in Mullaitivu presents unique and difficult practical challenges;
- I chose to assess MH using the locally devised and validated SLIPS-C, and PH using measures of BFA, HFA and the MFT;
- Inter-rater reliability was excellent for all the PH assessment protocols;
- Successive adaptations to the SLIPS-C ensured a final measurement tool (SLIPS-GFP) that was reliable, as demonstrated by the excellent intra-rater reliability;
- Practical challenges in the organisation, training and retention of a research team will be crucial in the main study.

CHAPTER 6

Baseline measurements of mental and physical health

6.1 Purpose of chapter

The purpose of this chapter is to establish the baseline levels of MH and PH parameters before the start of the SDP intervention. This involves three separate sections: analysis of the sample distribution, comparison of the sample with global and regional normative values and examination of the association between the outcome variables.

6.2 Introduction

The Sri Lankan health experience is not uniform; there is considerable regional inequity. NCDs are now the leading cause of mortality in Sri Lanka and account for 75% of all deaths [209]. In contrast to much of the island, Mullaitivu is a predominantly rural, low-income district in which NCD prevention is not yet a top priority. There is a distinct lack of information on major health indices for the district, especially regarding MH and psychosocial well-being. The Ministry of Health is scaling up MH services, yet they remain inaccessible and insubstantial [210]. In the case of Mullaitivu, a MH deficit is assumed, rather than systematically measured and addressed appropriately.

There has been widespread exposure to potentially traumatising events in Mullaitivu [168]. Reactions to traumatising experiences vary among affected individuals, groups and communities [57]. Some individuals and groups will present normal stress reactions and, over time, recover without specialist intervention [69]. Others may cope less effectively and may develop symptoms of more severe mental illness. This is more likely to happen in individuals who have faced multiple exposures and those who lack basic support and psychosocial care, often provided informally in low-income settings [206].

There are many groups in Mullaitivu who have not yet obtained a durable solution, and face daily hardship and struggle; these individuals are particularly at risk of MH problems [2, 166]. The scale of the damage to the systems that support healthy life is widespread [210]. The establishment of a supportive, healthy community is an essential component of long-term recovery, human development and sustainable peace in Mullaitivu.

Mullaitivu remains a community divided, particularly between the Tamil and Sinhalese ethnic minorities. The Sinhalese residents of Mullaitivu were recently settled as part of a government programme in the division of Welioya, a previously Tamil-dominated area. Welioya is fertile land with important access to fishing areas and the relocation programme was widely opposed in the Tamil community. Welioya lies approximately 30km to the south of Mullaitivu town, and there is limited interaction between the two communities. A recent National Youth Survey

indicated that 80.6% of Sinhalese and 59% of Tamils believed that ethnicity was the most divisive factor in Northern Province [211]. This geographic distance is exacerbated by ethnic and linguistic differences, resulting in opposed communities.

Mullaitivu town and the adjacent area is a chiefly Tamil zone, pocketed with small Muslim communities. The two ethnic groups do not always interact harmoniously, and there is widespread distrust and dislike of Muslim cultural practice among the Tamil community (interview: K).

Drawing on the work of Blagojevic (2012), I hypothesised that the establishment of a healthy and engaged community might help to prevent the relapse into conflict and to promote widespread, equitable human development, as outlined in the theoretical framework in section 2.8.

The Community Development Organisation (CDO) planned and developed the parameters of a SDP intervention in Mullaitivu, prior to my involvement. In accordance with best practice advice for the development of effective programmes, local stakeholders were engaged in the planning process from the outset [57, 212]. The primary planning of the intervention was conducted in a Generations For Peace training week in Vavuniya in 2013. However, the intervention development was carried out without any formalised needs assessment, and is largely based on local opinion.

The planning group decided to implement a two stage schools-based sport for peace and inter-ethnic dialogue intervention. The first stage aimed to develop the confidence, self-esteem and interpersonal skills of children through sports-based games. In the second stage, CDO would host inter-ethnic sports events to promote acceptance and dialogue between youth from different ethnic groups.

The programme outlined by the CDO has the potential to improve the psychosocial well-being of young people. The National Youth Survey indicated that PA levels were decreasing in Sri Lanka, and that only 25% of students were actively involved in sports [211]. There is good evidence that engaging in sport and play can be beneficial for MH of young people [111, 114]. There is growing

evidence that fitter adolescents have better MH, but this evidence is primarily from developed contexts [119, 213]. Increased PF, as a result of increased PA, may lead to improved PH and subjective well-being [214]. If correctly harnessed, sport and play initiatives have the potential to contribute to increase positive social interaction, self-confidence and self-esteem [215]. They may also play a role in directly improving MH, through the prevention of depression or anxiety-related syndromes [126, 216].

A needs assessment is a vital part of any intervention planning process. Such an assessment should involve the collection and analysis of information to determine whether a response is necessary and, if so, then what kind of response [162]. In such a way, intervention planners can identify locally relevant priorities for action. Local knowledge is necessary, but not sufficient, in the planning, implementation and evaluation of interventions [105]. It is important to identify whether assumptions are based on beliefs, theories or experience and test them accordingly [212].

With this in mind, I conducted baseline measurements to test the assumptions behind my proposed theoretical model. The data collection was designed to test the following hypotheses:

- that there is a MH deficit in the target population;
- that fitness is associated with MH in the target population;
- that there are no significant differences in the MH or CF between the control and intervention groups included in the study.

There were considerable logistical, political and time constraints that limited the quantitative data I could collect. There were some potential confounding factors that I recorded during the baseline analysis:

- Presence of staff
- Facilities at schools
- School setting and degree of isolation

Through negotiation I was able to convince CDO and GFP to include schools which would eventually constitute the control group in the pre-post study, outlined in CHAPTER 7. I struggled

to recruit sufficient schools to ensure I had a large enough sample size for subsequent statistical analysis. GFP and CDO felt that to increase the intervention sample size any further was not possible, as it would place too much strain on the programme leaders.

6.3 Methods

6.3.1 Summary

The study design for the baseline assessments was cross-sectional. A local organisation (the CDO) and a delegated GFP Country Lead selected all the schools included in the study.

Participation was voluntary and informed consent was taken from each individual and their guardian.

6.3.2 Sample size calculations

The number of participants in the intervention and control group was pre-determined by the local partner organisation, based on the intervention capacity. However, I conducted independent sample size calculations to ensure that my evaluation was ethical and warranted.

Since no other study had used the SLIPS-GFP as an outcome variable, I have extracted the standard deviation from my own study conducted in Kilinochchi in 2014 (n=461). To do this, I extracted the 22 items that are included in the SLIPS-GFP from the longer SLIPS-C and calculated the standard deviation (11.81).

Since no previous project has outlined a value for significant change, I followed convention and chose a 5% change in mean value of the primary outcome variable. 5% of the total maximum SLIPS-GFP score is 3.52. Sample size calculations use a standard formula based on the probability of type I and type II errors, standard deviation and significant change.

EQUATION 6.1: Standard formula for calculating sample size

$$n = \frac{(Z\alpha + Z\beta)^2 \times \sigma^2}{\square^2}$$

- n represents the sample size (if all students at intervention schools are present and participate in the study)
- Z represents the standard normal distribution
- σ represents the within group standard deviation
- α represents the probability of type I error, taken here as 0.05

- β represent the probability of type II error, taken here as 0.20
- Δ represents the relevant difference in SLIPS-GFP score between M1 and M2 that qualifies as statistically significant change, taken here as 3.52

$$7.9 \times 11.81^2 / (3.52^2) = n$$

$$n = 88.93$$

Eighty-nine participants are required in each group to be 80% sure of being able to detect a significant change in MH score of 5% at the 95% significance level. Based on experience of working in northern Sri Lanka, I anticipated a rate of 10% loss to follow up between M1 and M2. As result, I increased the sample size required by 10% to 97 participants per study group.

The results of the 2014 Kilinochchi study also indicated that girls and boys had significantly different MH scores at baseline. I decided to stratify by sex, assuming a 1:1 ratio of boys to girls in Mullaitivu district. This gives a final sample size of 194 participants in the intervention and control group with a 50% gender split.

6.3.3 Recruitment

According to the CDO project lead, school selection was based on two criteria: (1) to reflect the different ethnic groups present in Mullaitivu (2) to be schools that were severely affected during the civil conflict. There was only one Sinhalese school that was of sufficient size and geographic proximity to be feasible for inclusion in the study. As there is no detailed mapping of Mullaitivu, the approximate location of the schools is outlined in FIGURE 6.1 below:



FIGURE 6-1: Map of Mullaitivu town and surrounding division with intervention and control schools marked on (scale 1cm = 2km)

TABLE 6-1: Identifiers for map

Number	School	Code	Intervention [V]/control status [C]	Number of children in Grade 8 & 9
1	Ampalavanpokhanai	AMB	[V]	46
2	Mullivaikal West	MVK	[V]	35
3	Vattappalai	VAT	[C]	104
4	Thanneerroottu	THN	[V]	76
5	Alampil	ALM	[C]	80
6	Chemmal	CHM	[C]	46
7	Kiribanwewa	KRB	[C]	21
8	Sampathnuwara	SAM	[V]	72

All schools contacted agreed to take part in the study. I made personal visits to each school Principal with signed letters from the Provincial, District and Divisional Directors of Education. A translator outlined the aim of the study and a data collection date was agreed. I took the personal contact number of the principal to allow them to contact me with any concerns and so that I could provide a reminder two days before data collection commenced.

Participants were recruited through their school. Participation was voluntary and subjects were free to drop out at any time. All students in attendance were invited to take part in the study.

6.3.4 Identification and tracking

All students in years 8 and 9 were eligible for inclusion in the study. Participants provided basic demographic data to allow for tracking over time (SLIPS-GFP form in attached in APPENDIX C.3). Data were coded, encrypted and stored on a password-protected hard drive. Hard copies were kept in a locked cabinet, separate from identifiers.

6.3.5 Outcome variables

The outcome variables were the same as those included in the test-retest reliability analysis: the final 22-item SLIPS-GFP, MFT, BFA and HFA (APPENDIX C3-C5).

6.3.6 Data collection

The data collection procedure was identical to that outlined in Chapter 5. Additionally, I took photos to visually assess the quality of the school facilities and the surrounding area.

Measurement took one day per school and commenced at 7:45 am. Where possible, we revisited schools to collect data on individuals absent from school on the collection date.

6.3.7 Data analysis

Statistical package for the Social Sciences 17 (SPSS) was used to enter and analyse data ^[217]. All data were cleaned and significant outliers were removed prior to analysis. Descriptive statistics (mean, SD and 95% CIs) were calculated for all outcome measures. All assumptions underlying statistical tests were verified prior to commencing analysis (APPENDIX D.1).

The data analysis was split into three sections. In many cases normality assumptions were violated. In this instance, I used a conservative statistical test, which remain robust despite normality assumptions [218]. In addition, I performed non-parametric tests to support the results of the ANOVA.

1. Sample distribution:

This section examines the composition of the sample. To analyse differences in performance between groups, I used ANOVA, which requires that samples are normally distributed and have HOV [219].

2. Comparison with global norms:

In order to compare the nutritional status of participants with global norms, height and weight were converted into BFA and HFA Z scores. The WHO provides software online for the calculation of BFA and HFA Z scores [220]. Global references and thresholds are available, against which the prevalence of thinness and stunting in the test sample can be compared [189]. There are no statistical assumptions behind these comparisons.

A previous meta-analysis of multiple international MFT studies has provided global normative values [191]. In this meta-analysis, performance is reported as the speed at the last completed level of the MFT. I converted the Mullaitivu beep test scores into the same metric to allow for comparison of the samples. The samples were compared using an ANOVA.

3. Association between outcome variables:

I assessed the association between the primary outcome variables in order to determine the ability of HFA, BFA and fitness level to predict psychosocial health score. In addition, I assessed whether BMI, HFA and BFA are associated with MFT score. This was to assess whether proxy indicators of acute and chronic nutritional status played an important role in the fitness of adolescents in Mullaitivu. Association was initially assessed using scatterplots. In the event that a linear relationship was identified by plotting the data, simple linear regression was used to quantify the association between the independent variable and the dependent variable [219].

6.3.8 Assumptions

The statistical methods used in this study assume that the samples were normally distributed and had HOV. Normality was initially assessed using histograms for the crude results of each of the MH and PF variables. Initial assessment warranted further tests of normality and statistical tests were also completed. HOV was assessed using the Levene's Test ^[219].

The use of Pearson's R assumes that the samples being compared have homoscedasticity and the residuals are normally distributed. These assumptions were assessed by plotting the regression-standardised residuals against the dependent variable, and constructing a Probability-Probability (PP) plot of the regression standardised residuals for the independent and dependent variables.

6.4 Results

6.5 Sample Distribution

383 out of a possible total of 480 students were present on the day of data collection. This is an absence rate of 20.1%, which was higher than predicted. 367 participants (76.5% of total eligible sample) were included in the final analysis, see diagram for flow of participants through study. The biggest cause of missing data was absenteeism, followed by those individuals who dropped out of the beep test without completing a single level. Intervention schools are denoted with a [V] and control schools a [C].

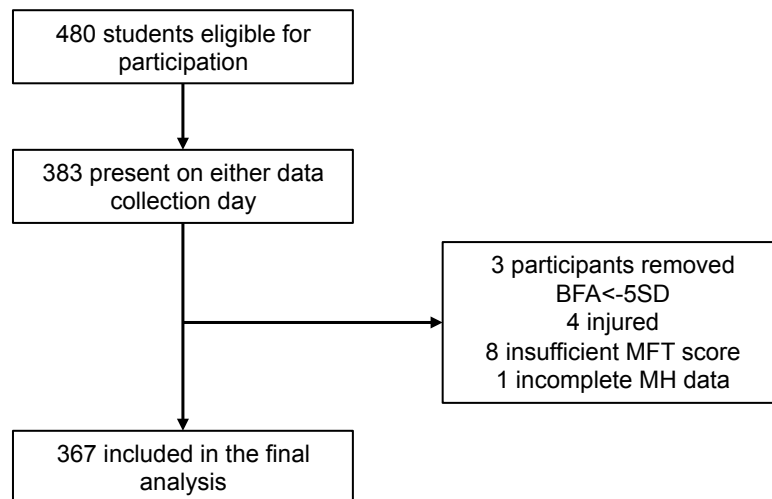


FIGURE 6-2: Study summary diagram

The total sample is evenly distributed across the sexes (45% female) and 96.4% of the sample is either 13 or 14 years old. Girls and boys were both on average 13.50 years old.

TABLE 6-2: Number of participants in each school as an absolute figure and as percentages of the entire sample, stratified by sex (n = 367)

Ethnic Group	School	Gender	(n)	% of total sample
Tamil	Ambalavanpokhanai	B	18	5%
		G	21	5.7%
	Mullivaikal	B	14	3.8%
		G	16	4.3%
	Alampil	B	35	9.5%
		G	30	8.2%
	Vattapali	B	37	10.1%
		G	32	8.7%
	Chemmalai	B	28	7.6%
		G	13	3.5%
Sinhala	Sampathnuwara	B	32	8.7%
		G	22	6.0%
	Kiribanwewa	B	3	0.82%
		G	9	2.5%
Muslim	Thanneerroottu	B	34	9.3%
		G	23	6.3%
Total			367	100%

Schools are segregated into ethnic groups, and the Tamil ethnic group are over-represented in the overall sample. 66.5% of the entire sample is Tamil, 18.02% is Sinhala and 15.6% is Muslim.

When the sample is stratified in [V] and [C] cohorts, it is evident that the gender balance remains similar (TABLE 6.3 outlines the breakdown of the [V] and [C] groups by sex), 46.9% of the [V] group, and 43.6% of the [C] group are female.

TABLE 6-3: Number of participants in [V] and [C] groups, stratified by gender (n = 367)

Independent variable		Intervention [V]		Control [C]	
		n	% of [V] sample	n	% of [C] sample
Gender	B	95	53.1%	106	56.4%
	G	84	46.9%	82	43.6%
	Total	179	100	188	100

TABLE 6-4: Number of participants in the [V] and [C] groups stratified by ethnic group (n = 367)

Independent variable		Intervention [V]		Control [C]	
		n	% of [V] sample	n	% of [C] sample
Ethnicity	Tamil	68	37.9%	176	93.6%
	Sinhala	54	30.2%	12	6.4%
	Muslim	57	31.8%	0	0%
n		179	100	188	100

The [V] group is well distributed across the ethnic groups (37.9% Tamil, 30.2% Sinhala and 31.8% Muslim). The control sample is 93.6% Tamil and 6.4% Sinhalese, and there are no Muslim participants in the control group.

6.6 Descriptive statistics

Descriptive statistics were calculated for all the primary and secondary outcome variables. The sample was initially stratified by gender to see if there were any gender-based disparities. After this, the sample was stratified into the [V] and [C] groups to identify if there was any significant differences between in the outcome scores.

TABLE 6-5: M1 mean and 95% CIs for all major outcome variables (n = 367)

Boys-girls	N		SLIPS-GFP	INT	EXT	WTH	MFT	BFA	HFA
Boys	201	Mean (95% CIs)	26.13 (24.90 - 27.38)	14.54 (13.90 - 15.19)	6.37 (5.86 - 6.88)	4.02 (3.70 - 4.34)	5.58 (5.37 - 5.80)	-2.10 (-2.25 - 1.96)	-2.13 (-2.30 - 1.97)
		SD	8.9	4.64	3.64	2.32	1.56	1.02	1.14
Girls	166	Mean (95% CIs)	23.3* (22.03 - 24.60)	13.42* (12.72 - 14.11)	4.89* (4.35 - 5.42)	4.04 (3.71 - 4.37)	3.74* (3.56 - 3.91)	-1.50* (-1.64 - 1.35)	-1.32* (-1.50 - 1.12)
		SD	8.4	4.56	3.51	2.16	1.23	0.935	1.24

NB: * denotes a significant difference between boys and girls to the 0.05 significance level.

There were significant gendered differences in scores. Girls outperform boys in the mental health testing (SLIPS-GFP score $F = 9.63$, $p = 0.002$) and have significantly better anthropometric indicators (BFA: $F = 33.94$, $p = <0.001$; HFA: $F = 43.51$, $p = <0.001$). However, boys outperformed girls in the MFT ($F = 163.36$, $p = <0.001$).

TABLE 6-6: Mean, 95% confidence intervals and standard deviation of all outcome variables at M1, stratified by intervention/control status (n = 367)

[V]/[C]	N		SLIPS-GFP	INT	EXT	WTH	MFT	BFA	HFA
V	179	Mean (95% CIs)	24.65 (23.42 - 25.89)	13.98 (13.31 - 14.66)	5.45 (4.99 - 5.90)	4.06 (3.72 - 4.39)	4.51 (4.29 - 4.74)	-1.82 (-1.97 - 1.67)	-1.74 (-1.84 - 1.55)
		SD	8.38	4.57	3.10	2.28	1.53	1.05	1.25
C	188	Mean (95% CIs)	25.06 (23.74 - 26.38)	14.08 (13.40 - 14.75)	5.94 (5.34 - 6.53)	4 (3.68 - 4.32)	4.98* (4.72 - 5.23)	-1.83 (-1.97 - 1.69)	-1.78 (-1.97 - 1.60)
		SD	9.16	4.69	4.14	2.22	1.74	1.01	1.26

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

The control subjects outperformed the intervention group in the MFT ($F(1, 365) = 7.28, p < 0.05$). There were variable differences between groups at M1 when comparing stratified groups according to age, gender and ethnic group (APPENDIX D.1). School level analyses indicated that there were school level differences in MH, MFT and BFA performance. The small sample size meant that any analyses at the school level are underpowered, and excluded any possible multi-level analysis. School level data are outlined in APPENDIX D1. To control for these differences in subsequent analyses, data were stratified according to intervention/control status and a model that adjusted for age, gender and ethnicity. Of particular interest is that Muslim (mean = 3.97, +/-0.352) children significant underperform in the MFT when compared to Tamil children (4.88, +/- 0.21 ($F(1, 299) = 14.74, p < 0.001$)).

6.7 Comparison to global norms

This section compares the PH performance of adolescents in Mullaitivu to global normative scores. I have excluded the results for 12 year olds and 15 year olds from the analysis, as it is likely that they have misreported their age during the test.

6.7.1 Descriptive statistics

The proportions and 95% confidence intervals of the Mullaitivu sample in each HFA and BFA category were calculated. The expected global norms were drawn from the WHO global reference population (TABLES 6.7 and 6.8).

TABLE 6-7: Proportions and 95% confidence intervals of the Mullaitivu sample at M1 in each of the HFA categories according to the WHO global normative values, stratified by age and gender (n = 354)

	Boys		Girls		Global norm
	13	14	13	14	
Age	13	14	13	14	-
N	110	82	77	85	-
Severely Stunted	16.4 9 <-> 23.7	19.5 10.3 <-> 28.7	9.1 2 <-> 16.2	4.7 0 <-> 9.8	0.14
Stunted	61.8 52.3 <-> 71.4	58.5 47.3 <-> 69.8	42.9 31.2 <-> 54.6	22.4 12.9 <-> 31.8	2.14
Normal height	20.9	21.1	48	72.9	81.86
Tall	0.9 0 <-> 3.1	0.9 0 <-> 3.1	0 0 <-> 0.6	0 0 <-> 0.6	13.59
>+2SD	0 0 <-> 0.5	0 0 <-> 0.5	0 0 <-> 0.6	0 0 <-> 0.6	2.28

NB: I discuss the potential inappropriateness of the WHO cut offs for assessing BMI in a group of Sri Lankan adolescents in section 6.10.1, as the WHO cut offs may overestimate levels of malnutrition and stunting.

TABLE 6-8: Proportions and 95% confidence intervals of the Mullaitivu sample in each of the BFA categories according to the WHO global normative values, stratified by age and gender (n = 354)

	Boys	Boys	Girls	Girls	Global norm
Age	13	14	13	14	-
N	110	82	77	85	-
Severely thin	22.7 14.4 <-> 31	28 17.7 <-> 38.4	9.1 2 <-> 16.2	4.7 0 <-> 9.8	0.14
Thin	55.5 45.7 <-> 65.2	61 49.8 <-> 72.1	35.1 23.8 <-> 46.4	22.4 12.9 <-> 31.8	2.14
Healthy weight	20.9	11	51.9	68.2	81.86
Overweight	0.9 0 <-> 3.1	0 0 <-> 0.6	3.9 0 <-> 8.9	4.7 0 <-> 9.8	13.59
Obese	0 0 <-> 0.5	0 0 <-> 0.6	1.3 0 <-> 4.5	0 0 <-> 0.6	2.28

6.7.2 Statistical findings

The results indicate that there is a high prevalence of stunting and thinness among adolescents in Mullaitivu when compared to global normative values. Boys and girls also underperform in the MFT when compared to global normative scores.

There are no assumptions to be verified when comparing proportions. When considering HFA, adolescents in Mullaitivu have a high prevalence of stunting (boys aged 13: 61.8% and boys aged 14: 58.5%; girls aged 13: 42.9% and girls aged 14: 22.4%) and severe stunting (boys aged 13: 16.4% and boys aged 14: 19.5%; girls aged 13: 9.1% and girls aged 14: 4.7%). See FIGURE 6.3 and FIGURE 6.4 for graphical representation of results.

When considering BFA, both boys and girls have a substantial prevalence of thinness (boys aged 13: 55.5% and boys aged 14: 61%; girls aged 13: 35.1% and girls aged 14: 9.1%) and severe thinness (boys aged 13: 22.7% and boys aged 14: 28%; girls aged 13: 9.1% and girls aged 14:

4.7%). There is very low prevalence of overweight or obese in either sample. See FIGURE 6.5 and 6.6 for graphical representation of results.

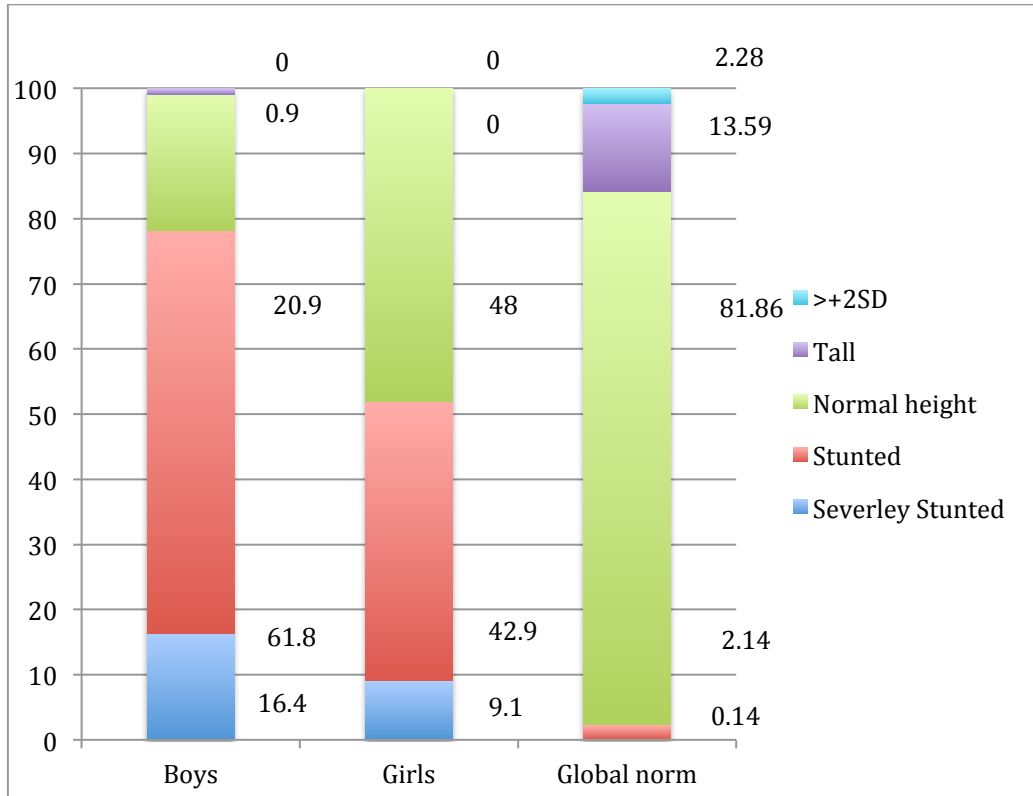


FIGURE 6-3: Graph comparing the HFA Z proportions of 13 year olds in Mullaitivu at M1 compared to global normative values (n = 187)

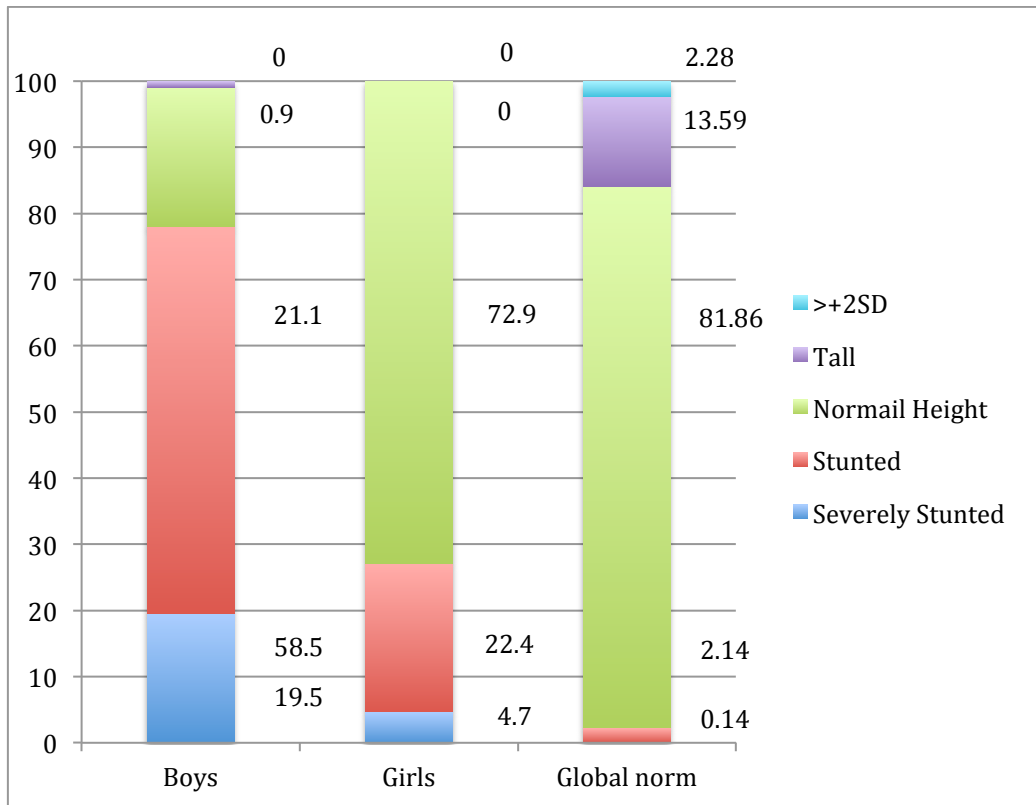


FIGURE 6-4: Graph comparing HFA Z score proportions of 14 year olds in Mullaitivu at M1 and global normative values (n = 167)

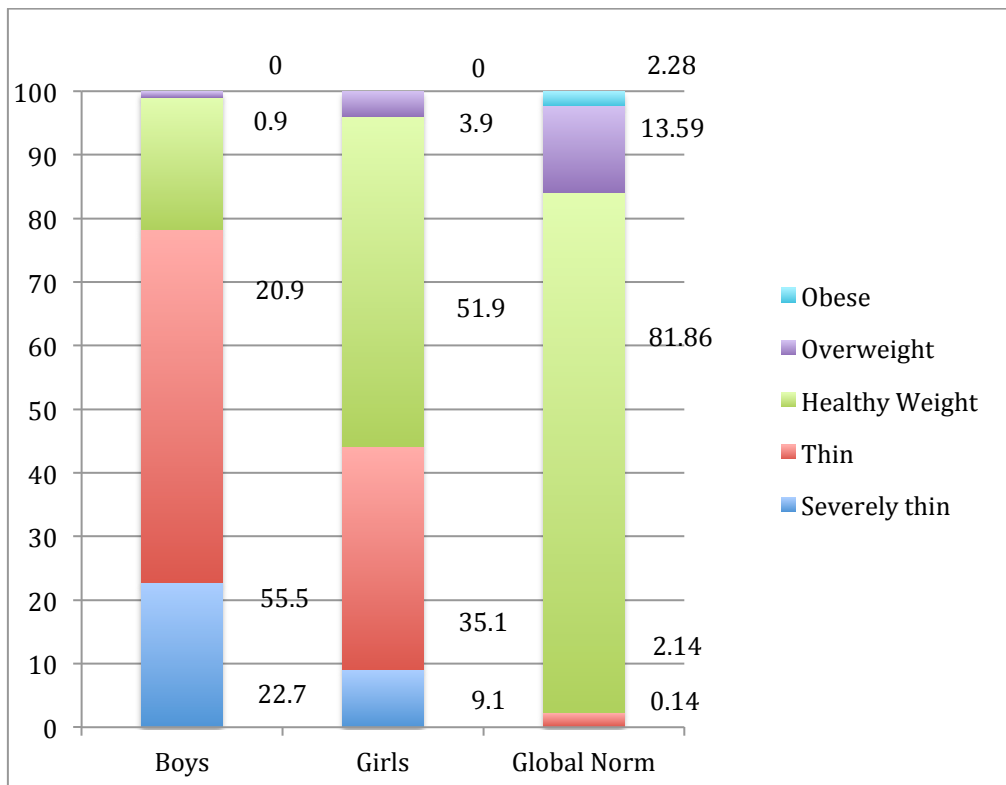


FIGURE 6-5: Graph comparing the BFA Z score proportions of 13 year olds in Mullaitivu at M1 with global normative values (n = 187)

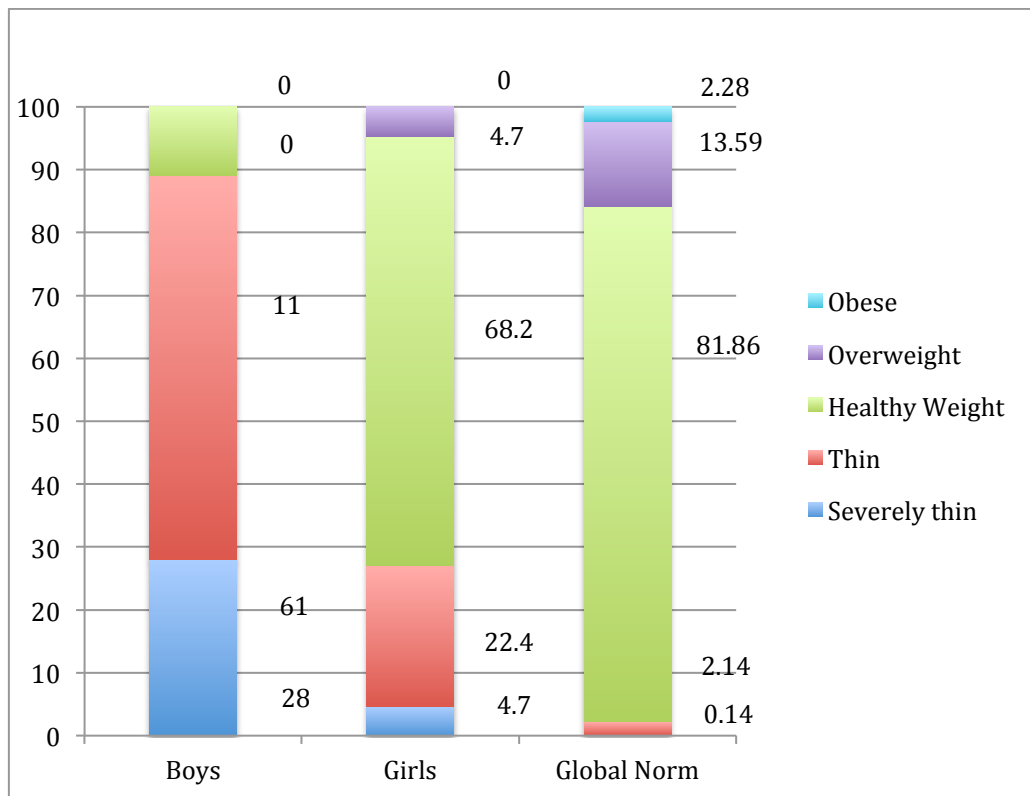


FIGURE 6-6: Graph comparing the BFA Z proportions of 14 year olds in Mullaitivu at M1 with global normative values (n = 167)

The mean and 95% confidence intervals were calculated for the MFT and compared to global normative values (TABLE 6.9). It is evident that there is no overlap in the 95% confidence intervals in TABLE 6.9, and also FIGURES 6.7 and 6.8. Boys and girls both underperform compared to global averages.

TABLE 6-9: MFT scores in Mullaitivu at M1 compared with global normative values, stratified by gender and age (n = 354)

Sex	Age	n	Mullaitivu MFT	Global norms
Boys	13	110	9.96 <i>9.80 - 10.12</i>	10.95 10.94 - 10.97
	14	82	10.07 <i>9.88 - 10.25</i>	11.16 11.151 - 11.81
Girls	13	77	8.94 <i>8.77 - 9.10</i>	10.22 10.21 - 10.15
	14	85	9.11 <i>8.95 - 9.28</i>	10.31 10.30 - 10.32

NB: Bold and italic indicates a significant difference between the Mullaitivu MFT score and the global norms

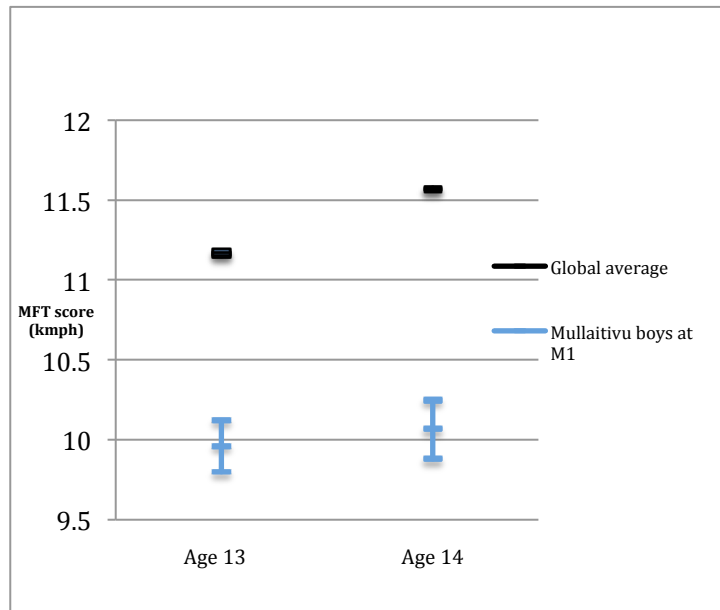


FIGURE 6-7: Graph of Mullaitivu boys MFT score at M1 compared to global averages (n 13 = 110, n 14 = 82)

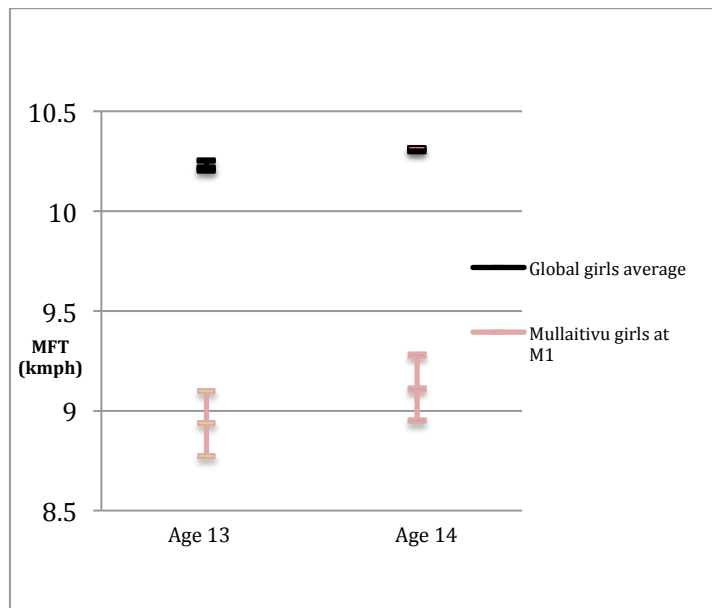


FIGURE 6-8: Graph of Mullaitivu girls MFT scores at M1 compared with global averages (n 13 = 77, n 14 = 85)

6.8 Comparison with adolescents in Kilinochchi

This section compares the MH and PH performance of the Mullaitivu sample at M1 with adolescents in the neighbouring district of Kilinochchi.

6.8.1 Descriptive statistics

The proportions and 95% confidence intervals of the 13 -year-old Mullaitivu sample in each HFA and BFA category were calculated and compared to the Kilinochchi results (TABLE 6.10 and 6.11). The results are displayed graphically in FIGURES 6.9 and 6.10. Mean and 95% confidence intervals were calculated for the SLIPS-GFP and MFT for the Mullaitivu sample and compared to the Kilinochchi results (TABLE 6.12).

TABLE 6-10: Proportions and 95% confidence intervals of 13-year-olds in Mullaitivu and Kilinochchi in each of the BFA categories according to the WHO global normative values, stratified gender (n = 631)

	Mullaitivu		Kilinochchi	
	Boys	Girls	Boys	Girls
Age	13	13	13	13
N	110	77	221	223
Severely thin	22.7 <i>14.4 <-> 31</i>	9.1 2 <-> 16.2	10 <i>5.7 <-> 14.2</i>	4 1.2 <-> 6.8
Thin	55.5 <i>45.7 <-> 65.2</i>	35.1 23.8 <-> 46.4	31.8 <i>25.2 <-> 38.3</i>	15.2 10.3 <-> 20.2
Healthy weight	20.9	51.9	56.3	75
Overweight	0.9 0 <-> 3.1	3.9 0 <-> 8.9	1.9 0 - 4	5.4 2.2 - 8.6
Obese	0 0 <-> 0.5	1.3 0 <-> 4.5	0 -0 <-> 0.5	0 0 - 0.4

NB: Bold and italics indicate a likely significant difference between the rates of thinness between Kilinochchi and Mullaitivu, as the 95% confidence intervals do no overlap.

TABLE 6-11: Proportions and 95% confidence intervals of 13-year-olds in Mullaitivu and Kilinochchi in each of the HFA categories according to the WHO global normative values, stratified gender (n = 631)

	Mullaitivu		Kilinochchi	
	Boys	Girls	Boys	Girls
Age	13	13	13	13
N	110	77	221	223
Severely stunted	16.4 <i>9 <-> 23.7</i>	9.1 2 <-> 16.2	3.8 <i>1 <-> 6.6</i>	3.1 0.6 <-> 5.7
Stunted	61.8 <i>52.3 <-> 71.4</i>	42.9 <i>31.2 <-> 54.6</i>	36.5 <i>29.8 <-> 43.2</i>	17.9 <i>12.7 <-> 23.2</i>
Healthy height	20.9	48	59.7	79
2SD < HFA < 3SD	0.9 0 <-> 3.1	0 0 <-> 0.6	0 0 <-> 0.2	0 0 <-> 0.2

NB: Bold and italics indicate a likely significant difference between the rates of stunting between Kilinochchi and Mullaitivu, as the 95% confidence intervals do not overlap.

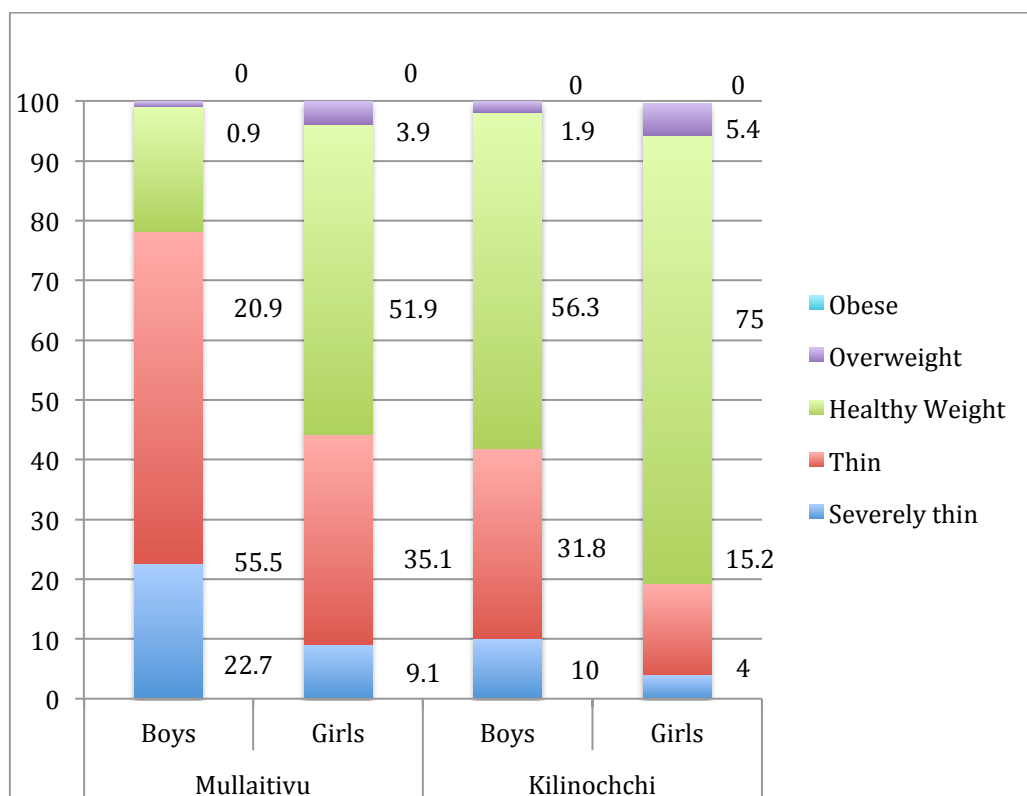


FIGURE 6-9: Graph comparing the BFA Z score proportions in 13 year olds Mullaitivu at M1 with Kilinochchi values (n= 631)

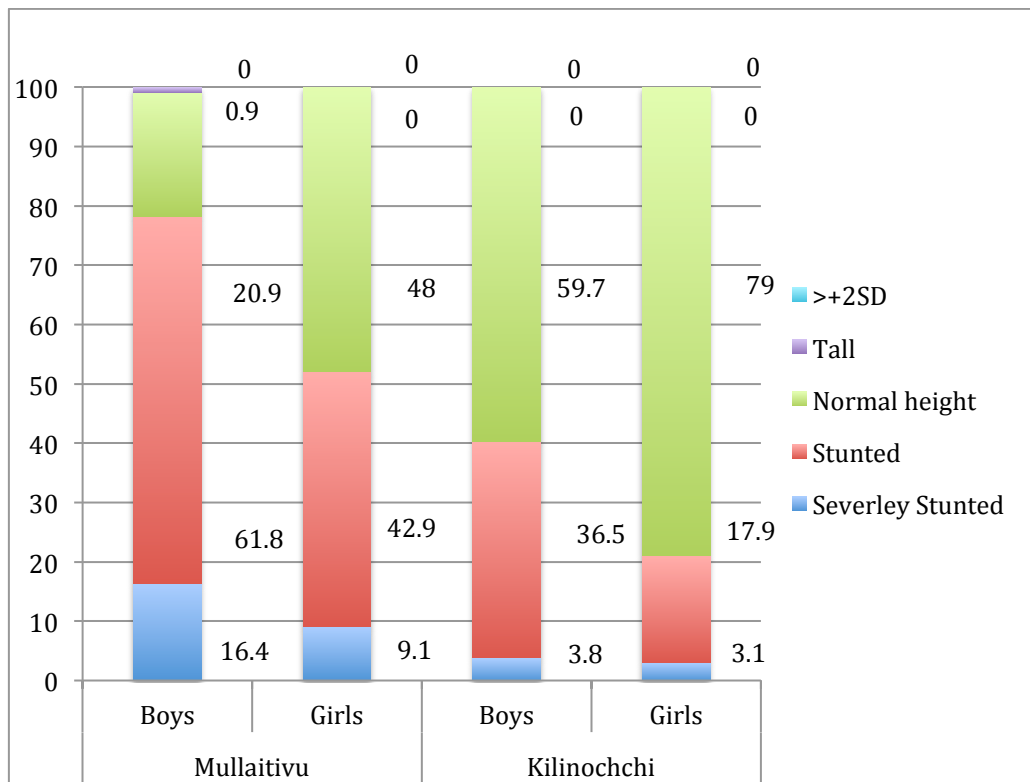


FIGURE 6-10: Graph comparing the HFA Z score proportions in 13 years olds in Mullaitivu at M1 with Kilinochchi values (n = 631)

TABLE 6-12: Mean and 95% confidence intervals for SLIPS-GFP and MFT in Mullaitivu at M1 and Kilinochchi (n = 828)

Location	N	Outcome	Mean
Mullaitivu	367	SLIPS-GFP	24.86
			23.96 - 25.76
		MFT	4.74
			4.59 - 4.92
Kilinochchi	461	SLIPS-GFP	27.30
			26.22 - 28.38
		MFT	4.79
			4.64-4.95

NB: Bold and italics indicate a possible significant difference between Kilinochchi and Mullaitivu SLIPS-GFP scores

6.8.2 Statistical findings

Thirteen-year-old boys in Mullaitivu have a significantly higher prevalence of thinness (Mullaitivu boys: 55.5%; Kilinochchi boys: 31.8%) and severe thinness (Mullaitivu boys: 22.7%; Kilinochchi boys: 10%) than boys in Kilinochchi. Mullaitivu boys have a significantly higher prevalence of stunting (Mullaitivu boy: 61.8%; Kilinochchi boys: 36.5%) and severe stunting (Mullaitivu boys: 16.4%; Kilinochchi boys: 3.8%). This relationship holds for girls too, in which adolescents from Mullaitivu have significantly higher rates of stunting and thinness than adolescents in Kilinochchi. This relationship does not hold for the prevalence of severe thinness and severe stunting, in which the 95 CIs overlap among girls. Both samples have very low proportions of overweight individuals (all samples <6%).

An ANOVA was used to assess whether there were any significant differences between the SLIPS-GFP score and MFT of adolescents in Mullaitivu and Kilinochchi. The assumptions underlying the use of an ANOVA are discussed in APPENDIX D.3.

Adolescents in Kilinochchi scored higher on the SLIPS-GFP than Mullaitivu adolescents ($F(2) = 10.85, p=0.01$). There was no significant difference in the MFT score between adolescents in Mullaitivu and Kilinochchi ($F(2) = 0.132, p = 0.717$). As the SLIPS-GFP samples did not have HOV, I used a Mann Whitney U test to verify that there were a significant difference between SLIPS-GFP scores in Mullaitivu and Kilinochchi ($U(2) = -2.764, p=0.006$).

TABLE 6-13: Levene's test for HOV and ANOVA results for Mullaitivu and Kilinochchi (df=1)

Outcome	Levene's	sig.	F	sig.
SLIPS-GFP	38.24	<0.001	10.85	0.01
MFT	0.002	0.962	0.132	0.717

6.9 Association between outcome variables

I assessed the association between the physical and MH outcome variables in the Mullaitivu sample at M1. The scatterplots in APPENDIX D.4 indicate that there were no significant relationships between any of the physical and MH variables in any of the subgroups (APPENDIX D.4). The association between PH variables was initially assessed using scatterplots.

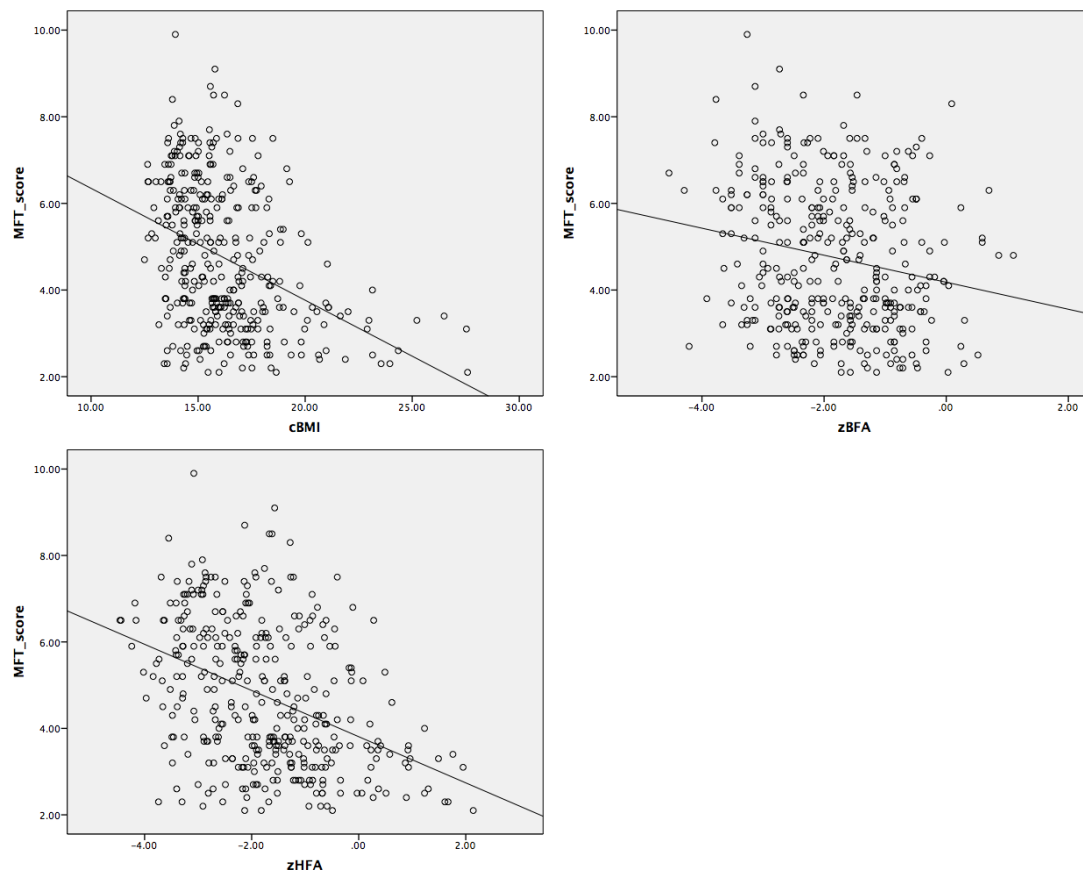


FIGURE 6-11: Scatterplots plotting BMI, BFA and HFA against MFT score in Mullaitivu at M1 (n = 367)

The scatterplots indicate an association between BMI and MFT score and for HFA and MFT score. I plotted scatterplots for the age, sex and ethnic subgroups and the association is similar.

A standard linear regression was performed to assess the ability of HFA to predict MFT score.

The assumptions for using Pearson's R were verified. Using the enter method, I found that HFA predicted a significant amount of variation in MFT score ($F(1, 365) = 71.07, p < 0.05, R^2_{adjusted} = 0.161$). This indicates that HFA explained 16.1% of the variation in MFT scores.

6.10 Discussion

The principal aim of this chapter was to assess the baseline performance of participants in Mullaitivu and, where possible, to contextualise the baseline scores by comparing them with global and regional equivalents. In doing so, this chapter aimed to verify some of the assumptions outlined in the theoretical framework in CHAPTER 4.

The sample is predominantly male and Tamil ethnic origin, reflecting the makeup of the regional population. As a result, there are almost no controls for either Sinhalese or Muslim students. There were no differences in the performance of the intervention and control groups in any of the outcome variables except the MFT, in which the control group outperformed the intervention group. There were variable differences in outcomes when the sample was stratified by gender or ethnicity. The Muslim population underperformed significantly in the MFT when compared to the Tamil population. This may reflect a fitness deficit in the Muslim population, but it is of note that the conditions in which the MFT was taken were by far the worst in the Muslim school. Similarly, Muslim girls wore full-length clothing that appeared to obstruct their running and may have contributed to the poor results among this ethnic group (FIGURE 6.12).



FIGURE 6-12: Girls at Thanneeroottu School running on loose ground, barefoot and in full-length clothes. Photo taken by A. Hamilton in June 2015

I was unable to make any international MH comparisons, as the SLIPS-GFP is an emic measure. However, adolescents in Mullaitivu performed better than their counterparts in Kilinochchi. Boys had poorer MH than girls, which is reflective of the Kilinochchi results.

The results indicate a PH deficit in the sample population. Adolescents in the sample have a very high prevalence of stunting and wasting when compared to global normative values. 80% of boys are either stunted or severely stunted and 90% of 14-year-old boys are thin or severely thin, when the proportions predicted by the WHO are 2.28% for both categories.

Studies assessing BMI in Sri Lankan adolescents have had mixed results. According to a small study conducted by Ramanayaka et al (2013), prevalence of adolescent underweight status in Mullaitivu was 34%, whereas the Sri Lankan national average figure is 17.7% [221]. Hettiarachchi et al. (2006) assessed thinness and stunting prevalence in Galle; 37.3% of grade 8 boys and 20.9% of girls met the criteria for wasting (weight for age Z score > -2SD), and 25.3% of boys and 19.1% of girls were stunted [222]. Jayatissa and Randbanda (2006) conducted a large scale, representative study of BFA and HFA in Sri Lanka. They recorded rates of underweight in 13 year-old boys and girls as 66.5% and 26.4% respectively; prevalence of stunting was 41.3% and 40.1% [223]. The results of the baseline analysis support previous evidence that malnutrition and stunting is prevalent in Sri Lanka, and particularly in northern regions.

Adolescents in Mullaitivu underperform when compared to other South Asian countries. Within South Asia, there is considerable regional variation, and notable rural-urban differences in performance. The performance of adolescents in Mullaitivu is most similar to that of rural Bangladeshi students (Mullaitivu: stunting 67.4%, thinness 64.2%, Bangladesh: stunting 48%, thinness 67%, Myanmar: stunting 39%, thinness 32%) [224]. Thinness and stunting are based on percentiles in these studies, rather than the WHO HFA and BFA categories, which only allows for broad comparisons.

BFA and HFA studies in South Asia report that girls perform worse on average than boys [224, 225]. The Mullaitivu data are not consistent with this trend, as girls performed better than boys across all categories. A Sri Lankan nationally representative study of stunting and wasting identified a

similar trend [223]. The high prevalence of underweight, and that boys perform worse than girls, might be a result of a sexually determined differential effect on the delayed onset of puberty.

Mullaitivu girls and boys perform worse than adolescents in Kilinochchi in all BFA and HFA categories. Mullaitivu is a predominantly rural district and the Kilinochchi sample was taken from a semi-urban area. The semi-urban nature of Kilinochchi town may have a protective effect, although this is counter intuitive to the typical Western experience. The previously mentioned nationally representative study identified a similar trend in Sri Lanka: a protective effect of urban living on malnutrition status, which, according to the authors, may be the result of variable coverage of food programmes [223].

It is perhaps unsurprising that adolescents in Mullaitivu show high proportions of stunting given the socio-political history of the region. 13 and 14-year-olds would have been 7 or 8 years of age at the end of the civil conflict, and are likely to have suffered multiple displacements throughout the period of 2005-2009. These results are reflected in the wider literature, which suggest that exposure to conflict at a young age has a causative and permanent effect on rates of stunting [226].

Adolescents in Mullaitivu demonstrate a significant fitness deficit, with both sexes on average underperforming by two levels of the MFT compared to global normative scores. They do not, however, perform any worse than adolescents in Kilinochchi. It is of interest that neither sex improved their fitness performance with age, as is seen in global norms [191].

The use of an emic measure of MH limits the external validity of the findings. Adolescents performed significantly better than adolescents in Kilinochchi, despite their rural and lower-income status. This difference could also be as a result of systematic error. In Kilinochchi the 22 SLIPS-GFP items were extracted from the 49-item SLIPS-C. The different layout and length of the questionnaire may have significantly affected the answering behaviour of participants. These differences should be read with caution. Generally, rural living status is associated with poorer MH outcomes [33].

There was no association between PH scores and MH outcomes. Fitness and level of thinness or stunting do not predict MH in this context. HFA predicts a significant amount of the variation in MFT performance ($R_{adjusted} = 0.16$), as HFA increases MFT performance decreases.

6.10.1 Weaknesses:

There were some considerable methodological challenges to operating in northern Sri Lanka that impacted on the study design and implementation. As the study is “piggybacking” a pre-designed intervention, the sample size and composition were pre-determined. This influences the capability to sample effectively, and the statistical power of analyses. There are several changes that could be made to improve the representativeness of the study. Ideally, the study would have randomised at the school level and have used probability proportional to size to identify students for measurement [203].

There were several environmental and political idiosyncrasies that were difficult to control for, which may have affected the quality of the data collection and reliability of the results. The quality and size of the schools in which I was working was variable. Many schools were isolated and in deprived rural areas. As a result there was variation in the quality of the facilities, behaviour and attitude of the participants.

I experienced difficulty in accessing THN School in the morning, as I had with the other schools. This meant that the temperature was near 40 degrees centigrade when participants took the MFT, many dropped out early. Access to schools is at the discretion of the School Principal. According to informal interviews, there are competing political parties within the Muslim community in Mullaitivu. Local political parties put pressure on principals to toe the line. Allowing research during school time, especially near exams, is considered a poor use of time. This limits access to participants during formal school hours and many eligible students leave school before the testing takes place.

One of the principal weaknesses in this study was the amount of absenteeism. Informal conversations with teachers and the data collection team revealed that children often do not attend school if they think that it will rain that day. The explanation for this was that children

would get wet on their way to school and consequently fall ill. Further, children and young adolescents often help out with the planting and harvesting of their parents paddy fields; although this wasn't of major concern in March, There may have been significant differences in the outcome variables between those children who participated and those who were absent on the day of collection.

I had a limited ability to communicate effectively with my research team, primarily due to language. A lack of research experience in the volunteer team often meant that speed was given priority over precision and exactitude. My inability to communicate effectively with the team one-to-one meant that it was hard to harness teachable moments effectively. The staff members that had proficient English language skills were often too busy to participate and involved in several projects at one time. Since their skills were needed to communicate the plan to the research team, this reduced my degree of control. All communications had to go through a conduit that inherently mixed messages, plans and outcomes.

The WHO BFA and HFA cut offs may be inappropriate for this sample, as they may overestimate the level of thinness and stunting ^[227]. The cut offs may be doubly inappropriate because the population in question is both South Asian and living in post-conflict setting. Epigenetics factors may further constrain the applicability of the WHO cut offs, as Rannan-Eliya et al. (2013) found a strong relationship between maternal and child height ^[228]. Although this is not a suitable justification for the very high prevalence of under nutrition and stunting observed in Mullaitivu. I have also calculated the number of participants in each weight and height category using the International Obesity Taskforce; the results are similar and indicate very high levels of stunting and malnutrition in the sample.

6.10.2 Strengths

The study indicates that 'rigorous pragmatism' in study design and implementation can produce worthwhile insight into mental and PH in a low-resource, post-conflict context. I assessed clearly defined, measurable MH outcomes that had been pre-tested, refined and validated. In addition, these outcomes were locally relevant and contributed to the local ownership of the evaluation.

In Sport for Development and Peacebuilding research, theoretical frameworks, assumptions and models often remain poorly defined and un-tested ^[212]. This study outlined a theoretical framework, and then systematically tested assumptions behind the model, using public health methods to ensure methodological rigour and valid results. According to the Organization for Economic Cooperation and Development (OECD), *“90% of evaluation done on peacebuilding programmes reasserts assumptions...therefore [practitioners] must make theory explicit, not reassert pet theories and systematically test theories underlying programmes”* (OECD: 2007). In outlining and testing a feasible underlying theory I am contributing to the limited evidence base for SDP.

To my knowledge, it is the first study to assess the difference in MH among the different ethnic groups in northern Sri Lanka, in a particularly remote and isolated region. Similarly, it provides a stark insight into the state of undernutrition in a neglected corner of Sri Lanka.

In some cases, needs assessments are often poorly implemented, typically relying on assumed deficits in a recipient population without true consultation of local partners. Subsequent interventions are then based on value judgements, rather than theory and systematic testing. Both poles are unsatisfactory, true dialogue in the development of an evaluation will help both parties to ensure validity and relevance of their results.

Knowledge that the current sample would be underpowered might have provided more leverage to argue for an increased sample size, along with other improvements to the intervention and the evaluation design. A larger sample would allow for further stratification of the sample and more detailed analyses, which is important, as there were significant differences in performance when comparing the various subgroups.

6.10.3 Implications for study:

The results of the baseline study have identified factors that must be controlled for in any further analyses. Similarly, the results indicate that further work should increase the sample size to allow for stratification along age, gender and ethnic lines. The baseline study has highlighted some practical methodological issues that required attention in the follow up data collection

(M2). Particularly that absenteeism was high in this area, despite the advice of District Officials, and further work should expect a rate of absenteeism as high as 20%. To account for this I made sure to visit each school at least twice to ensure maximum participation.

This baseline study indicated that it is both important and feasible to test commonly held assumptions before the beginning of an intervention. The results indicate that there is in fact no association between PF and the MH in this particular population. The intervention should focus on the relational elements, as improving PF alone is unlikely to have any effect on the MH of the participants. Unfortunately, due to time constraints, I was unable to adjust the theoretical framework for the study to consider other potential influential intervention elements and measure them.

With the realisation that the use of an emic measure of MH inherently limits the external validity of any conclusions, I decided to introduce another MH outcome. I chose to implement the measure of subjective happiness outlined by Lyubomirsky and Lepper (1999) ^[208]. My intention was to contextualise the SLIPS-GFP scores, and see if subjective happiness and psychosocial wellbeing were associated. As the happiness scale has been used in many countries, introducing it should have allowed me to make comparisons between the sample population and other international groups.

SUMMARY BOX

- A needs assessment and baseline study is a crucial part of the study design process, local knowledge is necessary but not sufficient for the planning, implementation and evaluation of programmes;
- I conducted a baseline study in 7 schools in Mullaitivu district, assessing MH and PH in Grade 8 and 9 students;
- The sample is predominantly male and Tamil, and the control group is almost entirely Tamil;
- Young adolescents in Mullaitivu had better MH than their counterparts in Kilinochchi;
- The study identified a considerable PH deficit, and adolescents in Mullaitivu underperformed when compared to local, regional and international norms for BFA and HFA;
- There were no significant differences in the performance of the intervention and control groups;
- Both boys and girls underperformed when compared to global average MFT scores, indicating a CF deficit in the target population;
- There was no association between the MH and PH variables, contrary to predictions;
- In recognition of the limitations of the SLIPS-C I decided to translate and introduce another MH measure that had been used internationally;
- There was a higher rate of absenteeism than predicted, which has implications for the follow up data collection and future programs.

CHAPTER 7

Impact of the sport for peace intervention

7.1 Purpose of chapter

The purpose of this chapter is to analyse the impact of the sport for peace intervention on the MH and PH of the participants. I examine the sample distribution, change within the intervention and control groups, and conduct between group analyses.

7.2 Introduction

Sport for Development and Peace is the use of sport as a mechanism for social development and change. It now covers a wide catalogue of interventions, in a multitude of settings, which claim an even broader range of outcomes [229]. Continued investment in SDP increases the demand for accountability and additional in depth evaluation and outcome analysis [229]. Despite progress in the field, there is little rigorous impact and process evaluation or peer reviewed literature in the *sport for peace* sector.

That which does exist is largely grey literature, typically manuals or toolkits that describe superficial levels of programme reach and anecdotal success stories. Research is conducted under the assumption that sport will have positive benefits, to *prove* claims, rather than to objectively generate evidence [230].

Evaluating peacebuilding faces special challenges: it does not lend itself easily to causal analysis, as it is an inherently complex process [61]. This can be tackled through effective planning and theory development, which is often lacking in sport for development and peace evaluation [230].

It appears as if funding bodies, evaluators and practitioners have competing priorities, and ideas concerning the best way to realise these goals [98]. Funding bodies are typically characterised as wanting short-term, quantifiable impacts of their input. Practitioners are forced to compete for funds and must establish their effectiveness in the short term, while peacebuilding is widely recognised as a long-term process [61]. The necessity for both international and local relevance adds complexity to the design and implementation of peacebuilding evaluations [105].

Evaluation is a sub-type of the broader category of research; there are many elements that cannot be controlled, especially in a post-conflict context. An evaluator must satisfy the demand for scientific rigor, adapt to non-ideal conditions and account for the competing interests of stakeholders. Contrasting with this, the role of an academic is to maximise the quality of the evidence collected. The role of the academic evaluator is about compromise; identifying and answering valuable questions within pragmatic boundaries.

I was aware of the role I needed to play in this project and attempted to maximise the utility of the data gathered in this study for all stakeholders by selecting, testing and adapting the outcome variables, design and methods to suit the context. By consulting with international and local organisations I designed a contextually relevant evaluation of their planned programmes in Mullaitivu.

The constantly shifting political situation in northern Sri Lanka between May 2014 and June 2015 caused significant delays in the implementation of the evaluation, and detracted from the quality of the study design. The previous Government of Sri Lanka was sceptical of foreign researchers and any NGO activity in Mullaitivu district, especially those concerned with mental and psychosocial health. At the time of writing, all NGOs wishing to conduct any programme in northern Sri Lanka had to register with the National NGO Secretariat, which came under the jurisdiction of the Ministry of Defence.

During the summer of 2014, the Ministry of Defence put the CDO under direct surveillance for illegally producing a short documentary about MH in northern Sri Lanka. As a result, the members of the CDO were harassed, interrogated and followed. Unable to register their programmes, the intervention was stalled. In 2014, I attempted to visit the field site and was forcibly removed from a train and sent back to Colombo. Unfortunately, these fluctuations in access are a common occupational hazard in restrictive post-conflict zones. This harassment continued until the change of the Government in December 2014. With the new Government, the CDO were taken off surveillance and allowed to restart the process of registration.

Despite the challenges, I managed to conduct an initial evaluation of the short-term impacts of the programme. In choosing a locally relevant and quantifiable metric that is feasibly on the 'path to peace', I have identified potential peace indicator [section 2.8]. Using this indicator as the primary outcome variable, this chapter concerns the impact of a sport for peace intervention on the psychosocial health of adolescents in Mullaitivu. The methods employed in this chapter took into account the previous testing and baseline analyses.

7.3 The Intervention

7.3.1 Training - Amman

In 2011, GFP identified four members of the northern Sri Lankan NGO, the CDO, who were invited to attend the Amman GFP basic training. GFP volunteers are expected to meet strict selection criteria including a mix of skills, values and community building experience.

During the Amman training the CDO volunteers were taught a specific GFP peacebuilding curriculum. This curriculum is devised by leading peacebuilding experts and, according to GFP, is regularly updated. The Curriculum currently (August 2016) covers:

- Theory of peace building and conflict transformation;
- Core and advanced-level conflict analysis;
- Skills and techniques for facilitating learning by different age groups;
- Behavioural change, and the use of sport, art, advocacy, dialogue and empowerment activities for change in communities;
- Volunteerism and social mobilisation;
- Programme design, monitoring and evaluation of impact and sustainability.

In addition to the training, the CDO volunteers worked with GFP mentors to a plan peacebuilding intervention for northern Sri Lanka. The plans were intended to be Sri Lanka specific, combining expert peacebuilding knowledge and local context awareness to generate sustainable and relevant solutions.

7.3.2 Intervention design in Sri Lanka

After the Amman training, the CDO staff returned to Sri Lanka finalise their intervention plan in consultation with other local volunteers, their GFP mentors and local partners. CDO identified a lack of awareness about peacebuilding and low capacity to deliver programmes. To address this, GFP funded, and CDO facilitated, a 4-day local training camp in Vavuniya for 25 volunteers. At

the training, the volunteers were taught about the 'message' of GFP, basic peacebuilding theory and peacebuilding games in action.

A significant segment of this training camp was designated to the design of an intervention.

Focus group discussions were held in groups of 4-5, the purpose of these was to identify the root causes of problems in Mullaitivu and Kilinochchi. Children and young people were identified as a crucial driving force for building peace in northern Sri Lanka. Group discussions highlighted the potential to change the opinions and values of younger people, who were less set in their ways than adults. By breaking down the fear of the other ethnic groups, and building up the confidence of young people they would better enable young people to develop positive relationships with members of the other ethnic group. Crucially, sport could be used as a 'common language' to bridge the gap between Sinhalese and Tamil speaking youth.

After the group discussion, teams were allocated the task of designing a suitable intervention for addressing the issues that had been highlighted. The following two proposals were made

- CDO/GFP should conduct some advocacy work in Kilinochchi and Mullaitivu in July 2013 and support a competitive sports league;
- CDO/GFP should deliver a schools-based sport for peacebuilding intervention in 2014.

7.3.3 Advocacy

At that time, international interventions and particularly peacebuilding programmes were unpopular with the central Government and military forces. To raise their profile, GFP supported a competitive open age sports tournament in Mullaitivu. Notably, this tournament was the first of its kind since the end of the war that successfully engaged all the local religious leaders. This advocacy work was intended to give legitimacy to GFP-related projects in Mullaitivu, and to ensure that volunteers working on their behalf were credible and perceived by the military as non-threatening.

7.3.4 The original SDP plan: Vavuniya 2013

The volunteer decided that the most suitable programme design was a year long, school based SDP programme for young people in Mullaitivu. This was with a view to expand the programme based on the evaluation of the first year. At this point, the programme elements that were agreed upon were the following:

- A school based sport for peace intervention in both Tamil and Sinhalese schools;
- Participants would be between the ages of 12-14;
- Each school would have 1-2 after school sport for peace sessions per week;
- After three months the CDO would facilitate inter-school sport for peace sessions;
- There was no ceiling placed on the number of sessions that could be held. Although this would approximate to 40 weeks of school-based sessions, excluding school holidays.
- Approximately 5 or 6 inter school meets were proposed.

The design of the intervention was loosely based around a theory of change, suggesting that using sport to increase contact between young people of different ethnic backgrounds, who were suitably primed, would lead to positive social interaction and youth development. Some important elements of the 'priming' were the development of confidence, self-esteem and the breaking down of the fear of the other. At this early stage, the programme did not have a coherent logic model, a clear programme delivery schedule or planned session elements.

At this point, I used my background literature review, experience in the field and conversation with GFP mentors and CDO staff to design and propose a simple evaluation of the intended intervention. My ideas were well received and we agreed to maintain discussions on the evaluation design and the development of the intervention elements going forward. After these initial discussions I returned to the UK to work on the further development of the intervention design and methods to share with GFP and the CDO.

What followed was a breakdown in communications between myself, the CDO and GFP. This was coupled with a total capitulation of GFP activities in northern Sri Lanka throughout 2014; the reason for this is discussed in Section 7.9.

During the period of December 2013 - December 2014, the programmes were scheduled to begin several times. However, due to the political instability, programmes were halted. In addition to these complications, I was unable to access any further information on the intervention elements. The only information that I was given in this period was the names of the intervention and control schools, and the number of attending students. GFP staff subsequently informed me (in June 2015) that the tussling with local and national authorities for permission to run programmes inhibited them from planning any particular sport for peace sessions. The constantly changing Sri Lankan political situation was the cause of uncertainty, and there was a considerable possibility of complete termination of GFP activity in Sri Lanka. In January 2015, due to the increasingly positive political situation, GFP informed me that programmes would begin in March 2015.

7.3.5 The intervention in Mullaitivu in 2015

During the two periods of data collection (March and June 2015) the CDO did not outline a plan of programme session content. At the baseline data collection in 2015, I could not obtain any detailed information on programme delivery, except a vague outline of the schedule of events; this is despite repeated requests to both GFP headquarters and CDO staff.

Between March-June, the CDO had planned to implement 36 programme sessions, the content of which had not yet been planned in March. Due to a shortage of volunteers, the CDO delivered introductory sessions in each intervention school and one or two SDP sessions.

According to the CDO, the introductory sessions involved sports-based games and some brief discussions on the schedule of programmes to come, there was no programme documentation available that detailed the contents of the sessions.

I observed one session, the third sport for peace session, delivered in Mullivaikal School on the 16th June 2015. The session was drawn from a GFP manual and was designed to develop co-ordination and teamwork. After school, the participants gathered at the back of the school in the playground and were briefed by two GFP volunteers on the aim of the game. A brief game

followed, described below, and then a very short discussion on the broader implications of the task.

The task:

Students worked in groups of 4 and suspended a tennis ball on pieces of string. The aim of the task was to coordinate dropping the ball into a bucket to encourage teamwork and co-operation. The session ran for about 20 minutes, involving fewer than 50% of the participants and then concluded quickly. The task only involved 6-8 students at any one time, despite the fact that there was a class of about 35 children present. This particular experience may not be representative of the other CDO programmes, as the capacity for programme delivery in Mullaitivu was very limited at the time.

The study design of the final evaluation differed from the original plan, which has previously been described.

7.4.2 Schedule of evaluation

The measurements taken in the needs assessment doubled up as baseline (M1) values for this pre-post study. The sample was a true baseline as, to my knowledge, there were no sports-based or MH programmes for youth running in Mullaitivu at the time. Likewise, since schools are divided along ethnic lines, there was no likelihood of contamination between intervention and control schools.

M2 data collection was initially scheduled 10 months after M1. However, because of problems registering the CDO with the NGO Secretariat, the intervention did not begin until 8 months after the intended start date. As a result, I was forced to adapt the study design to account for these changes and shorten the gap between M1 and M2. M1 took place between 2nd and the 20th of March 2015 and M2 took place between the 12th and the 26th of June 2015. This reduced the gap between M1 and M2 to 3 months. Tracking the participants as a cohort over time ensured that I could conduct within, and between, group analyses.

FIGURE 7.2 outlines a Gantt chart for the intended and the actual intervention and evaluation delivery process. It serves to highlight the considerable gap between the intention and the actual delivery. The considerable delays had widespread consequences for the study.

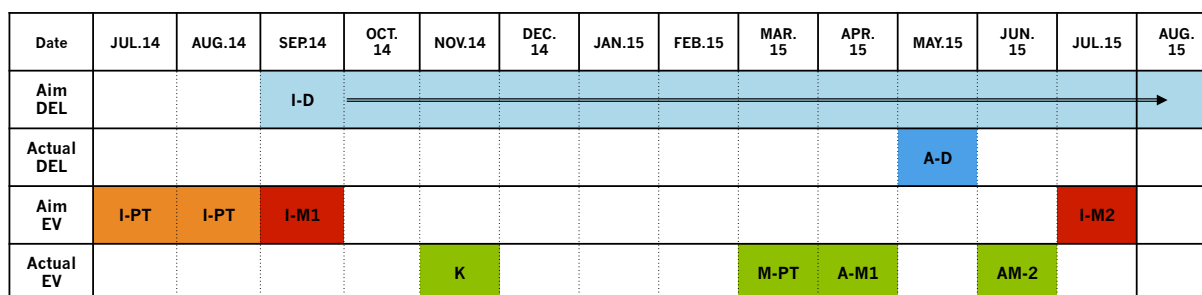


FIGURE 7-2: Gantt chart of programme elements showing intended and actual delivery schedule

TABLE 7-2: Legend for Gantt chart

Acronym	Detail
I-D	Intended delivery date
A-D	Actual delivery date
I-PT	Intended pre-testing of methods
I-M1	Intended baseline measurement
I-M2	Intended follow up measurement
K	Kilinochchi pre-testing
M-PT	Mullaitivu pre-testing
AM-1	Actual baseline measurement
AM-2	Actual follow up measurement

7.4.3 Sample size calculation

The sample size calculation is the same as for the baseline study (section 6.3.2). Accounting for 10% absenteeism, assuming a power of 0.8 and a significance level of 0.05 the number of individuals required in each group was 194.

7.4.4 Recruitment

Participants were recruited through their schools. The CDO and the GFP Sri Lanka Co-ordinator selected the schools that were included in the study and allocated them to an intervention or control group. Schools not selected for the first round of the intervention were wait-listed to

receive it in the following programme cycle, when the programme was to be expanded following a 1-year evaluation. The CDO chose the intervention schools based on their local knowledge. They included schools that had been badly affected by the civil conflict, and were close to former LTTE military encampments. Likewise, the intervention schools had to represent all the different ethnic groups in the district. There were a total of 7 schools selected for inclusion in the study; the 8th school, CHM, was dropped from the evaluation and only used in the needs assessment and baseline analysis.

7.4.5 Tracking

Simple demographic information was collected as part of the SLIPS-GFP. Participants name, age, grade and class were recorded. The information was entered into a spread sheet that was password protected on an encrypted laptop. All participants were given identifiers, kept separately from the hard copies, which allowed me to track them over time. All hard copies were kept in a locked cabinet.

7.4.6 Outcomes

The outcome variables were the same as in the baseline study, except that HFA was not recorded at M2. This is because HFA is an indicator of chronic malnutrition and unlikely to change over a three month period.

7.4.7 Procedure

The data collection procedure was identical to that described in CHAPTER 5 and CHAPTER 6. There were minor deviations from the plan in several schools, when we were not allowed access during the allocated testing period. The follow up data collection was particularly hampered in THN School. As we were having trouble accessing the target group, I offered to collect data in the evening, as the Principal was not allowing access during school hours. We attempted to collect data on 4 separate occasions in THN. 3/4 times the target group were not in school because it was a holy day, or the students had already left school for the day and not returned as promised.

On one occasion we were not allowed access as the school was having an inspection by the Ministry of Education and wanted students in class.

7.4.8 Analysis

The analysis is in two main sections: a sample distribution section and an impact section.

(i) Sample distribution

The sample at M2 is described in terms of gender, ethnicity and age in terms of absolute numbers and proportional representation. This section identifies any disparities between population distribution in the intervention and control groups.

(ii) Cohort study impact assessment

Descriptives were calculated for all major outcome variables at M2, stratified by intervention and control status. Paired t-tests were used to assess within group changes in crude means of all outcome variables. Repeated measures ANOVA were used to compare the magnitude of the change from M1-M2 for group means for intervention [V] and control [C] groups. Once the analysis was completed for crude data for each outcome, the analysis was completed adjusting for gender, age, and ethnicity. An additional model was constructed that controlled for age, ethnicity and gender, including baseline scores as a covariate. The within and between group results were summarised and tabulated.

7.4.9 Ethical approval

Ethical approval for the project was granted by OxTREC and the Faculty of Medicine Research Ethics Committee, Colombo University.

7.5 Results

434 students were enrolled in the 7 schools chosen for participation in the impact study. We collected measurements at M1 and M2 for 214 participants (49.3% of all eligible students). There was considerable loss to follow up between M1 and M2, particularly in the Muslim school, THN. Loss to follow up was 34.2%, not including the 8th school that was not involved in the impact study.

The sample size for the MFT measurement is smaller than the MH measurements; several participants did not take part in the MFT at M2. Very few MFT measurements were taken in the THN, because it was Ramadan and all the participants were fasting at the time of testing.

I had initially intended to stratify the sample by intervention status and gender. However, due to large loss to follow up, these analyses should be read with caution given that they are underpowered. Both gender and ethnic group were identified as important confounding factors in the baseline assessments.

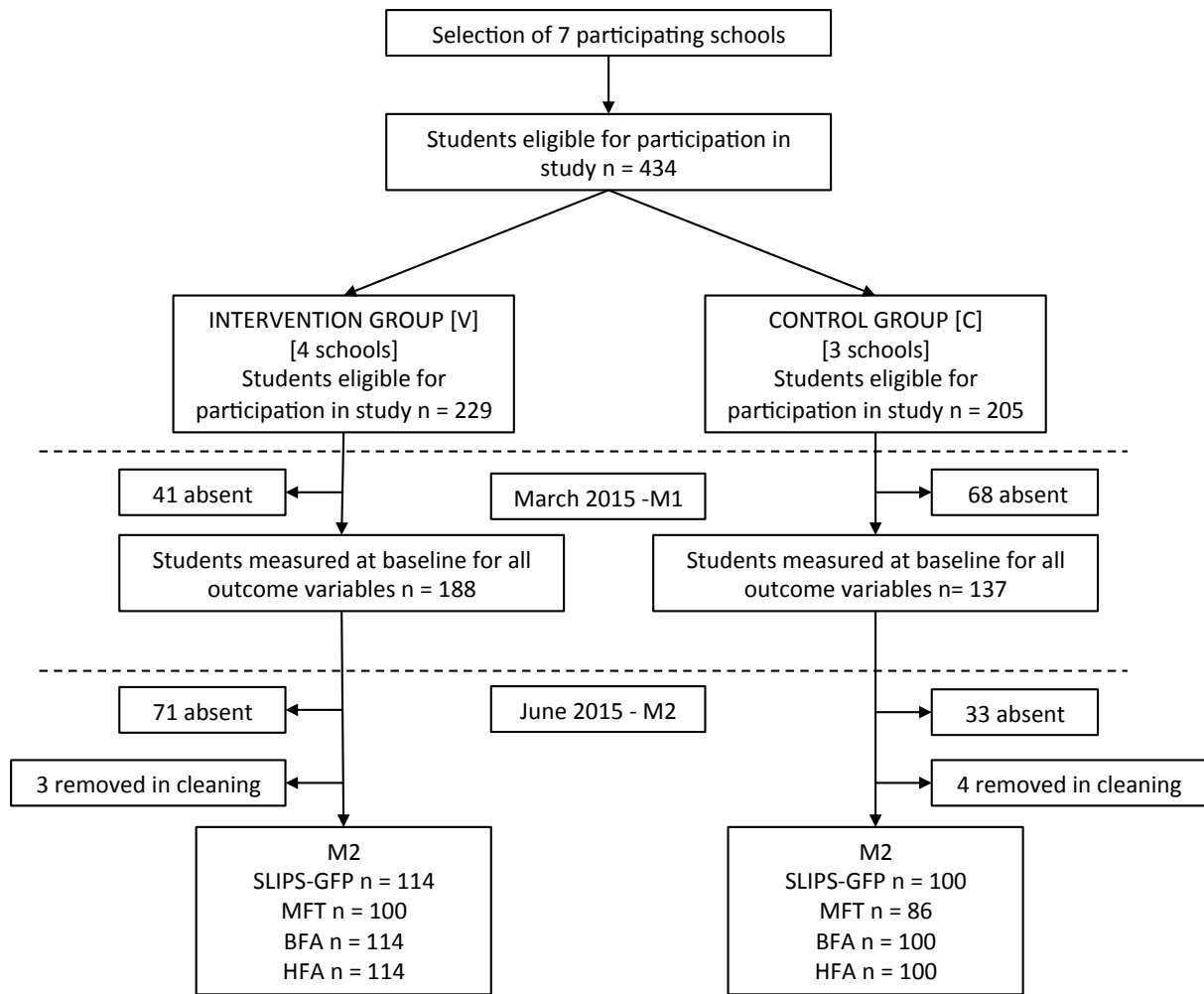


FIGURE 7-3: Study flow diagram

7.6 Sample distribution

This section outlines the sample distribution in absolute terms and proportionally. TABLE 7.3 outlines the sample distribution for all individuals who had an MH measurement at M1 and M2. The sample is evenly distributed across intervention and control groups, with similar proportions of boys and girls in each group. Likewise, both groups are almost entirely 13 or 14 years of age. There is a disparity in the ethnic composition of the intervention and control groups. Given the difference in MH performance at baseline, it is necessary to adjust for ethnic differences in later impact analyses.

TABLE 7-3: Table showing sample distribution across intervention and control groups (n = 214)

Independent variable		Intervention [V]		Control [C]	
		n	% of Int sample		% of Cnt sample
Age	12	0	0	2	2
	13	58	50.9	52	52
	14	54	47.4	43	43
	15	2	1.8	3	3
Gender	B	60	52.6	49	49
	G	54	47.4	51	51
Ethnicity	Tamil	50	43.9	95	95
	Sinhala	37	32.5	5	5
	Muslim	27	23.7	0	0
n		114		100	
Total n		214			

7.7 Impact of the intervention – within group analysis

Descriptive statistics for all the outcome variables at M1 and M2, stratified by intervention and control status, are in TABLE 7.4. MH scores decreased in the intervention and control groups, indicating an improvement in psychosocial health between M1 and M2. MFT scores in the intervention group deteriorated, but remained constant in the control group. There were negligible changes in BFA in both the intervention and control groups (TABLE 7.5, FIGURES 7.4-7.8).

TABLE 7-4: Crude mean and 95% confidence intervals for all outcome variables in the intervention group at M1 and M2 (n = 214)

Group	n	Time	BFA	MFT	SLIPS-GFP	INT	EXT	WTH
Intervention [V]	114	M1	-1.81 -1.99<->- 1.62	4.59 4.31<-> >4.86	23.58 22.01<-> >25.13	13.24 12.36<-> >14.10	5.30 4.72<-> >5.88	3.89 3.47<-> >4.30
		M2	-1.65 -1.83 <-> - 1.48	4.06 3.75<-> >4.37	21.46 19.84 <-> 23.07	12.11 11.22 <-> 13.01	4.28 3.62 <-> >4.94	4.29 3.89 <-> >4.69
Control [C]	100	M1	-1.72 -1.92<->- 1.52	4.53 4.22<-> >4.84	26.06 24.29<-> >27.82	14.40 13.50<-> >15.30	6.50 5.66<-> >7.34	4.15 3.70<-> >4.60
		M2	-1.65 -1.83<->- 1.43	4.58 4.22<-> >4.94	21.69 19.93<-> >23.45	11.91 11.03<-> >12.79	4.68 4.01<-> >5.35	4.21 3.72<-> >4.70

From the crude mean results alone it is not possible to determine if the change in outcome scores was statistically significant. I plotted the mean scores and 95% confidence intervals for all outcome variables at M1 and M2, stratified by intervention status. Having assessed the change over time visually, I used paired t-tests to identify any significant change in scores

between M1 and M2. The use of the statistical methods requires that the data fulfil certain assumptions, these are verified in APPENDIX E.1.

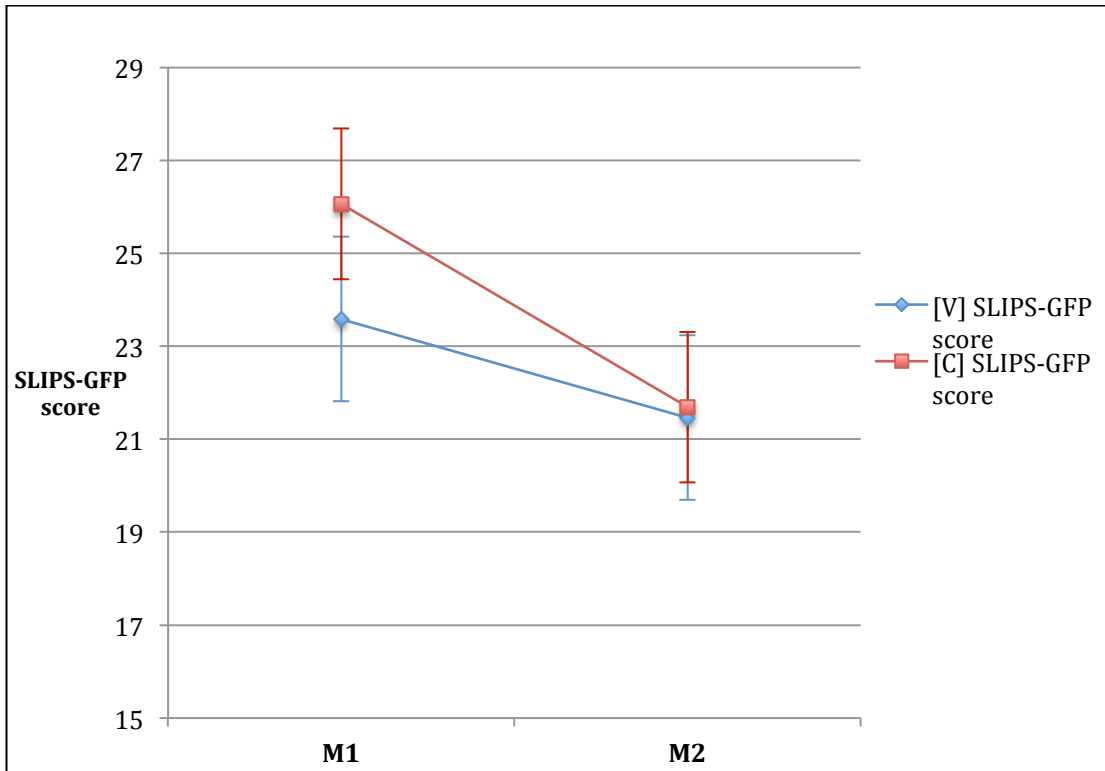


FIGURE 7-4: Crude mean and 95% CIs SLIPS-GFP scores at M1 and M2 for intervention and control groups (n = 214)

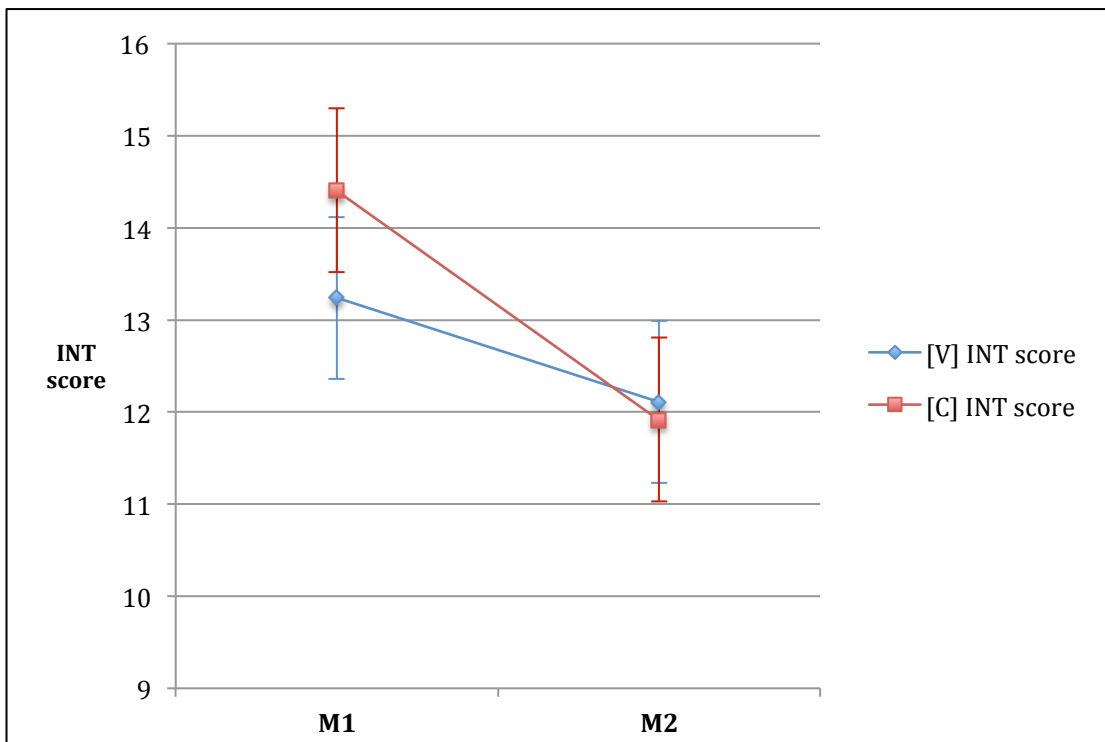


FIGURE 7-5: Crude mean and 95% CIs of INT subscale score for intervention and control groups (n = 214)

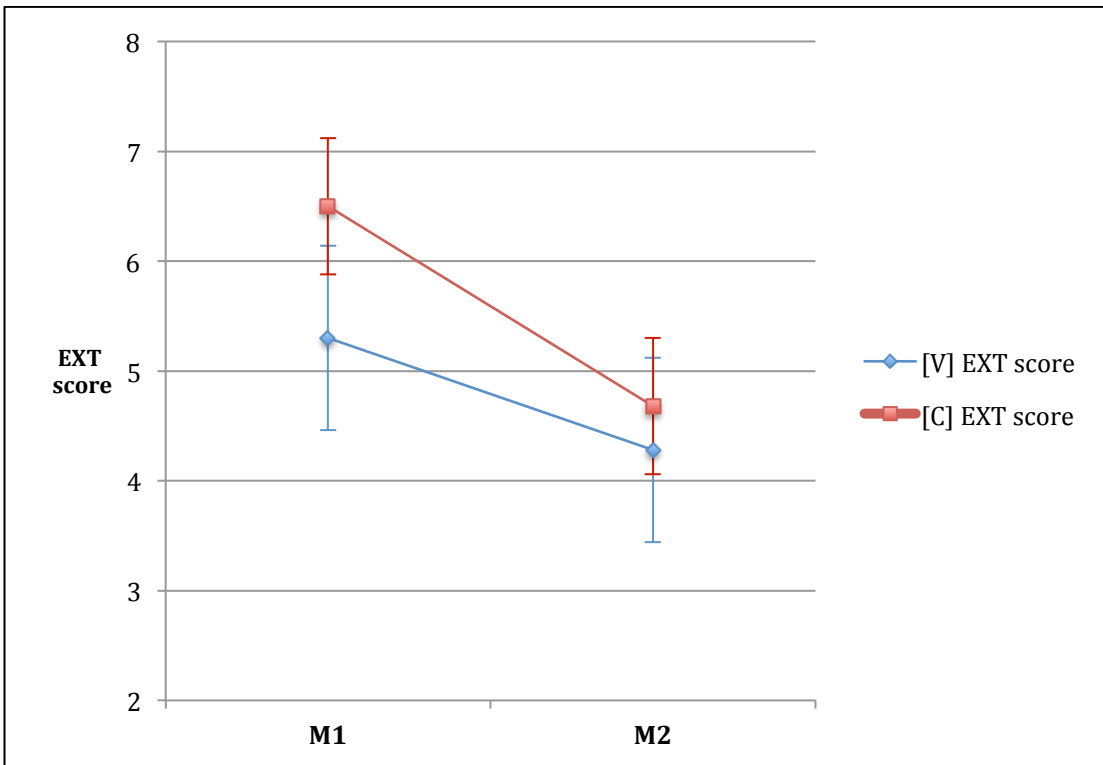


FIGURE 7-6: Crude mean and 95% CIs of EXT subscale scores in intervention and control groups at M1 and M2

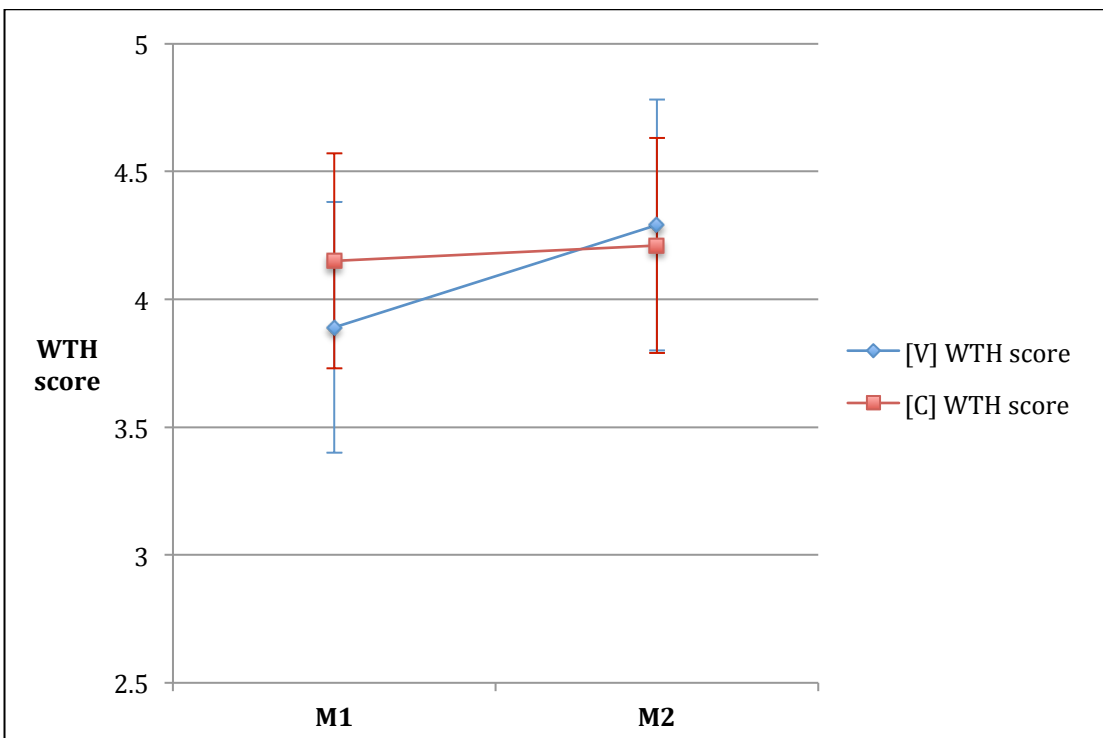


FIGURE 7-7: Crude mean and 95% CIs for WTH subscale score in intervention and control groups at M1 and M2

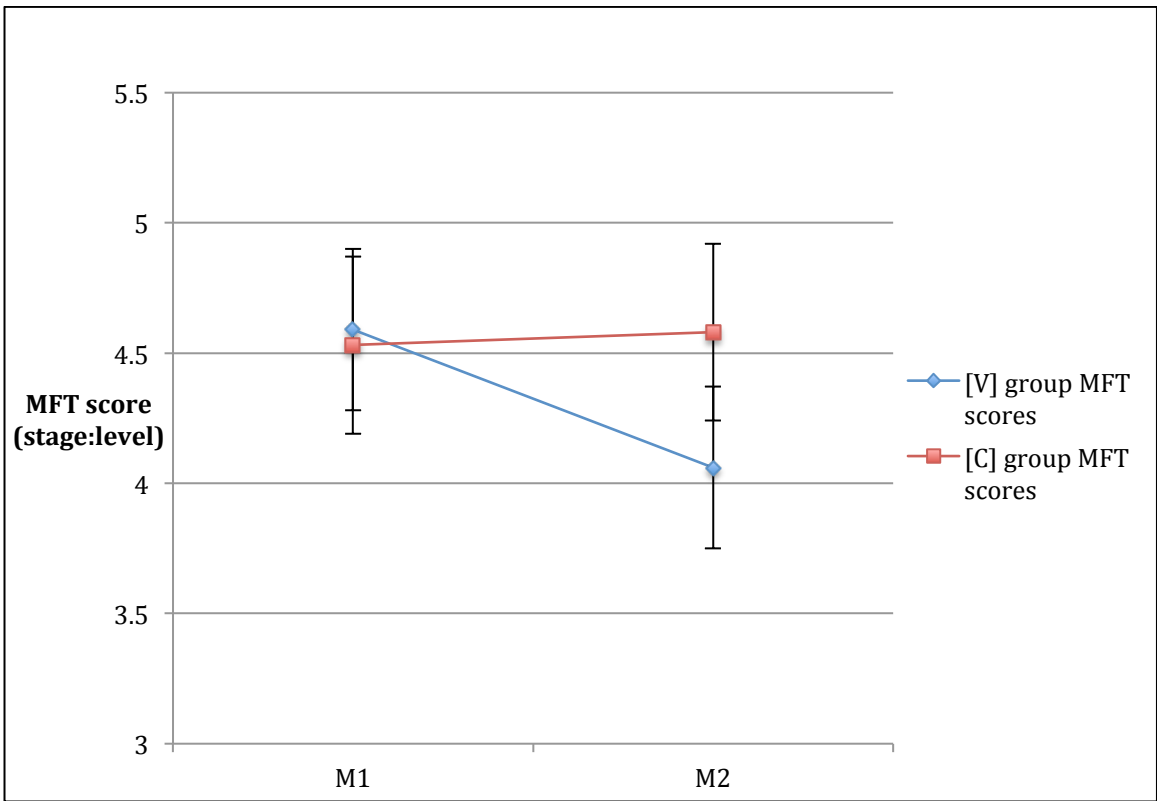


FIGURE 7-8: Crude mean and 95% CIs for MFT in intervention and control groups at M1 and M2

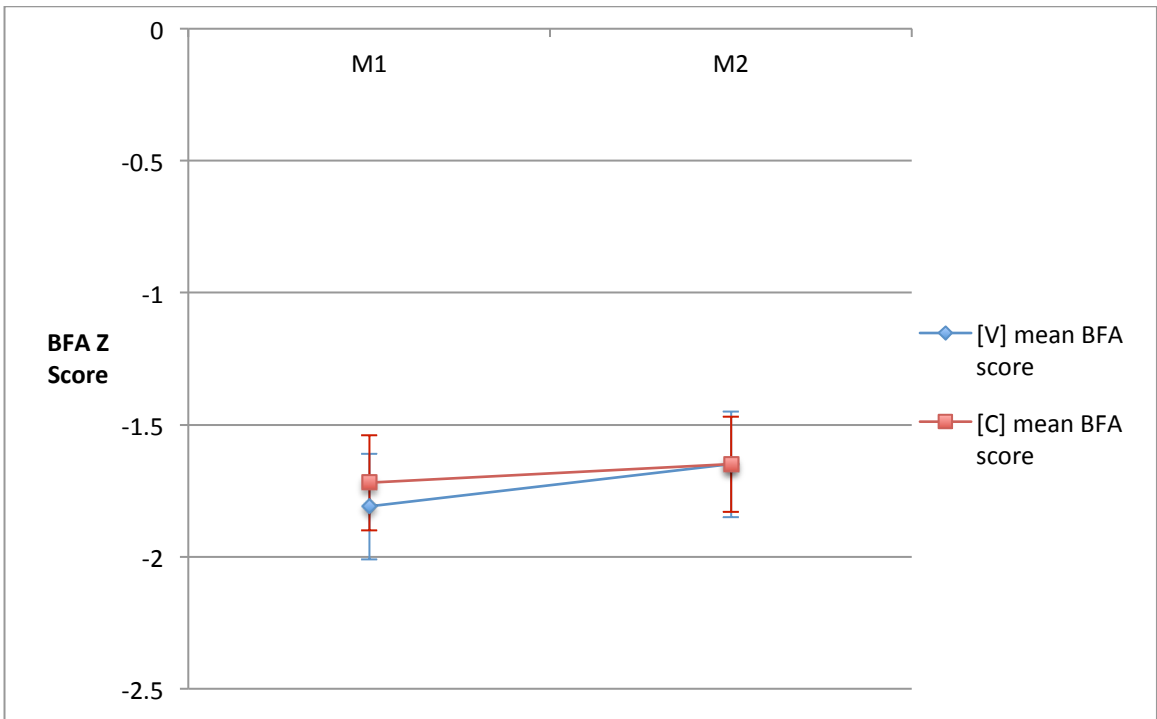


FIGURE 7-9: Crude mean and 95% CIs for BFA score in intervention and control group at M1 and M2

TABLE 7-5: Results of paired t-tests all outcome variables comparing M1 and M2 crude means scores in intervention and control groups (n =214)

	Independent variable	d.f	t statistic	Sig.
Intervention	BFA	113	-4.72	<0.001**
	MFT	99	3.39	0.001**
	SLIPS-GFP	113	2.46	0.016*
	INT	113	2.03	0.044*
	EXT	113	2.74	0.007**
	WTH	113	-1.82	0.071
Control	BFA	99	-2.02	0.046*
	MFT	86	0.35	0.724
	SLIPS-GFP	113	5.04	<0.001**
	INT	99	4.80	<0.001**
	EXT	99	4.61	<0.001**
	WTH	99	-0.215	0.83

NB: Bold and italic indicates a significant result

NB: * indicates 95% confidence level, ** indicates 99% confidence level.

Analysis of the crude results suggests that there was a significant change in performance between M1 and M2 in almost all of the primary outcome variables, in both [V] and [C] groups. Intervention crude mean SLIPS-GFP score improved significantly (M1 mean = 23.58, M2 = 21.46)(t (113) 2.46, p = 0.016), as did control SLIPS-GFP score (M1 mean = 26.06, M2 mean =21.69)(t (99) 5.04, p = <0.001), indicating a significant improvement in MH. In addition, both groups recorded significantly lower INT and EXT subscale scores.

Mean MFT score in the intervention cohort significantly deteriorated between M1 and M2 (M1 mean = 4.59, M2 mean = 4.06, t (99) = 3.39 p = 0.004), yet did not change significantly in the control group. Visual assessment of the BFA scores does not indicate a significant improvement in either the intervention or control groups (FIGURE 7.9). However, ANOVA indicated a significant improvement in the BFA scores in the intervention and control group BFA scores (Intervention: t (113) -4.72, p <0.001; control t (99) -2.02, p = 0.046).

TABLE 7-6: Table of mean and 95% confidence interval MFT scores at M1 and M2, stratified by intervention status and gender (n = 214).

Dependent variable	Gender	n	SLIPS-GFP		MFT score		BFA	
			M1	M2	M1	M2	M1	M2
[V]	Boys	60	23.83 21.62- 26.04	21.35* 19.21- 23.49	5.40 5.01 - 5.79	4.52** 4.10 - 4.95	-2.10 -2.35 <-> -2.18	-1.93* -2.15 <-> -1.70
	Girls	54	23.30 21.06- 25.53	21.57 19.06- 24.10	3.8 3.53 - 4.17	3.45 3.07 - 3.84	-1.48 -1.73 <-> -1.23	-1.35** -1.59 <->- 1.10
[C]	Boys	49	28.65 26.30- 31.00	23.10** 20.35- 25.84	5.48 5.08 - 5.87	4.60** 4.09 - 5.11	-1.99 -2.28 <-> -1.70	-1.86 -2.16 <-> -1.57
	Girls	51	23.57 21.06- 26.08	20.33** 18.09- 22.58	3.65 3.27 - 4.03	4.56** 4.04 - 5.07	-1.46 -1.72<-> -1.18	-1.40 -1.67 <-> -1.13

NB: * indicates a significant difference between M1 and M2 to the 0.05 level, ** indicates significance to the 0.01 level, MFT score n's are smaller than those listed in the table.

TABLE 7.6 outlines the results, stratified by gender. Boy's MFT scores deteriorate significantly in both the intervention and the control groups [V] boys: $t(56) 2.99, p = 0.004$, and [C] boys: $t(47) 2.80, p = 0.007$. Intervention girls MFT scores remain the same, and control girls MFT scores improved significantly ($t(37) = -2.23, p = 0.01$). BFA significantly improved in both [V] group boys and girls: [V] boys ($t(59) = -3.24, p = 0.002$) [V] girl's $t(53) = -3.76, p < 0.001$.

FIGURES 7.4-7.9 suggest that there were no significant differences between [V] and [C] outcome scores at M2. An ANOVA verified that there were no significant difference in MH score between [V] and [C] groups at M2. The [C] group significantly outperformed the [V] group in MFT ($F(184) = 1.39, p = 0.0029$). The assumptions for implementing an ANOVA were verified in APPENDIX E.2.

TABLE 7-7: ANOVA comparing outcome variable scores at M2 for intervention and control groups (n = 214)

Dependent variable	F	df1	df2	Sig.
BFA_M2	0.036	1	212	0.849
MFT_M2	1.71	1	168	0.029*
SLIPS_M2	0.038	1	212	0.846
INT_M2	0.103	1	212	0.748
EXT_M2	0.708	1	212	0.401
WTH_M2	0.064	1	212	0.801

A summary of the MH and PH scores, stratified by gender, ethnicity and age are reported in

APPENDIX E.3

7.8 Impact of intervention – between group analysis

Mean change between M1 and M2 scores were calculated for the major outcome variables for [V] and [C].

TABLE 7-8: Crude mean, 95% CIs and ANOVA for mean change between M1 and M2 for [V] and [C] groups (n = 214)

Mean change M1-M2	Intervention		Control		F.	Sig.
	n	Mean	n	Mean		
SLIPS-GFP	114	2.12 0.41 - 3.84	100	4.37 2.65 - 6.09	3.34	0.069
INT		1.11 0.03 - 2.20		2.49 1.46 - 3.52	3.27	0.072
EXT		1.02 0.28 - 1.75		1.82 1.04 - 3.52	2.19	0.141
WTH		-0.40 -0.84 - 0.03		-0.6 -0.61 - 0.50	0.947	0.332
BFA		-0.15 -0.22 - -0.09		-0.09 -0.18--0.001	1.25	0.264
MFT	100	0.642 0.21 - 1.08	84	0.0871 -0.40 - 0.58	3.25	0.733

An ANOVA verified that there was no significant difference in the magnitude of the change between M1 and M2 between [V] and [C] groups. ANOVA confirmed that there was no significant difference in MH score change between [V] boys (mean change = 2.48) and [C] boys (mean change = 5.55), or between [V] girls (mean change = 1.72) and [C] girls (mean change = 3.24). The assumptions for the use of the ANOVA are verified in APPENDIX E.4.

I repeated the ANOVA comparing only the Tamil participants in [V] and [C] groups, to ensure that ethnicity did not affect the results. When comparing Tamil intervention participants (n = 50) and control Tamils (n = 95) there was no significant difference in the magnitude of the change between M1 and M2 across any of the outcome variables (TABLE 7.9).

TABLE 7-9: Mean and 95% confidence intervals for mean change between M1 and M2 in only Tamil participants (Intervention Tamils n = 50; Control Tamils n = 95).

Dependent variable	Mean change M1-M2		Sig.
	[V] Tamils	[C] Tamils	
SLIPS-GFP	2.680 0.322-5.039	4.715 2.945-6.487	0.077
MFT	0.343 -0.137-0.822	-0.057 -0.555-0.441	0.290
BFA	-0.117 -0.234--0.0004	-0.100 -0.194--0.007	0.742

7.8.1 Multivariate analysis

I constructed a multivariate model that adjusts for gender, ethnicity and age to confirm that sample distribution did not affect the statistical findings, and that there is no difference in the magnitude of the change between M1 and M2 for any of the outcome variables. Gender, age, and ethnicity were entered into the model as covariates.

Model: [difference in score] = gender, ethnicity, age, [V]-[C] status.

TABLE 7-10: ANOVA table adjusting for gender, age, ethnicity (n = 214).

Dependent variable	Mean square	F	Sig.
SLIPS-GFP	2.24	0.03	0.862
INT	0.004	<0.001	0.991
EXT	9.33	0.595	0.441
WTH	7.42	1.18	0.280
MFT	2.89	0.796	0.373
BFA	0.21	1.41	0.237

The adjusted values confirm that there appears to be no significant difference in the magnitude of the change between M1-M2 in [V] and [C] groups when controlling for gender, age and ethnic group; although the study is underpowered which limits the strength of the conclusions that can be drawn. Additionally, I constructed a model that controlled for ethnicity, gender, age and baseline score (treating baseline score as a covariate). This confirmed that there was no significant difference between the magnitude of the change in the intervention and control group when adjusting for baseline and other factors.

TABLE 7-11: Table of results of ANOVA adjusting for baseline, gender, ethnicity and age (n =214).

Dependent variable	Mean square	F	Sig.
SLIPS-GFP	38.85	0.708	0.401
MFT	1.03	0.463	0.497

7.8.2 Intention to treat analysis

As there was considerable loss to follow up in the study, I assessed the need for loss to follow up calculations. I assessed the difference in baseline scores for the main outcome variables in participants with and without M2 scores. I used an ANOVA to assess whether there were any significant differences between mean baseline scores in participants with and without M2 data. The statistical assumptions for using the ANOVA are verified in APPENDIX E.5. An ANOVA verified that there were no differences in the baseline scores of those individuals who had follow up data and those who did not.

TABLE 7-12: Mean and 95% confidence intervals at M1 for participants with and without M2 data.

Group	n	Dependent variable	Mean score
Participants with no follow up data	105	SLIPS-GFP	26.50 24.93-28.06
		MFT	4.78 4.44-5.11
		BFA	-1.74 -1.97--1.51
Participants with follow up data	214	SLIPS-GFP	24.74 23.57-25.91
		MFT	4.56 4.36-4.77
		BFA	-1.75 -1.92-- 1.58

Given that there were no differences between those individuals with M2 data and those without, I did not complete an intention to treat analysis on the grounds that there were no differences in the baseline scores for all outcome variables between those individuals included in the full study and those who were lost to follow up. Secondly, as the degree of loss to follow up was so high, imputing baseline values as M2 values would have almost certainly diluted any changes seen in those individuals who participated in the intervention.

7.9 Discussion

7.9.1 Principal findings

The sample was evenly distributed between the intervention and control groups, and the gender and age distribution was uniform. There was a difference in the ethnic distribution of the intervention and control groups. Loss to follow up varied considerably by ethnic group (Tamil 27%, Sinhala 32% and Muslim 53% loss to follow up).

MH scores significantly improved between M1 and M2. When the results are stratified by gender, the largest improvement in MH is seen in the control boys. This largely explains the improvement in MH in group [C]. The only group to not improve significantly were the [V] girls, who experienced a small, non-significant improvement. When ethnic differences in MH are considered, the [V] Sinhalese experienced almost no change in MH, while all other ethnic groups improved considerably in both the [V] and [C] cohorts (APPENDIX E.5).

CF deteriorated significantly between M1 and M2 in the intervention cohort, but remained unchanged in the control group, as a result the control group significantly outperformed the intervention group. When the results are stratified by gender, boy's performance significantly decreased in both the intervention and control cohorts. The performance of the control girls, who improved significantly, masked the effects of the control group boys decrease in MFT score. The control cohort girls were the only group to improve their MFT scores significantly. Performance in Tamil and Muslim boys only decreased slightly, whereas the Sinhalese boys had a marked drop in CF.

There were small, significant improvements in BFA in the intervention and control groups. When the results were stratified by gender, all groups BFA scores improved.

7.9.2 Between group changes:

There were no significant differences in the MH of the intervention and control groups at M2. When the results were stratified by gender, there were no significant differences in the M2 MH

scores of intervention boys when compared to control boys, or between intervention girls and control girls.

There was no significant difference in the CF of intervention and control boys, or the intervention and control girls at M2. Boys in the intervention group outperformed the intervention cohort girls significantly in the MFT ($f(1) = 5.34, p = 0.023$), but did not significantly outperform the control girls at M2 ([V] boys mean = 4.52, [C] girls mean = 4.56).

There were no significant differences for any of the outcomes when comparing the amount of change in the intervention and control groups. Adjusting for gender, ethnicity, age and baseline scores, did not affect the significance of the change between M1 and M2 for any of the outcome variables in the intervention and control group.

There were significant improvements in the MH of all participants enrolled in the study, regardless of intervention or control status. Control cohort boys MH improved the most during the study, while the Sinhalese participants experienced the smallest change. The results suggest that the intervention had no additional effect on the MH of participants. There was no difference in the magnitude of the change in CF between the intervention and control groups. When the results are stratified by gender it is apparent that boy's CF, irrespective of intervention status, decreased significantly. The intervention had no effect on girls CF. Further, the intervention did not affect their body shape any more than if they had not taken part in the intervention.

7.9.3 Strengths and limitations of the study

The results of this build on the data described and analysed in the previous studies. The principal weakness of this study is that it is underpowered; as a result there are limited conclusions that can be drawn from the findings. An underpowered study increases the risk of drawing false positive conclusions.

The reliance upon the CDO was both a strength and a limitation of this study. The CDO had limited capacity, and as a result could only operate in a limited number of schools. This had several implications for the power of the statistical analysis conducted in the study, and the sample distribution.

The CDO selected the schools that participated in both arms of the study ([V] and [C]). Their selection was based on the need to include several different ethnic groups in the intervention, and also to include those students most affected by the civil war and tsunami. The local geography constrained the CDO in their choice, as there were only two Sinhalese schools within a feasible distance. As a result the intervention and control groups were unevenly distributed according to ethnicity.

Intervention and control group size were heavily influenced by the CDO capacity; at the time of the study they could not operate in more than 4 schools simultaneously. This meant that the margins for error were very small, given the sample size calculations conducted in Chapter 6. The large numbers of pupils absent at M1 contributed to the sample size shortfall. Nesting the study in schools was an attempt to minimise loss to follow up in the study, and provided a convenient way to track students over time. However, the high rate of absenteeism and loss to follow up meant that we did not achieve the target numbers, and the analyses were all underpowered.

The reliance upon local knowledge and the presence of an assumed deficit in the selected schools, rather than the systematic testing of assumptions, could have led to a school-level selection bias. There is variation in the quality, size and wealth of schools in northern Sri Lanka. On account of the sample size, it was not possible to conduct meaningful school-level analysis. There may have been important clustering at the school level that we cannot feasibly explore with the current dataset.

The high rate of absenteeism at M1 raises concerns over the representativeness of the sample. In addition, informal interviews with programme staff indicated that those children who were most severely affected by the war are unlikely to be in school regularly [Interview T]. A cross sectional study by Siriwardhana et al. (2013) indicates that absenteeism is associated with poorer MH status and conflict exposure^[158]. As a result the sample may poorly reflect the experience of those most in need of assistance, and who stand the most to gain by involvement in an intervention. I would like to coin the phrase the ‘inverse intervention law’; it states that those who have the most to gain from an intervention, are the least likely to engage with it. I

tried to account for the high rate of absenteeism by visiting each school on two separate occasions to collect data, but it is evident that this was not sufficient in this context.

The consultation and involvement of the CDO in the design, planning and implementation of the evaluation aimed to increase the local relevance of the results, making sure the aims of the evaluation were feasible, while contributing to the community capacity to conduct scientifically rigorous research in low-resource settings. This is in keeping with current best practice, which encourages participatory planning and evaluation of programmes ^[103]. The adaptation of the design and the methods is a considerable strength in this study, and I aimed to optimise the level of evidence obtained given the pragmatic constraints of the context.

The selection of a locally devised MH assessment tool limited the external validity of the results. I translated and introduced Lyubomirsky and Lepper's (1997) four-item happiness scale at M2, to assess whether SLIPS-GFP score was associated with subjective happiness. This was an attempt to contextualise the SLIPS-GFP results. However, there were considerable problems with the translation of the happiness scale, and the internal consistency of the scale was very poor in this population ($\alpha < 0.3$). Earlier translation and initial validation studies in the formative evaluation would have been potentially beneficial for the inclusion of the subjective happiness measure.

The clinical significance of the SLIPS-C has never been established, as it has only previously been used in cross-sectional associative studies. This study established the utility of the SLIPS-C as a potential metric to track MH over time in low-resource settings. Using a previously validated MH assessment tool, such as the SDQ, may have provided results that could be more broadly compared. I selected the SLIPS-C as the primary outcome variable for the study as I felt it best suited the research question. In this study, the priority was to assess the impact of an intervention, rather than to identify clinically significant cases or conduct population surveillance. In my opinion, at the time of planning, the need for a locally specific and relevant measure exceeded the need to make international comparisons. In hindsight, given the limited implementation of the intervention, a measure that produced externally valid results may have been of more scientific value.

Local norms in time management and planning caused considerable methodological problems during the project. I constantly found myself following unpredictable, shifting schedules and the changing of plans without my consent. These issues were compounded by the language barrier, which should not be underestimated. It is not uncommon for people to be involved in several 'volunteer' jobs at once in northern Sri Lanka; it is a practical response to the poor employment opportunities in the region. However, as a result, volunteers have constantly shifting agendas and competing priorities. The inability to effectively communicate and co-ordinate meant that data collection opportunities were missed, and there was deviation from the intended data collection schedule.

The constantly changing data collection team meant several threats to measurement protocol, and ensuing validity. The inability to train, maintain and develop a research team meant that the same mistakes were made time and time again and the protocol implementation suffered as a result, most likely increasing the rate of random error. It is widely recognised that the staff are a crucial element in producing successful peacebuilding or MH programmes [84, 100]. An additional, related point is that this contributes to a lack of capacity and skills for M&E in SDP agencies in the field [92].

During November 2014 I conducted a cross-sectional study in Kilinochchi. I formed a partnership with Jaffna University, which provided 3rd year Medical Students to collect data as part of a community health programme. Each year there is a new group of students willing to learn about data collection methods. All students were required to attend a certain number of data collection sessions to ensure that they received a certificate of participation. This ensured the team remained the same throughout the Kilinochchi data collection, and could ensure a new set of students regularly. This agreement does not, however, ensure that skills are retained in a specific set of team members, rather it ensures a reliable source of new recruits. Training physical education staff members in schools to collect data, or staff members at the District Officer for Sport, might be beneficial as these individuals tend to remain in place longer than volunteer staff.

The effect of local norms on the measurement validity was compounded by the socio-political context of the region. As a result of military interference and control, I was limited in the methods, data collection tools and scale of the evaluation I could conduct. Access to schools for data collection was highly variable and unpredictable. Political involvement in schooling meant that certain principals were hesitant to allow us access to schools during school time. This was particularly relevant for the Muslim community and responsible for the large loss to follow up in the intervention cohort. Restricted access meant that we could not collect data before the start of Ramadan and many participants had left for another region of Sri Lanka.

An additional weakness in the design of this study was the lack of process data. I scheduled to attend all possible intervention sessions, and designed evaluation forms accordingly. Only $\frac{1}{4}$ programme sessions were held while I was in Mullaitivu. During the period of March-June a total of 6/36 planned sessions were delivered. I attended, took notes and informally interviewed the programme staff that conducted a session. This information alone is not enough to explain any of the effects of the intervention. This is reflective of the field of SDP, there is poor process data collection and the articulation of 'mechanisms' behind proposed interventions remain undisclosed ^[230].

The loss to follow up in the study was very large, and I did not complete an intention to treat analysis. As a result, the study was an efficacy study, as it only considered the effect of the intervention on 'completers'. There was no systematic measurement of adverse effects of the intervention. However, as with the process evaluation, as there was so little of the intervention delivered this was not of major concern.

There are four common factors noted to complicate the evaluation of peacebuilding programmes:

1. the outcomes of peacebuilding interventions are long-term and difficult to quantify;
2. it is difficult to isolate the effects of a peacebuilding intervention from wider social impacts;

3. all the actors involved have different expectations of the intervention and evaluation process;
4. there is no common definition of what peacebuilding is, and should be trying to achieve [103].

In light of the study, I have added an additional factor:

5. it is difficult to specify the 'dose' received and delivered in a peacebuilding intervention;

The design of this study addresses many of these issues head on. Firstly I outlined a potential theory linking the intervention and long-term peace and identified measurable short-term indicators. There is a lack of empirical application of peacebuilding theory; this study systematically tested assumptions outlined in a theoretical framework. To my knowledge it is one of few studies that collected primary data on the association between sport for peace programmes and MH to systematically test a pre-defined theory.

Secondly, the use of a control group meant that the potential effects of the study could be isolated from wider social changes. If there were no control group in this study, the conclusion would be that even a limited dose of the programme caused a significant improvement in MH. In reality, there was a background increase in MH across all participants included in the study.

The final intervention was designed in association with the local partner organisation to ensure that it was sensitive to local context, culture and resources. In this sense, the evaluation was a success, all the elements to repeat the process are available locally, and several members of the CDO have had experience in implementing the data collection protocol. GFP, the funding body, were consulted extensively throughout the process and provided logistical and financial support where necessary. However, this is a potential conflict of interest, as GFP have a large amount at stake in this thesis.

7.9.4 Mechanisms and implications of study

There are several potential explanations for the results of the study. Before attempting to explain any of the results it is important to recognise that the study was underpowered, and any

of the conclusions must be read with caution. However, the repeatability and validity of the measures has already been established in previous chapters, and the changes witnessed in the participants should not be written off as noise or random error. Another important point is that there was considerable deviation from the original intervention implementation plan.

GFP were caught filming in a Sinhalese school without permission in August 2014. As a result, GFP were forcibly ejected from Sri Lanka and the CDO staff were harassed, intimidated and followed. As a result, the CDO could not apply for a permit to begin the intervention in September 2014 [interview: KP]. The breakdown of the relationship between GFP, the CDO and the Military Police in Northern Province meant that the CDO failed to retain sufficient volunteers to conduct programmes effectively.

The CDO were only granted permission to begin programmes in March 2015, 9 months after the intended start date. The effects of this were manifold: only 4/32 trained CDO volunteers remained active at the beginning of the intervention in March 2015, the time between M1 and M2 was reduced from 10 to 3 months, and the M2 data collection schedule was compressed into a tight two week period before the start of Ramadan and summer holidays.

From 1st July, no further intervention sessions were scheduled until the start of the new school year. As a result, I felt it necessary to conduct the follow up measurements when I could. This meant reducing the follow up to 3 months. It was a pragmatic decision based on the fact that if I delayed until September, the participants would have had 2 months of no intervention at all, increasing any potential contamination or noise.

An additional effect of the filming was that the intervention Sinhalese school refused any further involvement with GFP and CDO, severing all ties. The result was the almost complete failure to implement any of the intervention in the period between March and June 2015. As a result of the delays, the intervention participants received such a small 'dose' of the intervention that it was unlikely that any of proposed benefits to participation could have been realised at the point of M2. Additionally, it is unlikely that the CDO, GFP and the Sinhalese school will patch up relationships in the near future, which has implications for the reach of the future programmes.

Mental health

There are no clear explanations for the improvement in the MH of the study participants. There was a community-wide significant improvement in MH, apart from in the Sinhalese population.

I cannot credibly attribute the MH improvements to the impact of the intervention, as there was in fact a (non-significant) larger improvement in the MH of control school children, particularly among boys, when compared to the intervention groups. In addition to this, so little of the intervention was delivered in between M1 and M2 that it is unlikely that a significant change could have occurred as a result of the intervention.

As previously mentioned, very few sessions were delivered in the schools. SDP and psychosocial healing literature suggests that the selection, training and development of peer leaders, who can establish safe and rewarding relationships with participants is crucial to achieve MH objectives [231]. In addition, best practice indicates that peacebuilding interventions should be long term, a recent project in Sri Lanka suggested a project cycle of 2 years [84]. An intervention that consisted of two sessions over the space of three months is insubstantial and unlikely to have a lasting effect on MH.

It is worth noting that the second data collection took place at the end of term, just before the start of the summer holidays. This may have had an impact on the MH of the participants, as they were anticipating summer holidays, whereas in March they were just about to begin their summer term.

MH improved in all groups except the Sinhalese population. The MH of the Sinhalese population remained constant over the study. It is conceivable that the students were made aware that they were to receive a sports-based programme, which was subsequently withdrawn after GFP's run in with the Military Police. There is a chance that this 'disappointment effect' may have been a buffer for whatever event caused the MH improvements in the other groups.

Another potential explanation for the improvement in MH is test familiarity. Participants altered their behaviour in completing the SLIPS-GFP. However, this does not explain the fact that this learning effect was not witnessed during the pre-testing of the measure.

Physical health

Comparing performance with global norms (CHAPTER 6) indicated that there was considerable scope for improvement in CF in the sample. Despite this, there was no change in the CF of the intervention or control groups. This is unsurprising given the nature of the intended intervention activities, and the amount of the intervention that was actually implemented. The session I observed was low intensity and only involved a few participants at a time, there was no indication that the participants were exerting themselves. The control group outperformed the intervention group at M2, largely as a result of the contribution of the control group girl's improvement in CF.

CF in boys decreased significantly, in both intervention and control groups. This suggests a community wide decline in boy's fitness, regardless of intervention status.

However, the performance in the MFT of the Sinhalese boys dropped by two levels from M1 to M2, did not decline in Muslim boys and only marginally decreased in Tamil boys. The performance of the Sinhalese boys is dragging down that of the Tamil and Muslim boys when the groups are considered as a whole. An explanation for this might be that the Sinhalese group were tested on an extremely hot day, at 2pm. The research team could not assemble early enough in the morning due to other commitments, and problems we were having in accessing the school after the fallout between the CDO and the school.

There was large loss to follow up for the MFT, particularly among the girls. Although there was no difference in CF at M1 between girls who did and didn't have follow up data, there may have been an attitudinal difference. Those girls who participated in M2 may have held a different attitude towards the testing and competed more fiercely. Many girls refused to participate in the beep test, particularly in the Muslim community due to the start of Ramadan.

The small, limited improvements in BFA across all groups may be the result of several factors. This might be the result of measurement error as different weighing scales were used at M1 and M2. The set used at M1 were broken by the CDO during routine use in other programmes. As a result, the M2 data collection team used a different set of scales. Given that these were regular

bathroom scales, there is always a limited amount of variation in accuracy. The alternative is that there was a genuine increase in the BFA scores for all participants. To me this seems unlikely, given the fact that June is nearing the middle of the dry season in the North East of Sri Lanka, and food is generally less readily available during this period than in March.

7.9.5 Implications

The quality of the study and the conclusions that can be drawn were compromised by the lack of implementation of an intervention. As the dose of the intended intervention received was insignificant, I cannot comment on the impact of the intervention. At the very least, the programmes are not negatively impacting on children's MH, which has been recorded in other SDP programmes ^[93]. Despite this, there are implications for the planning, delivery and measurement of SDP interventions and evaluations in this context. Investing the time and resource in the planning and design of interventions is crucial to their success, a theme to which I will return in the wider thesis discussion ^[128].

The rate of loss of volunteers calls into question the sustainability of this method of intervention delivery in this context. GFP claims to have 15% of volunteer staff active at any one point in time, in Mullaitivu it was even less. If constant externally driven training is required, the model is not sustainable, it relies upon the temporary importing of technology and human resources to bolster the supplies.

7.9.6 Future work

The improvements that could be made to this study come under two broad themes. Firstly, there are improvements that could be made to the evaluation design. Secondly, there are some crucial changes that could and should be made to the intervention design and implementation process. I will return to the process of planning and delivering the intervention in the wider thesis discussion.

Increasing the sample size would ensure that the statistical tests used in this study were of sufficient power. Expanding the number of schools, especially those with Muslim and Sinhalese

students would also improve the sample distribution. Likewise, sample size calculations should be updated to reflect the levels of absenteeism identified in this study.

The design of the study could also be improved if participants could be randomised to intervention and control groups. This is not always possible in practice, but there are examples from SDP literature and MH literature that have used randomisation successfully in post-conflict contexts [95, 232]. Randomisation minimises selection bias, and theoretically ensures that intervention and control groups are as similar as possible to one another. It also limits the capacity for intervention designers to engineer which groups of people receive an intervention.

As only 49% of eligible participants provided follow up data; there is scope for improvement in the data collection protocol. It would be of considerable scientific value to assess the MH of those children who were not attending school in Mullaitivu district. Several individuals in Mullaitivu suggested that, in fact, these children were a higher at risk of MH problems. Other NGOs and organisations have used home-based studies to assess the MH of young people, although this method has its own limitations. A longer period of data collection in schools might capture a broader array of participants, however this must be weighed up against the cost of researchers and participant burden. 'Hut-to-hut' sampling has been used in post-conflict contexts before and would ensure that a broader array of participants was involved in the study. However, it is a much more intrusive method of data collection, and may not be suitable in Mullaitivu.

The use of objective measures for the measurement of PA, such as accelerometers, may have helped to explain and contextualise the CF test results. I chose not to use accelerometers for several reasons, but primarily because this would have involved the importing of technology not available in the local context, and as such would have detracted from the sustainability of the project. In addition, the technology, storage capacity and skills for interpreting the data are not yet established in Mullaitivu. Likewise, since the technology would have been entirely new in this context, there may have been considerable reactivity to the device, which might have detracted from their validity in this context. Further, even if accelerometers were provided for PA assessment their added value to the current project is limited; the scientific community has not established accelerometer cut offs appropriate for the particular population. Further, we do

not know the capacity of accelerometers to identify and measure the types of activities that young adolescent Sri Lankans engage in. Given my experiences in March 2014, I felt that the addition of accelerometers in the follow up measurement would have added un-needed logistical and practical complications for the purpose of the current study.

Anecdotally, it appears that young northern Sri Lankans are not engaged in a lot of moderate to vigorous PA, which may explain the fitness deficit. However, without further evidence no conclusions can be drawn.

The use of internationally validated MH measures would have helped to contextualise the psychosocial health scores of the participants. As the SLIPS-C (and so SLIPS-GFP) has no cut offs that indicate when a participant is at high risk of psychosocial stress, it is difficult to see how it would function well in a clinical setting. However, a study that assessed the association of SLIPS-C score, or the SLIPS-GFP, with known mental illness tests, or positive measures of MH, would be of scientific value.

In theory, the use of qualitative methods would enhance any potential findings from the evaluation. It is not sufficient to simply know that an intervention has an effect, if the results are to be generalised, we need to know the specific mechanisms, processes and relationships that generate positive change. In this case, participant qualitative data was not a major issue, as so little of the intervention was implemented. Qualitative data is crucial in the determining the 'how' and 'why' an intervention has had an effect, rather than the 'what' effect it has had. In this case, informal interviews were essential in establishing the reasons behind the failure of the intervention.

Drawing on my own experiences, I would recommend that any future work in the region should ground itself in a local academic institution. The association that I had with Jaffna University when conducting research in Kilinochchi town was crucial for the establishment of a research team. I recognise that this reduces the opportunities for personal development for those of a disadvantaged background. However, one has to balance the need for methodological rigor against the humanitarian aspirations in evaluation in this context.

SUMMARY BOX

The study identified interesting trend data in a low-resource and complex setting;

There was a community-wide improvement in MH during the period of the study; the only group not to benefit were Sinhalese boys;

CF decreased throughout the intervention, particularly the Sinhalese boys in the intervention group and the control Tamil boys;

The sample distribution and size affected the strength of the conclusions that could be drawn from the study;

The association with a local organisation can be a double-edged sword as it can involve compromising the rigor of a study, but can give a programme much needed local legitimacy;

The study developed local capacity for conducting research and trained several members of the CDO in basic data collection methods;

Future programmes should:

- be wary of local military rules and be context sensitive
- train staff members that are local to the intervention schools
- plan sessions in advance to ensure sufficient staff are available
- form a mutually agreed upon contract for delivery
- discuss and outline a more coherent plan of programme elements to ensure and evaluation is relevant.

Future evaluations should:

- apply methods that optimise evidence generation
- increase the sample size and adjust for a higher rate of absenteeism
- sample evenly from schools of different ethnic groups
- include a process evaluation.

CHAPTER 8

Discussion

8.1 Purpose of chapter

The purpose of this chapter is to discuss the evaluation as a whole, and to discuss the overarching themes. First of all I present a summary of the principal findings, laid out in the order proposed by Bauman and Nutbeam (2014) in their 'rocket model', the model and a short description is attached in APPENDIX: F.1.

This structure provides a good framework with which to compare the actual process of implementation and evaluation. I then discuss the strengths and limitations that span my thesis, and provide suggestions for further policy, practice and research.

I draw on my own experiences in Mullaitivu to interpret the results and indicate what I feel are the implications for the SDP field at large. The chapter closes with a summary of the contributions my work has made to the field and the implementation of SDP programmes in post conflict settings.

8.2 Principal findings

The studies did not follow the sequence laid out by Bauman and Nutbeam (2014) in the ‘rocket model’, as is often the way with health promotion evaluations. Despite this, I present my principal findings and discussion in the order outlined in the framework [179]. In previous chapters I formatted my discussion according to the BMJ framework, this was more appropriate for the constituent studies [182]. I feel the framework outlined by Bauman and Nutbeam (2014) reflects the purpose of the evaluation as a whole and provides a logical sequence for the discussion.

8.2.1 Stage 1: Problem Definition

The purpose of this stage is to identify the magnitude, prevalence and public health impact of a problem. It should include analysis on the potential for intervention, and evidence that the community recognises the problem and gives it priority [179]. I have outlined a summary of the findings from this stage of the evaluation.

Epidemiology and Demography – Chapter 2

- MH disorders are a considerable contributor to the global burden of disease;
- Over 1 billion children live in countries affected by armed conflict;
- Strong evidence suggests an increased rate of mental disorder in post conflict settings, despite issues of measurement and definition;

Social, behaviour and organisational research - Chapter 3

- Sri Lanka endured decades of violent conflict, and Mullaitivu was one of the most affected districts;
- Mullaitivu is neglected, and MH care is of low priority, despite evidence of a considerable MH burden;
- Current MH services and support are insufficient to cover the area appropriately.

Community needs analysis (Baseline study) – Chapter 6

- My results indicate that young adolescents in Mullaitivu have better MH than their regional peers in similarly affected, post conflict district Kilinochchi;
- Young adolescent boys have worse MH than young adolescent girls;
- There is no difference in baseline MH between the ethnic groups in Mullaitivu;
- High rates of absenteeism mean that the sample is unlikely to be representative of the background population; this was also influenced by the selection of the participating schools;
- Young adolescents in Mullaitivu underperform compared to global norms for the MFT, which indicates poor CF;
- Very high proportions of boys and girls are stunted and thin, indicating high levels of current and chronic malnutrition;
- Young adolescent boys suffer from higher levels of malnutrition than girls;
- There is no association between MH and PF for young adolescents in Mullaitivu, which indicates that purely exercise-based MH programmes may be ineffective.

8.2.2 Stage 2: Solution Generation

This second stage of the evaluation ‘clarifies the potential content and methods for the intervention, and further defines the different needs of populations’ [179]. The results should contribute towards the intervention and evaluation objectives, and identify the theoretical links to their achievement. I have outlined a summary of the findings from this stage of the evaluation.

Literature review – Chapter 2

- There is a growing body of evidence for various positive impacts of sport and PA on the physical, mental and social wellbeing of young people, including ‘at-risk’ populations in developed, peaceful settings;

- There is little or no evidence that substantiates these associations in a youth population in a post conflict setting;
- SDP organisations make broad sweeping claims about the benefits for participants in their programmes, which are not backed up with sufficient evidence;
- Published evidence from post-conflict settings is usually grey literature reporting programme reach.

Theory development – Chapter 2

- Theory and mechanisms linking SDP actions, outputs and outcomes, and peace are under-researched and have not been empirically tested;
- The outcomes of peacebuilding programmes are cited as immeasurable and long term, which inhibits meaningful investigation;
- Commonly cited peacebuilding theory is simplistic and does not provide grounds for programme development or impact measurement;
- I have developed a theoretical framework, based on the work of Blagojevic (2012) and a literature review that defines psychosocial health (referred to as MH) as a short-term, operational indicator of peace.

Pre-testing of methods and materials – Chapter 5

- Assessment of MH and PH is pragmatically feasible and ethically viable in Mullaitivu using locally adapted methods;
- The adapted MH assessment tool had very good internal consistency and intra-rater reliability.
- The adapted PH measurement protocols have very good inter tester reliability;
- The measures were not assessed alongside gold standard measures and as such the validity of the adapted protocols is unknown;
- Identifying, training and retaining a research team in remote and rural conditions is challenging, and has implications for the quality of the data collection procedure;

- Context should be a main driver in the selection of assessment methods. Support from all levels of administration was crucial in effective implementation of the measurement protocols in Mullaitivu.

8.3 Stage 3: Innovation testing

The testing of a successful, sustainable solution is the aim of stage 3 of the 'rocket model'. The focus in this stage of intervention development should be whether the intervention 'worked', whether it achieved the intended objectives in an optimised context. However, the intervention design required that schools were sampled from multiple regions of Mullaitivu and the intervention delivery conditions were non-optimal. I have outlined a summary for this stage of the evaluation.

Impact and outcome assessment – Chapter 7

- I can draw limited conclusions about impact of the GFP intervention, as it was implemented to a negligible degree;
- The intervention at least 'did no harm' and successfully advocated the use of sport for peace and development in the region;
- The GFP intervention cohort did not experience MH improvements that exceeded those of the control group;
- There was a community wide decrease in boys CF;
- CF increased for girls in the control group;
- The lack of process data inhibited any assertions about the 'dose' of the intervention received, although this is not problematic given the minimal implementation;
- There were very high rates of absenteeism and loss to follow up, despite sampling each school twice to minimise this problem;
- I observed that the programmes were not heavily aerobic and may have been insufficient to cause change in CF, although this does not explain the reduction in boys fitness;

- The change in MH and PH are unlikely to be the result of random or measurement error, as the methods testing section indicated the reliability and validity of the assessment methods;
- Drawing on the results of the baseline analysis, it is likely that the MH improvements were not the result of increased activity, or the 'sporting' element of the programmes.

8.4 Implications

The implications of the various components of this evaluation are discussed in previous sections. This section will discuss the findings of the study in relation to the research objective and the field of SDP at large. The research objective for the study was:

“To contribute to the evidence base for sport for development and peace programmes on the mental and physical health of adolescents living in a post-conflict region”.

8.4.1 Impact on MH

The results of the evaluation neither support nor disprove the supposition that SDP programmes can have an impact on MH. Previous studies that have assessed the MH impact of SDP interventions have had mixed results [95, 233], and the role of sport and PA in promoting MH in post-conflict contexts remains unclear [91]. Richards et al. (2014) Bolton et al. (2007) conducted RCTs in Northern Uganda. Both studies were of good methodological quality and used a locally validated MH assessment tool. The evaluations reported mixed results: Richards et al. (2014) reported a significant deterioration in intervention boys MH scores, and concurrent improvements in intervention girls and control and wait-listed boys MH scores; Bolton et al. (2007) reported no significant positive findings in the study on either gender. The results of my evaluation are subject to bias, as I have identified in previous chapters, and neither support nor refute the claims that SDP agencies can have a positive impact on MH.

According to local anecdotal evidence, children suffering from MH problems and living in difficult home conditions are less likely to attend school. As a result, they are less likely to engage with school-based SDP programmes. This phenomenon may entrench or exacerbate existing inequalities in MH among children, as the healthy children become healthier, and the rest remain left behind or perhaps deteriorate as seen among boys in the Sinhalese school. It was not possible to assess this hypothesis in this study, but is certainly an important consideration for future research.

Few SDP publications have reported MH findings, despite grandiose claims [91]. The evidence from psychosocial literature supports the inclusion of recreation and creative-expressive therapy elements for conflict-affected children and youth. Further, results from non-conflict affected regions support the inclusion of sport and PA in the promotion of MH and prevention and treatment of mental illness. There is a call for more research into the impact of PA and sports based programmes [70, 71], the role of PA in the amelioration of mental illness or the promotion of MH in post-conflict settings is unclear at this time.

The results of previous studies suggest that the relationship between SDP projects, and psychosocial programmes, and their impact is nuanced and dependent on multiple factors including:

- The relationship between the participants and the programme delivery staff [100];
- The nature of the activities [80];
- The programme context and the relevance of the programmes to the needs of the participants [9, 87];
- The creation of a safe and neutral space in which participants feel comfortable and happy [108];
- The skills of the delivery staff [79, 82, 84];

See Lyras (2011) for a comprehensive list of important programme elements. According to Hartman (2013), the impact of a sports-based intervention is largely determined by its non-sport components (section 2.7). My systematic review of the grey literature indicated that PA programme components are widespread in psychosocial assistance. However, the measurement of the frequency, intensity and dose of the PA programme elements is poor, as is the corresponding reporting.

Identifying the effective and transposable elements of SDP programmes is the crucial next step. It is commonplace to comment on the shortcomings of anecdotal evidence and to point out the absence of causal evidence needed to justify SDP; as it is to stress the need for better monitoring and evaluation processes [99, 104, 234]. In this study, I aimed to add to the evidence base using

pragmatic methods that could be drawn upon widely within the field of SDP, at limited cost and researcher input. Evaluation of SDP programmes is a complex task. In evaluating the impact of sport, researchers face an array of theoretical, methodological and epistemological challenges. These challenges are amplified when conducting analyses in post conflict settings, [235].

8.4.2 Delivery conditions

The limited delivery of the intervention in Mullaitivu is indicative of an issue that plagues SDP projects in general: how do you know when the conditions are right for intervention? Post conflict contexts are transient, and often politically unstable regions. As was the case in Mullaitivu, snap-elections in Sri Lanka caused heightened military sensitivity, and seemingly insignificant transgressions led to damaging consequences. Establishing suitable conditions for the implementation of sustainable and locally relevant interventions is crucial; in some instances it may be a process of trial and error [104]. The delays in implementation caused by unpredictable political shifts are uncontrollable forces, unavoidable in the implementation of natural experiments [236].

Several SDP publications refer to the complexity of the intervention landscape. Effective planning and flexible funding arrangements, despite being difficult to realise, should empower local stakeholders to ensure that is relevant. An understanding of the local context and what is or isn't feasible at the time is crucial to beginning and sustaining interventions [237].

8.4.3 Stakeholders

GFP commits to the involvement of the local implementing partners in the design, delivery and evaluation of programmes, and at the time of writing provides an extensive framework by which to do so. An uncritical evaluation of their method would reason them a 'reactive' organisation; they co-design programmes to suit the context, rather than impose pretested delivery methods. However, as authors have commented, local partners' perceptions of peace are not completely free of dominant discourses [89, 238]. Do the methods, interventions and evaluation methods chosen really reflect the dominant discourse and narratives of GFP, rather than local perspectives?

As GFP now operates in over 52 countries, this ‘bottom up’ approach has considerable implications for producing widely applicable evaluation results. Independent evaluators are often deployed in the SDP field to contribute to the methods and quality of evaluation [239], and their involvement adds another set of priorities and skills to the process.

Many authors have commented on the complexity of the relationships that exist between funders, evaluators and implementing agencies [80, 88, 98], a phenomenon not confined to the field of SDP [171]. I have outlined a summary of the relationship, with participants at the core, who are bound by their geo-political and socio-cultural context, drawing on the work of Richards (2011).

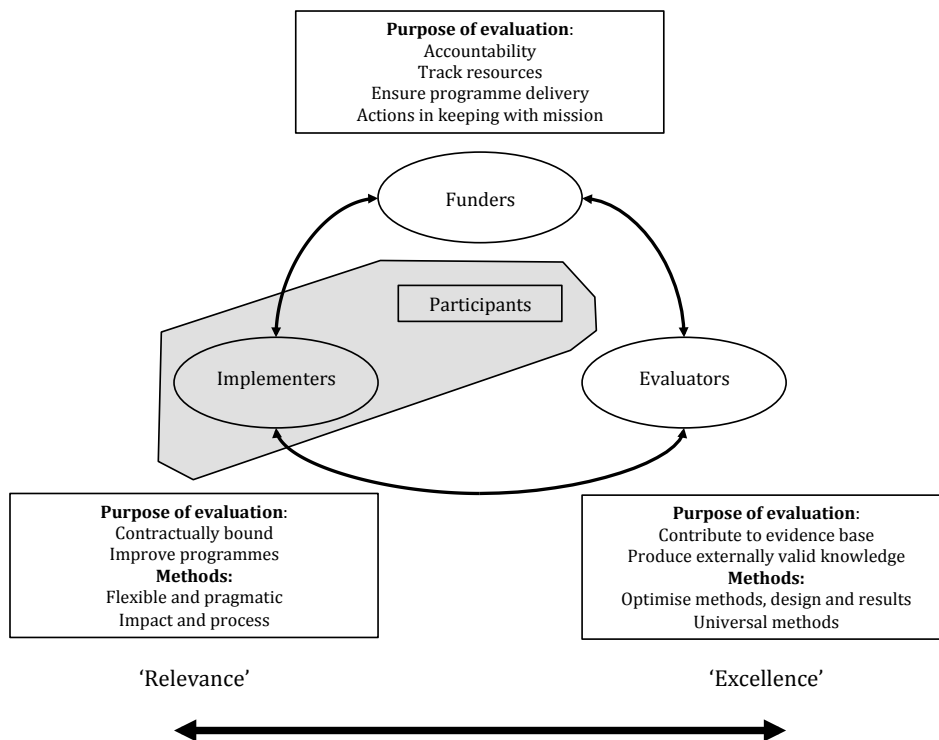


FIGURE 8-1: Diagram indicating the relationships between funders, evaluators and implementing partners. The area highlighted in grey represents the location of a project. Implementing partners and beneficiaries are both acutely affected by the programme location. Drawing on work of Richards and Tol et al.[171]

As indicated in FIGURE 8.1, each stakeholder brings its own set of values, expectations and objectives to an evaluation. The increasing need for accountability in the field of SDP has led to increased demands on implementing partners to provide evidence of programme

implementation and impact [77, 104]. Evaluators are often put under pressure to deliver high quality results in a challenging context, within a limited time frame. Implementers are characterised as seeing evaluation as an adjunct task, necessary for the acquisition of funds, but not central to their core purpose, which is programme delivery [64, 240]. Many authors have stated that the increasing demands placed on implementers by funding bodies are responsible for the overly ambitious claims made in the SDP sector; they are a means by which to obtain funding [77]. Tol et al (2012) suggest that evaluation in these settings is a balance of 'excellence' and 'relevance': practitioners are concerned with 'relevant' information that will contribute to programme delivery, while academics are concerned with 'excellence', the scientific validity of the work and generalizability [171, 241].

GFP encourage the development of reflexive and locally relevant measures for evaluation. In the case of Sri Lanka, they had assured funding without comprehensive formative work. The intervention was designed around simple parameters: it was to be weekly school-based programmes that aimed to increase the contact between the ethnic groups through the use of sports-based games. At the time of the intervention design, there was no associated evaluation. I designed my evaluation around these loosely defined programme elements, seeking to contribute to the evidence base of SDP by addressing key peacebuilding evaluative challenges, which I have described previously.

The CDO and GFP designed an evaluation for the project in my absence. The intended evaluation of the SDP programme was a single item non-validated question that assessed the frequency of an individual's interaction with a child from a different ethnic background. Whether data are observational or based on perceptions, they should be carefully pre-tested to ensure the capture the intended phenomena appropriately [240].

This contrasts with the evaluation that I designed, which consisted of multiple rounds of formative testing and a carefully designed MH impact assessment study. Despite this, my work was not integrated into the programme design and was an adjunct research project. This is indicative of the competing priorities and expectations of the stakeholders; my aim was to obtain the highest level of evidence given the pragmatic restraints of the context, the CDO

developed something within their means and was not resource-intensive. The evaluation and intervention designed by the CDO did not follow logical steps of programme design; there was no formal assessment of needs or methods testing prior to the design of the intervention. Once the intervention was designed, the evaluation was tacked on at a later date.

The aim of formative work is to align the aims and methods of all the parties to ensure a mutually beneficial result ^[179]. My hypothesised priorities for the evaluation did not align with those of the implementing partners, and as a result our work ran in parallel, rather than being mutually reinforcing. Wallis and Lambert warn of 'research parasites', which engage purely on the basis to promote their own academic prestige, rather than to engage with programmes long-term ^[242]. Whilst I recognise this criticism, I believe external evaluators have a role in SDP evaluation, and should invest and engage with local actors to develop the capacity for reflexive and more rigorous evaluation.

The demand for evaluative rigour is influenced by the understanding that SDP projects exist in non-optimal circumstances, which do not lend themselves to standard evaluation techniques that many donors are familiar with ^[242]. GFP are atypical and allow a broad degree of evaluative freedom to their implementing partners. With this breadth comes a new challenge, most importantly this compromises the generalizability of results gathered from multiple projects, in different countries.

SDP funding agencies are forced to balance the need for standardised results, against the need to be site and context specific. This is especially true of peacebuilding agencies, which have the associated difficulty of operationalising measures of peace ^[130, 212]. The involvement of external evaluators should theoretically contribute to the front-end processes of design and planning, if effectively harnessed ^[57]. However, external agents have their own agenda and requirements, which can compromise their adaptability and flexibility ^[242].

Funders and evaluators must consider the consequences of their actions in any post-conflict state; they bring preconceptions of epistemology, acceptability and culture to any project ^[89].

The filming incident in Mullaitivu is indicative of the asymmetrical power relationships that are

often cited in the development field [77, 243]. The CDO require GFP assistance to begin and sustain programmes in Mullaitivu. This unequal power dynamic can lead to the compromising of locally appropriate practice [237]. In this case, the CDO were harassed as they acted unlawfully. Large INGOs and NGO must be conflict sensitive and aware of the implications of their actions for local partners. For local actors, as the 'research context' is a lived reality (indicated in FIGURE: 8.1) [244]. All the involved parties must balance the need to publicise the reality of a repressive context with the need for safety and security, particularly for local partners. Instances like this should encourage more effective conflict analysis and formative work, to ensure that there are no threats to human security.

A key learning point from this process is the need for increased investment in, and consideration of, formative work in the SDP programme planning process. There is a certain degree of groundwork that needs to be completed before roll out, which will be specific to each context and intervention. SDP programmes are complex interventions in difficult settings; they require an in-depth design and planning process to ensure that they are resilient to external threats. GFP and CDO did not invest enough time in the intervention development process, which led to poorly planned and managed programmes and an evaluation that was misplaced. Without sufficient intervention development and planning, the design of an impact evaluation is a superfluous detail, as the risk of intervention delivery failure is so high.

In this situation, the complex interplay between the priorities of the stakeholders and the difficult socio-political context drove this project into an untenable position, and as such a very limited amount of the intervention was delivered according to schedule. In Mullaitivu, a key consideration should have been to ensure that programmes were adequately resourced and sensitive to these political shifts. This would have provided a much firmer footing for the programmes. Failure invest sufficiently in design, and the lack of front end planning process, leads to interventions that are more likely to be inappropriately targeted, have little or no impact, or fail to get off the ground at all [179].

The competing interests and agendas of SDP funders, implementers and evaluators have practical implications for the design of SDP programmes. The fraught relationships identified in

this study are not isolated events, it is reflective of the SDP field [245]. Many SDP projects do not follow sufficient logical planning steps, and as a result cannot necessarily draw on theory to inform their delivery. My literature review indicated a shortage of proper, systematic needs assessment and formative evaluation in the field of SDP [see Chapter 2]. Clearly, there is a need to invest in formative evaluation to ensure that all parties are working in a coordinated manner, as ultimately the participants of the programmes are those who are subjected to poorly researched or managed programmes. To ensure successful intervention dissemination, evaluation must shift from the 'formative to the summative' [230]. However, sufficient formative work must be in place before any consideration of dissemination. Clear, equitable task delegation and communication facilitates a positive and productive outcome for all parties [236].

8.4.4 Theory development

Many SDP organisations state wide-ranging benefits to participation in their interventions, yet the theory and the mechanisms behind achieving the stated goals of such programmes remain under researched and undisclosed: there is a lack of theoretical knowledge on how sport leads to 'peace' [87]. Those theories that do exist have not been empirically tested, and assumptions remain unverified [233]. The effective use of theory can assist in the production of a clearly defined logic model; this in turn provides the basis for the development of indicators and an associated evaluation.

The process by which sport can be co-opted to produce individual level attitudinal and emotional changes in programme participants may well be understood at the GFP senior level. A limitation in their delivery model is how well this information is distributed, internalized and acted upon by their youth leaders. The fidelity of information transfer from their core-training model is unknown. On the other hand, it is unclear whether implementing staff requires detailed theoretical knowledge to facilitate change. But a common criticism of SDP programmes is that delivery staff have a poor knowledge of the mechanisms/impacts they are trying to achieve [233].

Many SDP programmes and associated evaluations are focused on individual-level change, rather than broader social development. Change does not occur in a vacuum; change achieved in

individual level interventions is unlikely to be maintained if not supported by additional changes in the supporting environment [235, 244]. Likewise, sport is only part of the wider social phenomenon that surrounds it [92]. Behaviour change is a complex science that requires considerable skill and theoretical development. Even in disciplines that have an extensive body of research behind them, such as health promotion, practitioners often fail to change behaviour effectively [128].

Sugden states that grassroots programmes can have to a 'ripple effect' and influence the broader social context [106]. The reverse is also true, the 'reverse ripple', the role of social norms and political context in determining behaviour is well documented [246]. Changes in behaviour that occur in a nurturing, safe space such as a SDP programme have no guarantee of replication in the broader social milieu [247]. It is questionable if narrow-scope programming that aims to change attitudes is sufficient to generate broader social change [230]. This is one of the key assumptions in SDP, that individual level change will facilitate change in wider society [79].

Programmes tend to focus on changing the skills, attitudes and behaviours of target groups, without concurrent commitment to changing the institutionalised injustice and socio-political structures that prolong their status as target groups [230]. The means by which to facilitate change in conflict-affected communities may largely be determined by external agencies.

Guilianotti and Armstrong (2014) state that a more influential role of radical social movements and NGOs in SDP should force more equitable relationships and creative partnership between funding bodies and user communities [243]. This may, in turn, contribute to more radical programmes that contribute to wider social change.

I believe that sport will play a limited role in re-defining institutional and systematic inequality: it is a social construct and cannot be the sole 'change agent' in establishing peace [248]. Despite this, major SDP actors should commit to theory driven, reflexive action research that focuses on multiple levels of impact, rather than solely on individual outcomes. A recent systematic review of psychosocial programmes for conflict-affected youth made similar assertions [70]. This is a clear limitation to my own evaluation as I only assessed individual outcomes; I attempted to

address this by the use of a measure of psychosocial well-being, which theoretically is determined by prevailing social and cultural conditions [50].

The theoretical framework for this study outlined assumptions and research questions that could be empirically tested within the timeframe of the study. It also provided a coherent, logical association between psychosocial health and peace: a potential link between individual level changes and broader positive peace [Chapter 2]. My theory and the associated evaluation stages are simplistic, but indicate that such a process is informative. I did not collect information on broader community or social indicators of peace, as it was not within the capacity of the implementing organisation, or within the scope of the project. Despite this, I think there is considerable scope for inter-disciplinary research and evaluation to contribute to the evidence base for SDP.

8.4.5 Indicators

A principal challenge in the evaluation of SDP projects is the identification and operationalization of peace indicators [103]. Identifying reliable, sensitive indicators on the causal pathway towards peace, and the subsequent use of these indicators for programme planning and implementation is critical for improving SDP practice. The lack of consensus on what constitutes peace, or peace indicators, restricts the ability to define and measure it effectively [64]. Where there are disagreements in the fundamental components of peace, it will be very difficult to operationalize the concept in a meaningful manner.

Wallis and Lambert (2014) state as the objectives of SDP programmes are attitudinal or emotional, and that attitudinal changes require a long time to “embed and catalyse impacts on an individual level”, a ‘leap of faith’ is required in peacebuilding evaluation [242].

I disagree with this for two reasons. Firstly, all interventions should at the very least conform to the ‘Do No Harm’ principal. Without measurement, of both intended and unintended consequences, practitioners cannot ensure their efforts are either contributing or undermining local initiatives, policies or practice. Secondly, MH research has successfully operationalised

feelings, attitudes and emotions. To shy away from measurement is to do a disservice to programme participants.

8.5 Strengths and limitations of the thesis

In previous chapters I have detailed the strengths and limitations of my studies. The principal limitation in this thesis is also one of its major strengths: the pragmatic adaptations made to account for the challenges of the research context. In designing a locally appropriate evaluation of an SDP intervention that was based on recognized needs, the availability of resources and the social-ecological context I had to make several methodological concessions, which affected the quality of the results.

The cause of many of the concessions, identified in previous chapters, can be grouped under the following headings:

- conflicting agendas, approaches and priorities of the programme stakeholders;
- socio-political context of the region;
- resources and capacity of the implementing agency;
- access to the local area and data collection sites;
- planning and communication (local and international).

These over-arching factors had an impact on many aspects of the study. Perhaps the most significant factor was the socio-political context, which had far-reaching consequences in the lead up to the planned delivery of the intervention. The limitations in this study can be grouped under several broad themes: evaluation design, measurement and metrics, intervention implementation, recruiting and retaining staff and participants, and the data analysis.

8.5.1 Design

The quality of the evaluation design was compromised by the pragmatic constraints of the research context. I aimed to maximise the level of evidence obtained, but was limited by the capacity and agenda of the implementing partners. Many publications state isolating the effects of a peacebuilding intervention from wider social change as one of the critical evaluative challenges ^[130]. By balancing evaluation design with pragmatic concerns, we can ensure that our conclusions are valid and causality can, in part, be attributed to an intervention ^[240].

I was not able to implement an experimental study design, despite my awareness of the success of other SDP researchers working in similar, post-conflict contexts [95, 249]. The randomisation of schools or participants was not feasible, but my introduction of a delayed intervention control group improved the quality of the results and conclusions. Several authors challenge the suitability of RCT design for post-conflict and low-resource settings due to their expense and complexity [240, 250].

Best practice literature encourages long term follow up to ensure that intervention effects can develop, as different health benefits will take varying amounts of time to materialise [104]. The initial evaluation plan had a 10-month follow up period, but was disrupted. The short time frame is grounds for criticism, and a common pitfall of evaluations. I began collecting impact data too soon after my baseline analysis, Whorley et al (2010) describe a 'shake down period', in which the intervention is finding its feet [240]. I may not have allowed sufficient time for the intervention to settle before collecting follow up data, but had a limited period to conduct data collection before the start of the school summer holidays.

The intervention designed by the CDO and GFP was not specifically designed to improve MH. It was based on broad, non-descript aims that are typical of many SDP implementers [233]. The intervention aimed to break down the fear of the other, develop confidence and introduce members of different ethnic communities to one another. They did not explicitly claim that the intervention would lead to improvement in MH; this was an assumption in my theoretical framework and a weakness in my evaluation design. I could be accused of 'measuring the wrong thing' by the CDO.

SDP and psychosocial assistance best practice advice suggests that all programme elements should have a stated purpose, which is rationalised through formative work and consultation with local partners [162]. In this case, MH improvements were an unintended consequence, rather than a specific objective of the programme.

The divergence from the ideal intervention and evaluation development process outlined by Bauman and Nutbeam (2014) was significant. I designed the evaluation without specific,

detailed knowledge of its constituent parts and intended mechanisms. I had a broad image of what to expect in the intervention, having only had limited exposure to GFP programmes in other contexts; but I had very limited information on the constituent programme elements, how these would lead to the desired changes within the participants, and how these changes would be measured. There was a disconnection between the intervention and my proposed evaluation objectives. The adaptations I made to the evaluation design limited the generalizability of the results for other SDP interventions.

My literature review highlighted a shortage of empirical testing of theoretical assumptions in the SDP sector, bar notable exceptions [87]. My theoretical framework, while being simple, provided feasible and testable assumptions for a SDP programme in a post-conflict zone. The specificity of my study was justified by the increasing recognition of the complexity of conducting SDP evaluations, and the limited research into the effect of sport on MH in post conflict states [183, 251].

My review also identified several reviews of evidence and best practice guidelines for psychosocial assistance to conflict affected populations [69, 70, 90, 252]. Few reviews critically assessed the quality of evaluation design, selection of measures and the use of theory. There were few examples of action research applying the principals outlined in best practice documents. There are examples of high quality studies in the field of psychosocial assistance to conflict affected populations, but these tend to be in clinical settings and therapy-based interventions that do not isolate the impact of sports activities [26, 71, 232, 253, 254]. Despite this, there is a growing recognition for the importance of high quality study design in the field of SDP [98, 255]. My evaluation of the programme methods, community needs and impact were specifically designed for the research context and target population, but perhaps not for the *intervention* itself.

As I have mentioned in previous chapters, the rate of absenteeism and the school selection process are threats to the quality of the results. The CDO selected participating schools, without any objective needs assessment. Selection was largely based on the compliance of school Principals and the need to represent all three ethnic groups in the intervention. As a result, the design of my evaluation was likely subject to considerable selection bias. The high rate of

absenteeism at baseline is a threat to the representativeness of the study. Participants who were absent at baseline were not included in the evaluation; this is a potential source of non-response bias. These two threats to the quality of the results highlight the fact that practical realities often constrain the level of evidence that can be collected.

Teachers and local staff members suggested that children who are absent from school are more likely to suffer from MH problems. These children have often lost multiple family members, and are often in difficult, informal foster care. Previous research in northern Sri Lanka supports these assertions; cross sectional evidence in Sri Lanka has identified associations between exposure to trauma and school absenteeism [158]. Similarly, left-behind children of international labour migrants report higher prevalence of mental illness, and nearly double the rate of malnutrition in children who have both parents living in Sri Lanka [256]. It is entirely conceivable that the absent participants would have had poorer rates of PH and MH than the included participants.

8.5.2 Methods

The quality of results is in part determined by the use of appropriate, valid and reliable measures before and after an intervention. Measurements must be valid and reliable, poor measurement leads to bias and erroneous conclusions about programme impact [240]. Best practice advice encourages the selection of methods that are relevant for funders and local implementing partners [66, 178]. This is another balancing act: the scientifically optimal against the locally feasible.

A key weakness in this study was the limited use of qualitative methods. Peacebuilding interventions are ‘complex interventions’: they consist of multiple components that act both inter-dependently and independently [257]. The use of quantitative methods alone risks being too reductive, and may not capture the important processes of change in a complex intervention. Indeed, qualitative methods can be used to address a number of SDP evaluative and programme design challenges, such as:

- The donor driven nature of intervention and programme design [234].

- The intangible nature of peacebuilding outcomes ^[130].
- The tendency to make deficit assumptions about the intervention target group ^[230].
- The capacity to ensure that programmes are context specific.

The use of mixed methods in study design can contribute towards ensuring that programmes are community controlled, locally relevant and reflect target group voices, especially when participatory methods are adopted ^[104]. Several authors state the value of mixed methods for peacebuilding evaluation, and there is encouraging SDP work using qualitative methodologies elsewhere in Southern Sri Lanka and in other contexts ^[137, 258].

Qualitative methods could have improved the design and evaluation of the Mullaitivu study, particularly during the formative stages. During the design phase, participatory qualitative methods with young people could have contributed to a more comprehensive needs assessment, and been used to identify what young people felt were the causes of divisions between Tamil and Sinhalese youth. Or, alternatively, to explore preconceptions, misunderstanding and fear of the other group. Once the root causes of these issues are understood, programme designers are better placed to design effective and targeted programmes. In theory, the use of qualitative methods in programme design leads to a more appropriate and participant-led design process, which can ultimately lead to more successful and sustainable programmes.

In theory, I could have used qualitative methods to identify 'how' and 'why' the programme had certain impacts, both of which are commonly cited challenges in SDP evaluation. Several authors state the need to identify the mechanisms of change within programmes so that the effective elements can be transposed into other contexts. In Mullaitivu specifically, focus groups or semi-structured interviews with programme staff, teachers and community members might have identified the reasons for staff dropout, volunteers leaving and for the delays in programme development. This information would be of considerable use to local partners, and to GFP more broadly, for identifying the key conditions that must be in place to initiate programmes.

Qualitative methods can also contribute to the process of programme improvement.

Importantly, qualitative methods can be used to critically assess the programmes, relationships and structures that create circumstances, rather than just identifying the circumstances themselves.

Despite the obvious value of qualitative work in this context, it was not feasible in Mullaitivu at the time of the study. In my first fieldwork trip to Mullaitivu and Kilinochchi, I tried one-to-one interviews with young people and sports coaches. This process was extremely fraught: soldiers followed me from site to site, listened into interviews, and spoke to participants after the interviews. This has clear implications for bias, but also is a considerable threat to human security in an area where there are on-going allegations of kidnap and torture at the hands of soldiers. At the time of the study, electronic recording equipment (especially cameras) required a licence, which would have drawn more undesired attention from the military. Given the practical complications in using qualitative methods in this context and the threats to the participant safety, I decided to limit the use of qualitative methods in this particular study. Despite this, I strongly support the use of qualitative methods in the field of SDP. In a field of study that is so strongly bound up in context and meaning, qualitative methods can be a key tool in the design, measurement and evaluation of interventions.

Local norms in planning and organisation, combined with the capacity of the CDO, meant that measurement protocols had to be quick to implement, require minimal training and easily reproduced. As a result, I was not able to use gold standard methods for assessing PH, as they were not feasible in Mullaitivu. I selected field measures that were suitable for assessing CF and BC with limited resources and equipment that had shown previous field reliability [190, 192, 198].

Through successive waves of adaptation and testing, I produced a locally relevant, internally consistent and reliable measure of MH. My selection of an emic MH measure limits the generalizability of the results of the study [259]. In reducing the number of items on the SLIPS-C, I further limited the external validity of the findings. I intended to produce a programme indicator; and I wanted a measure that was more repeatable and reliable than the 49-item SLIPS-C. Retrospectively, it would have been more useful to retain the 49-item version, so that I could conduct more comparative analysis with other regions of Sri Lanka. The clinical relevance

of the SLIPS-C (and consequently the SLIPS-GFP) has not been assessed, which further limits the usability of the findings.

The pre-testing of the measures in Chapter 5 was rudimentary, but was a practical field response to the research context. The validity of the BFA, HFA and MFT has been established, and the adaptations made to the protocols were minor. Reducing the items on the SLIPS-C may have affected the validity of the measure and I was unable to assess concurrent or criterion validity in the field. The threats to validity of the outcome measures limit the strength of the conclusions drawn from the MH and PH data.

Not being able to contextualise the MH scores was a limitation in the study. At M2 I introduced the Subjective Happiness Scale, translated into local language [208]. In doing so I hoped to assess the association between subjective happiness and MH in the target population. This would also give an indication of the predictive validity of the SLIPS-GFP [197]. The internal consistency of the happiness scale was very poor in the target population ($\alpha < 0.3$). During the delivery of the happiness scale, I observed considerable comprehension issues among the participants and the implementing staff. I believe this led to improper understanding of the emphasis of the questions and inconsistent results.

The selection, adaptation and pretesting of methods is a considerable strength in this study. Such extensive formative method testing is rare in the field of post conflict peacebuilding. Strong methodological integrity is crucial in determining the impact of an intervention, this is even more salient for interventions that exist in dynamic political situations such as post conflict zones [240].

8.5.3 Intervention implementation:

The implementation of an intervention is crucial for assessing programme effectiveness, as is systematic measurement of the implementation process [103]. GFP and CDO filming in Mullaitivu started a cascade of implementation issues that ultimately compromised the quality of the evaluation. But this, in turn, was a product of the strict, semi-military rule that exists in much of Northern Province. The CDO being reprimanded for filming is a stark reminder of the power

imbalance in the dominant aid-giving paradigm. Similarly, it is an important reminder that volunteers do not operate in a vacuum, their actions are politically charged and funders should consider their safety of paramount importance.

The shortage of process data was a key finding of the systematic review conducted in Chapter 2, and is a key weakness in this study also. Without process data it is not possible to determine the 'dose' of the programme required to produce a change, likewise we cannot explore the potential mechanisms by which inputs might lead to outputs. I designed a process evaluation, using participatory methods, which would contribute to the delivery of the programme by the CDO, but the idea was abandoned due to the availability of staff members and the poor intervention delivery.

In the event that the intervention was delivered, focus groups with the recipients would have provided insight into the delivery and outcomes of the programme. Interviews with the volunteer coaches who dropped out would have provided valuable information for the CDO and GFP, although I was not able to contact any while in Mullaitivu. Informal conversations with the CDO staff members proved crucial in understanding the failed delivery process. The lack of intervention implementation means that there are few conclusions that I can draw about the impact of the GFP programme on the MH of young adolescents in Mullaitivu.

8.5.4 Staff and participant recruitment and retention:

A key weakness in this study is the high rate of absenteeism and loss to follow up in the main arm of the study. Likewise, the constant turnover of staff was a threat to measurement validity and contributed to bias [260].

In consultation with GFP and the CDO, I selected schools as the location for the data collection. This was a pragmatic decision that aimed to recruit the maximum number of participants and enable me to track them over the course of the intervention. Local sports clubs were suggested as a site to recruit participants for the evaluation. Recruiting at sports clubs would have produced a selection bias. Additionally, association with a club is transient in Mullaitivu; many members fail to pay registration fees and drop in and out of clubs [Informal interviews: KP].

From my perspective, organising successive measurements at sports clubs would have been a logistical and organisational nightmare. Although absenteeism was high in the needs analysis, and loss to follow up in the impact study, I feel that locating the study in schools reduced selection bias and maximised participation.

The schools were pre-selected by the local partners, and as a result there is a chance of selection bias for participation in the study [203]. The uneven distribution of the ethnic groups in the control group was a limitation in the study, but was controlled for statistically in the analyses [203].

The sample size calculation for the impact study took into account the factors that affected subject recruitment. The main study was underpowered, even though I accounted for absenteeism, made multiple data collection site visits, and estimated loss to follow up. Despite efforts, underpowered studies are very common in scientific research [261, 262]. Primarily this was due to the capacity of the local organisation and the very tight margins I indicated in Section 7.8.3.

There are limited human resources from which to recruit a research team in Mullaitivu. Training a data collection team and maintaining their involvement over the course of a project is crucial. It was evident that the CDO was stretched, and the insufficient provision of data collectors contributed to non-optimal data collection procedures.

The programme relied on access to the selected schools and unpaid volunteers to deliver sessions in schools. Delays in delivery meant that many of the staff, which had been trained in 2013, had either taken up full time employment, left Mullaitivu or no longer wanted be involved in the programme. Other evaluations have indicated the difficulty in retaining trained individuals, as skills are in high demand and may lead to paid employment in other NGOs or Government agencies [263].

8.5.5 Data analysis:

There were several limitations to the data analysis that I conducted.

- Using field observations to explain results;
- Sample was both underpowered and I selected a small effect size;
- Multiple testing of outcomes;
- Potential sources of bias;
- Lack of clinical parameters for the effect size.

Personal interpretation of data is not a bias-free method. The limited implementation of the intervention means that my interpretations are of limited importance in this particular study. It is possible that I missed an important confounding variable that explains the results of the study.

Ioannidis (2005) states that small sample size and effect size are threats to probability of identifying true effects in research [264]. Since no previous project has outlined a value for significant change in the SLIPS-C, I followed convention and chose a 5% change in mean value to calculate my sample size. Likewise, as I was constrained in the sample that I could recruit, my sample size was accordingly small. Both of these were threats to the validity of analysis that I conducted, and raised the probability of drawing false conclusions.

The lack of previous research with the SLIPS-GFP meant that the parameters for the effect size had no clinical relevance. As a result, I could not assess whether there were any discernible health implications for the findings of the study.

SDP agencies are criticised for evaluating purely to satisfy the demands of donors [233]. I was at risk of data dredging to satisfy my own agenda for completing my thesis. Assessing multiple outcomes increases the chance that a result will be significant purely due to random variability. Some researchers adopt more stringent p values to account for the increased probability of false negative findings, although this raises the probability of type II errors if the sample size is not increased accordingly [265]. As I was not able to increase the sample size, I selected a primary outcome variable, a method outlined by Zhang et al. (1997) [266].

The greater the flexibility in design, outcomes and definitions used in research, the less likely relationships and results identified are to be true [264]. This is of particular concern to the field of

SDP in which agencies, academics and practitioners evaluate, but do not coordinate their results sufficiently, with lots of opportunity for 're-inventing the wheel'.

8.6 Recommendations: replication, dissemination and sustainability?

The latter stages of the rocket model are replication, dissemination and sustainability ^[179].

During these latter programme stages, focus should shift from assessing outcomes to understanding processes, implementation science and institutionalisation ^[179]. GFP are already at the stage of widespread dissemination and have an institutionalised method of implementation, as the field of SDP is currently a donor-driven entity that is riding on the good will of international bodies. The CDO and GFP, as organisations, are at different stages of maturity and differ considerably in size. The CDO is a small-scale delivery organisation in Sri Lanka that provides a degree of legitimacy to GFP programmes on the ground. GFP, in contrast, is a large-scale organisation that supports and implements peacebuilding training and programmes in 50+ countries.

In this section I will discuss the replication, dissemination and sustainability of the CDO programmes in Mullaitivu. I will also talk about the implications and provide some recommendations for GFP more broadly.

8.6.1 Replication and dissemination of CDO programmes

During the replication phase, developers should aim to reproduce the intervention effects in 'real life' settings. The intervention in northern Sri Lanka was trialled in real life settings from the beginning; it was the first attempt by the CDO to run a peacebuilding programme. At the time of writing, I cannot comment on whether this particular programme could, or should, be replicated in different regions of Sri Lanka. However, there are some small changes that could contribute to the design, implementation and evaluation of the intervention.

The CDO began programmes in March 2015, with a very limited number of volunteers to support delivery. As indicated in the evaluation, delivery on the intended scale was not feasible with their March staff numbers, as only 6/36 sessions were delivered. GFP funded the training of a new group of volunteers in June 2015, to ensure that the programmes could be delivered effectively throughout the year. Despite this, due to weather and staffing problems, programmes

began in Mullaitivu in January 2016. The CDO are conducting routine programme monitoring and a small impact evaluation. The decision to extend the programmes into Kilinochchi in 2016-7 is likely to be based on this small evaluation.

There are several conceivable ways in which the CDO could improve their current practice in Mullaitivu. Recommendations must balance the need for scientific rigour against the realities of the context, taking into account the amount of resource available to the CDO. Mullaitivu is a difficult and changeable context, however there are some simple steps that the CDO can take, supported by GFP, to improve practice and ensure that programmes have a better chance of getting off the ground and running effectively. The CDO can draw on extra resources from GFP, although this is typically for training and mentoring rather than direct financial support. The recommendations below are practical and achievable.

- CDO could identify an 'SDP lead'. The leads' time should be devoted to the planning, implementation and development of programmes in Mullaitivu. GFP can allocate a member of staff on secondment, or can support a CDO representative, but ensuring that the SDP programme is not an afterthought to other day-to-day CDO business is key to sustainability.
- The CDO would ideally outline a clear programme implementation plan. Sessions should be scheduled with schools and the content planned in advance to ensure that it can be delivered appropriately and in a timely fashion. So long as programme implementation remains ad hoc, it will be very difficult to evaluate effectively, or to hold a coherent intervention;
- The CDO and GFP should ensure that programmes can be suitably staffed before starting programme cycles ^[9]. CDO must ensure that they have sufficient staff on the ground to deliver their objectives. GFP have a considerable bank of experience in the design and implementation of SDP projects, they can give a steer on whether the time/resource is right. GFP have now supported an additional volunteer training session to address the staff shortages;

- The CDO should attempt to link up their objectives with broader donors and government programmes and policies. Involving community members, official bodies and schools in the intervention planning and delivery process can help to ensure that programmes are sufficiently staffed and are sustainable. The CDO have successfully engaged religious leaders in the process, and should continue advocating the use of sport for peacebuilding within official channels.
- The CDO could conduct a community baseline study to identify the needs of the local population ^[267]. This might include participatory methods and include the views of young people, who have not previously been engaged. GFP have experience using participatory methods and can support this process. This might also involve conducting a simple analysis of the barriers to positive contact between the different ethnic groups;
- The CDO and GFP should accept that this initial programme is a pilot and test of concept. In Mullaitivu, it is crucial to get broad support for peacebuilding programmes. By confirming that small-scale consistent delivery is possible, the CDO can consider putting in grant applications to GFP at a later date to expand the size of their programmes.

In addition to these practical design and implementation steps, there are a number of changes to the current evaluation plan that would provide much more valuable data. At the moment the intervention is a feasibility trial. However, if successful at this initial juncture, GFP permanent staff or an external consultant could be tasked with including a process evaluation to identify functional intervention elements. At the moment, the CDO does not have this capacity in house. Provided the CDO continue with their current objective, 'to increase contact between Muslim, Sinhalese and Tamil adolescents in Mullaitivu', they must be guided by more comprehensive formative evaluation. The current CDO theoretical framework is loosely based on the 'contact hypothesis', outlined by Allport (1954), which is necessary, but not sufficient, for the development of positive sustainable relationships ^[247]. Superficial multi-ethnic interaction will not lead to successful development of relationships: the use of more sophisticated theory is required, Lyras (2011) provides a comprehensive SDP programme development and

management framework [87]. Sport is only the 'site' of change, not an inherent mechanism of positive change. Only after confirming the feasibility of a long-term intervention, and having confirmed a positive impact on participants, should the CDO aim to reproduce their programmes on a broader scale.

A crucial learning point from this project was the use of a control group in the study. Without the use of a control group, we would have assumed that the project was having positive effect on the mental health of participants. Isolating the impacts of an intervention is one of the key evaluative challenges in peacebuilding. Simple adaptations and consideration of study design can improve the quality of the results obtained from research.

Perhaps the most crucial element for intervention delivery in Mullaitivu will be retaining the volunteer staff throughout 2015-16; the outcomes the CDO are hoping to achieve require careful cultivation and a consistent, sustained relationship between their coaches and participants [244]. The loss of intervention staff is a common theme in SDP programmes [268]. I recommend recruiting volunteers who are embedded in permanent structures within Mullaitivu.

Due to limited delivery of the CDO SDP programmes, there is limited information I can provide to guide the process of dissemination. Only programmes that have been identified to have a positive impact on participants should reach this stage [179]. However, more broadly, as a result of the donor-driven nature of SDP, GFP now delivers programmes in over 50 countries worldwide, despite limited evidence of their efficacy. The CDO programmes only represent a tiny portion of GFP's global activity. Section 8.7 considers some of the challenges facing GFP in establishing the impact of their portfolio of programmes.

8.6.2 Sustainability of CDO programmes?

Shediac-Rizakallah and Bone (1998) outline groups of factors that contribute to the sustainability of health promotion programmes, but the conceptual framework is also of use for analysing the position of the CDO in Mullaitivu. Sustainability refers to the continuation of an intervention, and of its effects [269].

As stated previously, to ensure the delivery of SDP programmes, the CDO must endeavour to advocate widely for peace and embed their programmes into more permanent structures. The CDO are not an exclusively SDP organisation, they are primarily funded to implement income generation support programmes. The results of the baseline analysis indicate a considerable PH deficit in adolescents in Mullaitivu (Chapter 6). These results can be used to lobby for more funding for income support interventions. The analysis suggested a considerable fitness deficit, which could be valuable information for extracting further support from official channels, such as regional and national Sporting, Health and Education bodies. The severe rates of malnutrition in the target population suggest that food insecurity is a concern in Mullaitivu. The CDO can utilise this information to draw on wider stakeholders for programme support.

The reliance on partnership brokering for continuity of funding, and on volunteer effort at the point of delivery are common factors in SDP programmes. Both of these factors are threats to sustainability in the future ^[89].

With regard to continuing the effects of the intervention, the vital next step is to establish the impact and feasibility of the current CDO programmes. If the programmes are deemed a success, CDO must ensure that their results are widely disseminated in official and policy-making circles.

8.7 Generations For Peace

GFP is an umbrella organisation that funds and supports many different types of programmes in an even broader set of post-conflict contexts. I have highlighted the over-arching GFP approach in section 2.10. Their basic model is to identify, train and support youth leaders in conflict-affected countries to run peacebuilding programmes and activities. Their remit is broad, and in many areas requires constant input of funds.

However:

- Assess the impact of the 2015-6 programmes;
- Assess how each programme element contributed to their stated objectives;
- Review the assumptions behind their theoretical framework;

- Conduct a systematically designed community needs analysis in Mullaitivu and any other target group;
- Develop a theoretical framework that accounts for the needs of local population, using the results of the formative evaluation and the results of the 2015-6 evaluation;
- Share this information with GFP mentors and identify commonalities between GFP projects;
- Pilot and assess the intervention in the new setting.

GFP, as with many SDP organisations, use individual level theory to determine community level change [116]. A key assumption in their current programmes is that a school-based intervention might lead to broader social change. This assumption is prominent in the field of SDP and requires higher-level thought, perhaps by the GFP Research Institute.

GFP should focus identifying the successful elements of programmes that should be developed and implemented in other contexts. Each post-conflict context is unique: the relationships and results identified in one setting may not be directly transferrable to another context. As a result, the production of generic principles is not applicable for GFP, who will continue to work in many different contexts. However, through supporting the CDO through this process, they are supporting the chronological and systematic development of a SDP programme, in accordance with best practice guidelines. GFP should invest theory-driven action research to identify working elements of a programme that might be transposable to other contexts.

8.8 Conclusions

The overarching objective of this study was to use public health theory to contribute to the evidence base for sport for development and peace; I feel that the project has advanced knowledge in both paradigms. Drawing on public health theory, I designed a locally appropriate evaluation of an SDP intervention that was based on recognized needs, the availability of resources and the social-ecological context. To do this, I selected a locally constructed psychosocial health assessment tool, adapted it, and conducted rounds of formative testing to assess its appropriateness for the target population, and the intended evaluation design. The theoretical approach behind this evaluation was one of its significant strengths, which was ultimately impeded by the practical implementation of the intervention.

Throughout this thesis I have identified several areas for further development. I have reiterated the challenges inherent in evaluating peacebuilding, but have provided logical and scientifically justifiable initial solutions to them. I have developed a simplistic theoretical framework linking peace and MH; this relationship requires further elucidation and research.

My work has practical implications for SDP practitioners, funding agencies and evaluators. It is a cautionary tale that highlights the complexity of evaluation in post-conflict zones and the relationships required to sustain it. Funders, implementers and evaluators aim to generate new, mutually beneficial knowledge, but this outcome is not inherent in the formation of partnerships.

I have highlighted the importance and feasibility of considering study design in the evaluation of impact in peacebuilding studies, and that this is necessary in future SDP evaluations. There are limitations to the work that I have conducted, but I have streamlined some of the nebulous claims of SDP organisations to practical and measurable metrics.

I have made a contribution to understanding of health priorities in Mullaitivu by collecting novel physical and MH data in a neglected and previously restricted-access region. I worked closely with the University of Jaffna, developing institutional ties, and providing field work experience for medical students. I provided training in basic data collection methods for the assessment of

MH and PH with limited resources, contributing to the research and evaluation capacity of the CDO and the community. My results could be of considerable use to local policy makers, especially District, Zonal and Provincial Officers of Education, who should be made aware of the long-term health implications of poor fitness among students. The study failed to reach non school-attending youth in the region; these people are potentially even more at threat of mental and PH issues.

The clinical significance of the SLIPS-C and SLIPS-GFP is unknown. Despite this, the MH of young adolescents in northern Sri Lanka is worse than that of their southern contemporaries. The results indicate the need for further research into MH in the region, to guide the process of future service development and NGO activity, in accordance with the Sri Lankan Ministry of Health research priorities.

I have drafted a GFP and international dissemination plan for the results of the studies (APPENDIX F.2). I will write a 1-page plain language summary of the results of the study (to be translated into the relevant local language). This will be disseminated within northern Sri Lanka by the CDO to the local School Principals, appropriate Government officials and Education authorities. I have reported verbally to GFP and the CDO, and have planned a longer report that provides more extensive feedback on the process and outcomes of the research. I have presented my results at international conferences in both oral and poster form, and, at the time of writing, have a paper accepted by a peer-reviewed journal that is in press, and another currently under review. I have previously published on the International Platform for Sport for Development and Peace and intend to publish the short GFP report online, given their permission.

Throughout the process I have learnt and grown a great deal as an individual, both in a professional and personal sense. I experienced several setbacks, and I have learnt a great deal about the need for compromise and patience when working in different cultural contexts. I have advanced my knowledge of evaluation, statistical methods and the field of peacebuilding. I have also developed a number of practical skills whilst conducting my fieldwork in Sri Lanka: most importantly the management of programmes, time and other people in a resource-poor, LMIC

setting. In planning, writing and refining this thesis I have developed my writing skills considerably.

I am optimistic about the field of sport for peace and development. Throughout the three years I have been working with GFP, I have seen their approach to programme development and evaluation mature considerably. The truth of the matter is that there are now hundreds, if not thousands, of agencies who are funding, designing and implementing SDP programmes. It requires immense foresight from agencies to invest in prospective peacebuilding research, rather than to implement prematurely. There is a need to expand existing knowledge and to embrace the complexities and creativity required in contemporary peacebuilding practice.

I am immensely grateful for the support I have received from my family, friends and supervisors throughout this process. Ultimately, without the generosity of GFP, and the hard work of the CDO, I would not have had the opportunity to write this thesis.

Bibliography

1. Prince, M., et al., *No health without mental health*. Lancet, 2007. **370**(9590): p. 859-77.
2. World Health Organisation, *Mental Health Action Plan 2013-2020*, World Health Organisation, Editor 2013, World Health Organisation: Department of Mental Health and Substance Abuse.
3. Whiteford, H.A., et al., *Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010*. Lancet, 2013. **382**(9904): p. 1575-86.
4. Murray, C.J.L., et al., *Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010*. Lancet, 2012. **380**(9859): p. 2197-2223.
5. Baxter, A.J., et al., *Global epidemiology of mental disorders: what are we missing?* PLoS One, 2013. **8**(6): p. e65514.
6. Simon, G.E., *Social and economic burden of mood disorders*. Biol Psychiatry, 2003. **54**(3): p. 208-15.
7. Hoffman, D.L., E.M. Dukes, and H.U. Wittchen, *Human and economic burden of generalized anxiety disorder*. Depress Anxiety, 2008. **25**(1): p. 72-90.
8. St John, T., L. Leon, and A. McCulloch, *Lifetime Impacts: Childhood and Adolescent Mental Health: understanding the lifetime impacts*, D.A. McCulloch, Editor 2005, The Mental Health Foundation.
9. Kieling, C., et al., *Child and adolescent mental health worldwide: evidence for action*. Lancet, 2011. **378**(9801): p. 1515-25.
10. Mental Health Foundation. *Mental Health Foundation: Children and Young People*. 2015 [cited 2016 22nd February]; Available from: <https://http://www.mentalhealth.org.uk/a-to-z/c/children-and-young-people>.
11. Knapp, M., et al., *The Maudsley long-term follow-up of child and adolescent depression: 3. Impact of comorbid conduct disorder on service use and costs in adulthood*. Br J Psychiatry, 2002. **180**: p. 19-23.
12. Hsia, R.Y. and M.L. Belfer, *A framework for the economic analysis of child and adolescent mental disorders*. Int Rev Psychiatry, 2008. **20**(3): p. 251-9.
13. Tramontina, S., et al., *School dropout and conduct disorder in Brazilian elementary school students*. Can J Psychiatry, 2001. **46**(10): p. 941-7.
14. Currie, J. and M. Stabile, *Mental Health in Childhood and Human Capital*, in *The Problems of Disadvantaged Youth: An Economic Perspective*, J. Gruber, Editor. 2009, University of Chicago Press: University of Chicago.
15. Naiker, K., et al., *Social, Demographic and Health Outcomes in the 10 Years Following Adolescent Depression*. J. Adolesc. Health, 2013. **52**(2): p. 5.
16. Hauenstein, E.J., *Depression in adolescence*. J Obstet Gynecol Neonatal Nurs, 2003. **32**(2): p. 239-48.
17. Cavanagh, J.T., et al., *Psychological autopsy studies of suicide: a systematic review*. Psychol Med, 2003. **33**(3): p. 395-405.
18. Merikangas, K.R., E.F. Nakamura, and R.C. Kessler, *Epidemiology of mental disorders in children and adolescents*. Dialogues Clin Neurosci, 2009. **11**(1): p. 7-20.

19. Halliwell, E., L. Main, and C. Richardson, *The Fundamental Facts. The latest facts and figures on mental health*, I. Booth, Editor 2007, The Mental Health Foundation: Colechurch House, UK.
20. Flisher, A.J. and A. Gevers, *Adolescence*, in *Promoting mental health in scarce-resource contexts: emerging evidence and practice*, I. Petersen, et al., Editors., HSRC Press: Cape Town.
21. C, K., et al., *Child and Adolescent Mental Health*, in *Global Mental Health: Principals and Practice*, V. Patel, et al., Editors. 2013, OUP: Oxford.
22. Belfer, M.L., *Child and adolescent mental disorders: the magnitude of the problem across the globe*. *J Child Psychol Psychiatry*, 2008. **49**(3): p. 226-36.
23. Patel, V., et al., *Treatment and prevention of mental disorders in low-income and middle-income countries*. *Lancet*, 2007. **370**(9591): p. 991-1005.
24. Patel, V., et al., *Promoting child and adolescent mental health in low and middle income countries*. *J Child Psychol Psychiatry*, 2008. **49**(3): p. 20.
25. Anthony, J., et al., *Promoting Mental Health: concepts, emerging evidence, practice: a summary report*, ed. World Health Organisation and Department of Mental Health and Substance Abuse. 2004, Geneva: World Health Organisation.
26. Barry, M., et al., *A systematic review of the effectiveness of mental health promotion interventions for young people in low and middle income countries*. *BMC Public Health*, 2013. **13**(1): p. 835.
27. Funk, M., et al., *Mental health and development: targetting people with mental health conditions as a vulnerable group*, ed. WHO. 2010, WHO.
28. World Health Organisation, *Social Determinants of Mental Health*, World Health Organisation, Editor 2014, World Health Organisation: Geneva.
29. UNICEF, *Machel Study 10-Year Strategic Review. Children and Conflict in a Changing World*, UNICEF, Editor 2009, UNICEF, Office of the Representative of the Secretary-General for Children and Armed Conflict.
30. Ghosh, N., A. Mohit, and R.S. Murthy, *Mental health promotion in post-conflict countries*. *J R Soc Promot Health*, 2004. **124**(6): p. 268-70.
31. Murthy, R.S. and R. Lakshiminarayana, *Mental Health consequences of war: a brief review of research findings*. *World Psychiatry*, 2006. **5**(1): p. 25-30.
32. Pederson, D., *Political Violence, ethnic conflict, and contemporary wars: broad implications for health and social well-being*. *Soc Sci Med*, 2002. **55**(2): p. 15.
33. Roberts, B. and J. Browne, *A systematic review of the factors influencing the psychological health of conflict affected populations in low and middle income countries*. *Global Public Health*, 2010. **6**(8): p. 814-29.
34. Roberts, B., et al., *Post-conflict mental health needs: a cross-sectional survey of trauma, depression and associated factors in Juba, Southern Sudan*. *BMC Psychiatry*, 2009. **9**: p. 7.
35. Murray, C.J., G. King, and A. Lopez, *Armed conflict as a public health problem*. *BMJ*, 2002. **324** (7333).
36. Silove, D., *The Challenges Facing Mental Health Programs for Post Conflict and Refugee Communities*. *Prehosp Disaster Med*, 2004. **19**(1): p. 6.
37. de Jong, J.T., I.H. Komproe, and M. Van Ommeren, *Common mental disorders in postconflict settings*. *Lancet*, 2003. **361**(9375): p. 2128-30.
38. Mollica, R.F., et al., *Mental health in complex emergencies*. *Lancet*, 2004. **364**(9450): p. 2058-67.

39. Ghobarath, H., P. Huth, and B. Russet, *The post war effect of civil conflict*. Am Polit Sci Rev, 2003. **2**: p. 189-202.
40. Roberts, B., P. Patel, and M. McKee, *Noncommunicable diseases and post-conflict countries*. Bull World Health Organ, 2012. **90**: p. 2-2a.
41. Marshall, S., *Developing countries face the dual burden of disease*. Bull World Health Organ, 2004. **82**(7).
42. Brown, G., A. Langer, and F. Stewart, *A Typology of Post-Conflict Environments*, Centre for Research on Peace and Development, Editor 2011, University of Leuven: Leuven.
43. Junne, G. and W. Verokren, *Post-conflict development: meeting new challenges*. 2005, London: Lynne Rienner.
44. Tol, W.A., J. Bass, and P. Bolton, *Introduction to the Special Issue: Combining qualitative and quantitative research methods to support psychosocial and mental health programs*. Intervention, 2012. **7**(3): p. 5.
45. Ryan, R.M. and E.L. Deci, *On happiness and human potentials: a review of research on hedonic and eudaimonic well-being*. Annu Rev Psychol, 2001. **52**: p. 141-66.
46. Keyes, C.L.M., *The Mental Health Continuum: From Languishing to Flourishing in Life*. J Health Soc Behav, 2002. **43**(2): p. 207-222.
47. Aboraya, A., et al., *The Validity of Psychiatric Diagnosis Revisited: The Clinician's Guide to Improve the Validity of Psychiatric Diagnosis*. Psychiatry, 2005. **2**(9): p. 48-55.
48. Gureje, O. and D.J. Stein, *Disorders, Diagnosis and Classification*, in *Global Mental Health: Principles and Practice*, A.V. Patel, et al., Editors. 2013, OUP: Oxford.
49. Miller, K.E. and A. Rasmussen, *War exposure, daily stressors, and mental health in conflict and post-conflict settings: bridging the divide between trauma-focused and psychosocial frameworks*. Soc Sci Med, 2010. **70**(1): p. 7-16.
50. Psychosocial Working Group, *Psychosocial Intervention in Complex Emergencies: A Conceptual Framework*, P.W. Group, Editor 2004, Psychosocial Working Group.
51. Fernando, G.A., *Assessing Mental Health and Psychosocial Status in Communities Exposed to Traumatic Events: Sri Lanka as an Example*. Am. J. Orthopsychiatry, 2008. **78**(2): p. 229-239.
52. Papadopolous, R.K., *Refugees, trauma and Adversity-Activated Development*. European Journal of Psychotherapy and Counselling, 2007. **9**(3): p. 301-12.
53. Bracken, P.J., J.E. Giller, and S. D., *Psychological Responses to War and Atrocity: The Limitations of Current Concepts*. Soc Sci Med, 1995. **40**(8): p. 1073-1082.
54. Somasundaram, D., *Collective trauma in the Vanni- a qualitative inquiry into the mental health of the internally displaced due to the civil war in Sri Lanka*. Int J Ment Health Syst, 2010. **4**(22).
55. Steel, Z., et al., *Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis*. JAMA, 2009. **302**(5): p. 537-49.
56. Neuner, F. and T. Elbert, *The mental health disaster in conflict settings: can scientific research help?* BMC Public Health, 2007. **7**.
57. Inter-Agency Standing Committee (IASC), *IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings*, ed. IASC. 2007, Geneva: IASC.
58. Betancourt, T.S., et al., *Interventions for children affected by war: an ecological perspective on psychosocial support and mental health care*. Harv Rev Psychiatry, 2013. **21**(2): p. 70-91.

59. Saxena, S., G. Thornicroft, and M. Knapp, *Resources for mental health: scarcity, inequity and inefficiency*. Lancet, 2007. **370**(878-889).
60. Blagojevic, B., *Sport and Peacebuilding. Healing the Wounds of War*, in *Sport, Peace and Development*, K. Gilbert and W. Bennett, Editors. 2012, Common Ground Publishing: Chicago, Illinois.
61. Lederach, J., *The Moral Imagination: the art and soul of building peace*. 2005, Oxford: OUP.
62. Galtung, J., *Violence, Peace and Peace Research*. J Peace Res., 1969. **6**(3): p. 167-191.
63. Schirch, L., *The Little Book of Strategic Peacebuilding*. 2004, USA: Skyhorse Publishing.
64. Development Assistance Committee (DAC), *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities*, Organisation for Economic Co-operation and Development (OECD), Editor 2008, DAC: OECD.
65. Purgato, M., et al., *Psychosocial interventions for children exposed to traumatic events in low- and middle-income countries: study protocol of an individual patient data meta-analysis*. Syst Rev, 2014. **3**(34).
66. Gschwend, A. and U. Selvaraju, *Psycho-social sport programmes to overcome trauma in post-disaster interventions. An overview*, Swiss Academy For Development (SAD), Editor 2006, SAD: Biel.
67. Henley, R., *Helping Children Overcome Disaster Trauma Through Post Emergency Psychosocial Sports Programs: A Working Paper*. , Swiss Academy for Development (SAD), Editor 2005, SAD: Biel.
68. Henley, R., et al., *How psychosocial sport and play programs help youth manage adversity: A review of what we know and what we should know*. Psychiatr Rehabil J, 2007. **12**(1): p. 7.
69. Kalksma-Van Lith, B., *Psychosocial interventions for children in war-affected areas: the state of the art*. Intervention, 2007. **5**(1): p. 14.
70. Jordans, M.J. and W.A. Tol, *Interventions for Children Affected by Armed Conflict: A Systematic Review of Mental Health and Psychosocial Support in Low- and Middle - Income Countries*. Curr Psychiatry Rep, 2016. **18**(9).
71. Jordans, M.J.D., et al., *Systematic Review of the evidence and treatment approaches: psychosocial and mental health care for children in war*. Child Adolesc Ment Health, 2009. **14**(1).
72. Fox, K., *The effects of exercise on self perceptions and self-esteem*, in *Physical Activity and Psychological Well-Being*, S. Biddle, K. Fox, and S. Boutcher, Editors. 2000, Routledge: London, UK. p. 88-117.
73. Biddle, S.J. and M. Asare, *Physical activity and mental health in children and adolescents: a review of reviews*. Br J Sports Med, 2011. **45**(11): p. 886-95.
74. Kay, T. and O. Dudfield, *The Commonwealth Guide to Advancing Development through Sport*. The Commonwealth Guide to Advancing Development through Sport. 2013, London, United Kingdom: Commonwealth Secretariat.
75. Levermore, R., *Sport in international development. Facilitating improved standard of living?*, in *Routledge Handbook of Sports Development*, B. Houlihan and M. Green, Editors. 2013, Routledge.
76. Kidd, *A new social movement: sport for development*. Sport Soc: CCMP, 2008. **11**(4).
77. Donnelly, P., et al., *Sport for Development and Peace: a public sociology perspective*. Third World Q, 2011. **32**(3): p. 589-601.

78. Colater, F., *Sports clubs, social capital and social regeneration: 'ill-defined interventions with hard to follow outcomes'?* Sport Soc: CCMP, 2007. **10**: p. 537-59.
79. Gasser, P. and J. Levinsen, *Breaking Post-War Ice: Open fun Football Schools in Bosnia and Herzegovina*. Sport Soc: CCMP, 2004. **7**(3): p. 457-472.
80. Coakley, J., *Youth Sports: What Counts as 'Positive Development'?*, in *Introduction to Sport and Development*, R. Schinke and S. Hanrahan, Editors. 2011, Routledge: London.
81. Kidd, B., *Cautions, Quesitons and Opportunities in Sport for Development and Peace*. Third World Q, 2011. **32**(3): p. 603-609.
82. Kunz, V., *Sport as a post-disaster psychosocial intervention in Bam, Iran*, Swiss Academy for Development (SAD), Editor 2009, Taylor & Francis
United Kingdom: United Kingdom. p. 1147-1157.
83. Sport for Development and Peace International Working Group, *Harnessing the Power of Sport for Development and Peace: Recommendations for Governments*, ed. Sport for Development and Peace International Working Group. 2008, UNSDOP: United Nations.
84. Wijesekera, R., T. Kirsch, and V. Schnorr *Sport and Play for Interethnic Dialogue among Children and Youth in Sri Lanka. Final Evaluation Report by the Swiss Academy for Development*. Swiss Academy for Development,,, 2012.
85. United Nations. *Sport for Development and Peace: Towards Achieving the Millennium Development Goals*. 2003 [cited March 2014; Available from: <http://www.un.org/themes/sport/reportE.pdf>.
86. Coalter, F., *A wider social role for sport : who's keeping the score?* 2007, London ; New York: Routledge. ix, 206 p.
87. Lyras, A. and J.W. Welty Peachey, *Integrating sport-for-development theory and praxis*. Sport Management Review, 2011. **14**: p. 15.
88. Kay, T., *Developing through sport: evidencing sport impacts on young people*. Sport Soc: CCMP, 2009. **12**(9).
89. Kay, T., *Sport and International Development. Introduction: The Unproven Remedy*, in *Routledge Handbook of Sports Development*, B. Houlihan and M. Green, Editors. 2013, Routledge.
90. Zakus, D., D. Njelesani, and S. Darnell, *The Use of Sport and Physical Activity to Achieve Health Objectives*. Literature Reviews on Sport for Development and Peace, 2007.
91. Hamilton, A., C. Foster, and J. Richards, *A systematic review of the mental health impacts of sport and physical activity programs on adolescents in post-conflict settings*. Journal of Sport for Development, Article in press.
92. Cronin, O., *Comic Relief Review. Mapping the reseach on the impact of Sport and Development interventions*, Comic Relief, Editor 2011: Orla Cronin Research.
93. Richards, J. and C. Foster, *The impact of a sport-for-development programme on the mental health of young adolescents in Gulu, Northern Uganda*. J. Sci. Med. Sport, 2012. **15**(Supplement 1): p. S345.
94. Bolton, P., et al., *Interventions for depression symptoms among adolescent survivors of war and displacement in northern Uganda: A randomized controlled trial*. JAMA, 2007. **298**(5): p. 519-527.
95. Richards, J., et al., *Physical fitness and mental health impact of a sport-for-development intervention in a post-conflict setting: randomised controlled trial nested within an observational study of adolescents in Gulu, Uganda*. BMC Public Health, 2014. **14**: p. 619.
96. Harris, D.A., *Dance/movement therapy approaches to fostering resilience and recovery among African adolescent torture survivors*. Intervention, 2007. **17**: p. 134-155.

97. Kaufman, Z.A., T.S. Spencer, and D.A. Ross, *Effectiveness of sport-based HIV prevention interventions: a systematic review of the evidence*. AIDS Behav, 2013. **17**(3): p. 987-1001.
98. Levermore, R., *Evaluating sport-for-development: Approaches and critical issues*. Prog Dec Stud, 2011. **11**(4): p. 339-353.
99. Spaaj, R., *The social impact of sport: diversities, complexities and contexts*. Sport Soc: CCMP, 2009. **12**(9): p. 1109-1117.
100. Meier, M., *Gender, Equity and Sport*, Swiss Academy for Development (SAD), Editor 2005, SAD: Biel.
101. Coalter, F., *The politics of sport-for-development: Limited focus programming and broad gauge problems*. Int. Rev. Sociol. Sport, 2010. **45**: p. 19.
102. Richards, J. and C. Foster, *Sport-for-Development Programme Objective and Delivery: A Mismatch in Gulu, Northern Uganda*, in *Global Sport-For-Development: Critical Perspectives*, N. Schulenkorf and N. Adair, Editors. 2014, Palgrave-Macmillan UK. p. 155.
103. Special EU Programs Body, *A Monitoring and Evaluation Framework for Peace-building*, in *SEUPB A Monitoring and Evaluation Framework for Peace-Building*, PriceWaterHouseCoopers (PWC), Editor 2007, PWC.
104. Preti, D., *Monitoring and Evaluation. Between the Claims and Reality*, in *Sport, Peace and Development*, K. Gilbert and W. Bennett, Editors. 2012, Common Ground Publishing: Chicago.
105. Mac Ginty, R., *Indicators +: A proposal for everyday peace indicators*. Eval Program Plann, 2013. **36**(1): p. 7.
106. Sugden, J., *Critical left-realism and sport interventions in divided societies*. Int. Rev. Sociol. Sport, 2010. **45**(3): p. 258-272.
107. Coalter, F., *Sport-in-Development: a monitoring and evaluation manual*, U. Sport, Editor 2008, University of Stirling: UK Sport.
108. Ley, C. and M.R. Barrio, *Movement, games and sport in psychosocial intervention: a critical discussion of its potential and limitations within cooperation for development*. Intervention, 2010. **8**(2): p. 14.
109. van Ommeren, M., S. Saxena, and B. Saraceno, *Mental and social health during and after acute emergencies: emerging consensus?* Bull World Health Organ, 2005. **83**(1): p. 71-5; discussion 75-6.
110. Scully, D., et al., *Physical exercise and psychological well being: a critical review*. Br J Sports Med, 1998. **32**(2): p. 111-20.
111. Eime, R.M., et al., *A systematic review of the psychological and social benefits of participation in sport for children and adolescents: informing development of a conceptual model of health through sport*. Int. J. Behav. Nutr. Phys. Act, 2013. **10**: p. 98.
112. Ekeland, E., et al., *Exercise to improve self-esteem in children and young people.*, in *Campbell Systematic Reviews*. 2005.
113. Lubans, D., R. Plotnikoff, and N. Lubans, *Review: A systematic review of the impact of physical activity programmes on social and emotional well-being in at-risk youth*. Child Adolesc Ment Health, 2012. **17**(1): p. 11.
114. Jansen, I. and A. LeBlanc, *Systematic review of the health benefits of physical activity in school aged children and youth.*. Int J Behav Nutr Phys Act, 2010. **7**(40).
115. Ahn, S. and A.L. Fedewa, *A meta-analysis of the relationship between children's physical activity and mental health*. J Pediatr Psychol, 2011. **36**(4): p. 385-97.

116. Patel, A.V., L. Swartz, and A. Cohen, *The Evidence for Mental Health Promotion in Developing Countries*, in *Promoting Mental Health: Concepts, Emerging Evidence, Practice*, H. Herrman, S. Saxena, and R. Moodie, Editors. 2004, World Health Organisation: Geneva.
117. Hurtig-Wennlof, A., et al., *Cardiorespiratory fitness relates more strongly than physical activity to cardiovascular disease risk factors in healthy children and adolescents: the European Youth Heart Study*. *Eur J Cardiovasc Prev Rehabil*, 2007. **14**(4): p. 575-81.
118. Plowman, S.A., *Physical Activity and Physical Fitness: Weighing the Relative Importance of Each*. *J Phys Act Health*, 2005. **2**: p. 143-158.
119. Bonhauser, M., et al., *Improving physical fitness and emotional well-being in adolescents of low socioeconomic status in Chile: results of a school-based controlled trial*. *Health Promot Int*, 2005(0957-4824 (Print)).
120. Aberg, M.A., et al., *Cardiovascular fitness in males at age 18 and risk of serious depression in adulthood: Swedish prospective population-based study*. *Br J Psychiatry*, 2012. **201**(5): p. 352-9.
121. Crews, D.J. and D.M. Landers, *A meta-analytic review of aerobic fitness and reactivity to psychosocial stressors*. *Med Sci Sports Exerc*, 1987. **19**(5 Suppl): p. S114-20.
122. Ruggero, C.J., et al., *Cardiorespiratory Fitness May Help in Protecting Against Depression Among Middle School Adolescents*. *J Adolesc Health*, 2015. **57**(1): p. 60-5.
123. Ortega, F.B., et al., *Physical fitness in childhood and adolescence: a powerful marker of health*. *Int J Obes*, 2008. **32**(1): p. 1-11.
124. Craft, L.L. and F.M. Perna, *The Benefits of Exercise for the Clinically Depressed*. *Prim Care Companion J Clin Psychiatry*, 2004. **6**(3): p. 104-111.
125. Monshouwer, K., et al., *Possible Mechanisms Explaining the Association Between Physical Activity and Mental Health. Findings From the 2001 Dutch Health Behaviour in School-Aged Children Survey*. *Clin Psychol Sci*, 2013. **1**(1): p. 67-74.
126. Babiss, L.A. and J.E. Gangwisch, *Sports participation as a protective factor against depression and suicidal ideation in adolescents as mediated by self esteem and social support*. *J Dev Behav Pediatr*, 2009. **30**(5): p. 8.
127. Bandura, A., *The self-system in reciprocal determinism*. *Am Psychol*, 1978. **33**(4): p. 344-358.
128. Nutbeam, D. and E. Harris, *Theory in a Nutshell. A guide to health promotion theories*, ed. R. MacDonald. 2004: McGraw-Hill.
129. Bhana, A. and A. Govender, *Evaluating interventions*, in *Promoting mental health in scarce-resource contexts. Emerging evidence and practice*, I. Petersen, et al., Editors., HSRC Press: Cape Town.
130. Fast, L. and R. Neufeldt, *Envisioning Success: Building Blocks for Strategic and Peace-Building Impact Evaluation*. *J Peaceb and Dev*, 2005. **2**(2).
131. Perera, H., *Mental Health of adolescent school children in children in Sri Lanka - a national survey*. *Sri Lankan Journal of Child Health*, 2004. **33**: p. 3.
132. Siriwardhana, C., et al., *Prolonged internal displacement and common mental disorders in Sri Lanka: the COMRAID study*. *PLoS One*, 2013. **8**(5): p. e64742.
133. Jayasuriya, D., et al., *Associations of mental distress with residency in conflict zones, ethnic minority status, and potentially modifiable social factors following conflict in Sri Lanka: a nationwide cross-sectional study*. *Lancet Psychiatry*, 2016. **3**(2): p. 145-153.
134. Rodrigo, C., et al., *Symptoms of anxiety and depression in adolescent students; a perspective from Sri Lanka*. *Child Adolesc Psychiatry Ment Health*, 2010. **4**: p. 10.

135. Northern Provincial Council, *Northern Provincial Council Statistical Information - 2013*, 2013, Provincial Secretariat: Office of the Deputy Chief Secretary - Planning.
136. United Nations Development Agency, *Human Development Index Report*, 2013: Geneva.
137. Schulenkorf, N., *Sports events and ethnic reconciliation: Attemption to create social change between Sinhalese, Tamil and Muslim sportspeople in war-torn Sri Lanka*. JoHLSTE, 2011. **45**(3): p. 21.
138. Arunatialke, N., S. Jayasuriya, and S. Kelegama, *The Economic Cost of War in Sri Lanka*. World Dev, 2001. **29**(9): p. 1483-1500.
139. International Crisis Group, *The Sri Lankan Tamil Diaspora After the LTTE*, in *Asia Report N.186*, International Crisis Group, Editor 2010, International Crisis Group: Washington.
140. Balasooriya, A., *Towards Positive Peace in Sri Lanka: Government of Sri Lanka's Post-Conflict Peacebuilding Efforts*. J Int Dev, 2012. **18**(3): p. 13.
141. International Crisis Group, *War Crimes in Sri Lanka*, in *Asia Report N. 1912010*, International Crisis Group: Washington.
142. United Nations News Centre. *Sri Lanka heading in 'authoritarian direction', says UN human rights chief*. 2013 15 April 2015]; Available from: <http://www.un.org/apps/news/story.asp?NewsID=45743-.VS4jdBPF-Sc>.
143. de Jong, K., et al., *Psychological trauma of the civil war in Sri Lanka*. Lancet, 2002. **359**(9316): p. 1517-8.
144. Norwegian Refugee Council, *Sri Lanka: A hidden displacement crisis*, Norwegian Refugee Council, Editor 2012, Internal Displacement Monitoring Centre: Internal Displacement Monitoring Centre.
145. Sooka, Y., *An Unfinished War: Torture and Sexual Violence in Sri Lanka*, 2014, The Bar Rights Committee of England and Wales: The Bar Rights Committee of England and Wales.
146. Northern Provincial Council: Chief Secretary's Secretariat, *Statistical Information of the Northern Province: 2014*, in *Northern Province Statistical Reports*, N.P. Council, Editor 2014, Northern Provincial Council Secretariat: Jaffna.
147. International Crisis Group, *Sri Lanka's North II: Rebuilding under the Military*, in *Asia Report N 220*, I.C. Group, Editor 2012, International Crises Group: Washington.
148. Fonseka, B. and M. Raheem, *Land in the Northern Province: Post-War Politics, Policy and Practices*, C.f.P. Alternatives, Editor 2011, Centre for Policy Alternatives: Colombo.
149. Pettitt, J., S. Dastan, and P. Cartland, *Tainted Peace: Torture in Sri Lanka since May 2009*, F.f. Torture, Editor 2015, Freedom from Torture: UK.
150. Group, I.C., *Sri Lanka: Women's Insecurity in the North and East*, in *Asia Report N 217*, I.C. Group, Editor 2011, International Crisis Group: Washignton.
151. Baiki, A., *Recovery Assistance to Resttled Populations Returning to Mullaitivu and Kilinochchi Sri Lanka: Focus Group Discussion*, UNICEF, Editor 2013, UNICEF: Kilinochchi.
152. Sriskandarajah, V., F. Neuner, and C. Catani, *Parental care protects traumatized Sri Lankan children from internalizing behavior problems*. BMC Psychiatry, 2015. **15**: p. 203.
153. Sriskandarajah, V., F. Neuner, and C. Catani, *Predictors of violence against children in Tamil families in northern Sri Lanka*. Soc Sci Med, 2015. **146**: p. 257-65.
154. Harris, S., *Humanitarianism in Sri Lanka: Lessons Learned?*, Fernstein International Centre, Editor 2010, Tufts University: Tufts University.
155. Siriwardhana, C. and K. Wickramage, *Conflict, forced displacement and health in Sri Lanka: a review of the research landscape*. Conflict and Health, 2014. **8**(1): p. 22.

156. World Health Organisation, *Risks to Mental Health: An Overview of Vulnerabilities and Risk Factors*, 2012, WHO: Geneva.
157. Siva, S., *Sri Lanka struggles with mental health burden*. Lancet, 2010. **375**(9718): p. 1.
158. Siriwardhana, C., et al., *Impact of exposure to conflict, tsunami and mental disorders on school absenteeism: findings from a national sample of Sri Lankan children aged 12--17 years*. BMC Public Health, 2013. **13**(1): p. 560.
159. Siritharan, J. and A. Sritharan, *Post Conflict Sri Lanka: The Lack of Mental Health Research and Resources among Affected Population*. Int. J. Human. Soc. Sci. Res, 2014. **4**(3): p. 151-6.
160. Tol, W.A., et al., *Outcomes and moderators of a preventative school-based mental health intervention for children affected by war in Sri Lanka: a cluster randomised trial*. World Psychiatry, 2012. **11**(2): p. 114-122.
161. Catani, C., et al., *Family violence, war, and natural disasters: a study of the effect of extreme stress on children's mental health in Sri Lanka*. BMC Psychiatry, 2008. **8**: p. 33.
162. World Health Organisation, *Assessing mental health and psychosocial needs. Toolkit for humanitarian settings*, World Health Organisation and UNHCR, Editors. 2012: Geneva.
163. Fernando, G.A., K.E. Miller, and D.E. Berger, *Growing pains: the impact of disaster-related and daily stressors on the psychological and psychosocial functioning of youth in Sri Lanka*. Child Dev, 2010. **81**(4): p. 1192-210.
164. Perera, B., et al., *Depressive Symptoms among Adolescents in Sri Lanka: Prevalence and Behavioural Correlates*. J Adolsec Health, 2006. **39**: p. 2.
165. Child Soldiers International, *Child Soldiers Global Report 2008*, 2008, Coalition To stop The Use of Child Soliders: London.
166. United Nations High Commision for Refugees, *Durable Solutions? A Protection Assessment of Sri Lankan Internally Displaced Persons who have Returned, Relocated or are Locally Integrating ("Tool Three")*, in *Data and Analysis*, United Nations High Commision for Refugees Sri Lanka, Editor 2013.
167. Siriwardhana, C., *Mental health of displaced and returnee populations: Insight from the Sri Lankan post-conflict experience*. Conflict and Health, 2015. **9**(22).
168. Somasundaram, D. and S. Sivayokan, *Rebuilding community resilience in a post-war context: developing insight and reccomendations - a qualitative study in Northern Sri Lanka*. Int J Ment Health Syst, 2013. **7**(3).
169. Siriwardhana, C., et al., *Changes in mental disorder prevalence among conflict-affected populations: a prospective study in Sri Lanka (COMRAID-R)*. BMC Psychiatry, 2015. **15**: p. 41.
170. Husain, F., et al., *Prevalence of war-related mental health conditions and association with displacement status in postwar Jaffna District, Sri Lanka*. JAMA, 2011. **306**(5): p. 522-31.
171. Tol, W.A., et al., *Relevance or Excellence? Setting Resaerch Priorities for Mental Health and Psychosocial support in Humanitarian Settings*. Harvard Rev Psychiat, 2012. **20**(1).
172. Elbert T, et al., *Trauma-related impairment in children: a survey in Sri Lankan provinces affected by armed conflict*. Child Abuse Neglect, 2009. **33**: p. 238-246.
173. Hallfors, D.D., et al., *Adolescent depression and suicide risk: association with sex and drug behavior*. Am J Prev Med, 2004. **27**(3): p. 224-31.
174. WHO, *Preventing suicide: a global imperative*, 2014, World Health Organisation,,: Geneva.
175. Catani, C., et al., *Tsunami, war, and cumulative risk in the lives of Sri Lankan schoolchildren*. Child Dev, 2010. **81**(4): p. 1176-91.

176. Siriwardhana, C., et al., *Reducing the scarcity in mental health research from low and middle income countries: a success story from Sri Lanka*. Int Rev Psychiatry, 2011. **23**(1): p. 77-83.
177. Siriwardhana, C., et al., *An intervention to improve mental health care for conflict-affected forced migrants in low-resource primary care settings: a WHO MhGAP-based pilot study in Sri Lanka (COM-GAP study)*. Trials, 2013. **14**: p. 423.
178. Right to Play International *SDPIWG Literature reviews on sport for development and peace*. SDPIWG Literature Reviews, 2007.
179. Bauman, A. and D. Nutbeam, *Evaluation in a Nutshell: A practical guide to the evaluation of health promotion programs: 2nd Edition*. 2nd Edition ed. 2014, McGraw-Hill Education Australia Pty Ltd: McGraw-Hill Education
180. Department of Census and Statistics, *Enumeration of Vital Events*, D.o.C.a. Statistics, Editor 2011, Ministry of Finance and Planning: Colombo.
181. Guyatt, G., S. Walter, and G. Norman, *Measuring change over time: assessing the usefulness of evaluative instruments*. J Chronic Dis, 1987. **40**(2): p. 171-8.
182. Docherty, M. and R. Smith, *The case for structuring the discussion of scientific papers*. BMJ, 1999. **318**(7193): p. 1224-5.
183. Ford, N., et al., *Ethics of conducting research in conflict settings*. Conflict and Health, 2009. **3**(7).
184. Lukumar, P., et al., *Validity and reliability of Tamil version of Strengths and Difficulties Questionnaire self report*. Ceylon Med J, 2008. **53**(2): p. 48-52.
185. Barenbaum, J., V. Ruchkin, and M. Schwab-Stone, *The psychosocial aspects of children exposed to war: practice and policy initiatives*. J Child Psychol Psychiatry, 2004. **45**(1): p. 41-62.
186. Betancourt, T.S., et al., *Psychosocial adjustment and mental health in former child soldiers-systematic review of the literature and recommendations for future research*. J Child Psychol Psychiatry, 2013. **54**(1): p. 17-36.
187. Miller, K.E., et al., *The Afghan symptom checklist: a culturally grounded approach to mental health assessment in a conflict zone*. Am. J. Orthopsychiatry, 2006. **76**(4): p. 423-33.
188. Bolton, P. and A.M. Tang, *An alternative approach to cross-cultural function assessment*. Soc Psychiatry Psychiatr Epidemiol, 2002. **37**(11): p. 537-43.
189. World Health Organisation Growth Reference Study Group, *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*, 2006, World Health Organisation: Geneva.
190. Public Health England, *National Child Measurement Program Operational Guidance*, Public Health England, Editor 2014, Public Health England: Wellington House.
191. Olds, T., et al., *Worldwide variation in the performance of children and adolescents: An analysis of 109 studies of the 20m shuttle run test in 37 countries*. J Sports Sci, 2006. **24**(10): p. 13.
192. Ortega, F.B., et al., *Reliability of health-related physical fitness tests in European adolescents. The HELENA Study*. Int J Obes, 2008. **32 Suppl 5**: p. S49-57.
193. Richards, J., *Evaluating the impact of a sport-for-development intervention on the physical and mental health of young adolescents in Gulu, Uganda-a post conflict setting within a low-income country*, in *Department of Public Health* 2011, Oxford University: OUP.

194. De Araujo, S.S., et al., *Association between body mass index and cardiorespiratory fitness as predictor of health status in schoolchildren*. Revista Andaluza de Medicina del Deporte, 2015. **8**(2): p. 73-78.
195. Miller, K.E., G.A. Fernando, and D.E. Berger, *Daily stressors in the lives of Sri Lankan youth: a mixed methods approach to assessment in a context of war and natural disaster*. Intervention, 2009. **7**(3): p. 187-203.
196. Fernando, G.A., et al., *The impact of exposure to traumatic events among children in Sri Lanka*, in *American Psychological Association*, G.A. Fernando, Editor 2008: Boston.
197. Golafshani, N., *Understanding Reliability and Validity in Quantitative Research*. Qual Rep, 2003. **8**(4): p. 597-607.
198. Leger, L.A., et al., *The multistage 20 metre shuttle run test for aerobic fitness*. J Sports Sci, 1988. **6**(2): p. 93-101.
199. McVeigh, S., A. Payne, and S. Scott, *The reliability and validity of the 20-meter shuttle test as a predictor of peak oxygen uptake in Edinburgh school children, age 13 to 14 years*. Pediatr Exerc Sci, 1995. **7**(1): p. 69 - 79.
200. Hallgren, K.A., *Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial*. Tutor Quant Methods Psychol, 2012. **8**(1): p. 23-34.
201. National Obesity Observatory, *A simple guide to classifying body mass index in children*, P.H. Observatories, Editor 2011, NHS.
202. Dumith, S.C., D. Van Dusen, and H.W. Kohl, *Physical fitness measures among children and adolescents: are they all necessary?* J Sports Med Phys Fitness, 2012. **52**(2): p. 181-9.
203. Coggon, D., D. Barker, and G. Rose, *Epidemiology for the Uninitiated*. 5th Edition ed. 2009, BMJ: BMJ Books.
204. Strauss, M.E. and G.T. Smith, *Construct validity: advances in theory and methodology*. Annu Rev Clin Psychol, 2009. **5**: p. 1-25.
205. Koepsell, T.D. and N.S. Weiss, *Epidemiologic methods: studying the occurrence of illness*. 2003, Oxford University: OUP.
206. UNICEF, *Inter-Agency Guide to the Evaluation of Psychosocial Programming in Emergencies*, U.N.C.s. Fund, Editor 2011: New York.
207. Claessens, L.F., et al., *Participatory evaluations of psychosocial interventions for children: a pilot study in Northern Uganda*. Intervention, 2012. **10**(1): p. 15.
208. Lyubomirsky, S. and H.S. Lepper, *A measure of subjective happiness: preliminary reliability and construct validity*. SOCI, 1999. **46**(2): p. 137-155.
209. World Health Organisation, *Country Profile: Sri Lanka*, World Health Organisation, Editor 2014: Geneva.
210. Arunatilake, N., et al., *Sri Lanka Human Development Report 2012: Bridging regional disparities for human development*, U.N.D.P.S. Lanka, Editor 2012, UNDP: Sri Lanka.
211. Amarasuriya, H., et al., *Sri Lanka National Human Development Report - Youth and Development*, 2014, United Nations Development Programme Sri Lanka: Sri Lanka.
212. O.E.C.D., *Encouraging effective evaluation of conflict prevention and peacebuilding activities*. OECD J Dev, 2007. **8**(3).
213. Smith, J.J., et al., *The health benefits of muscular fitness for children and adolescents: a systematic review and meta-analysis*. Sports Med, 2014. **44**(9): p. 1209-23.
214. Taylor, P., et al., *A review of the Social Impacts of Culture and Sport*, T.S.I.R. Centre, Editor 2015, The Culture and Sport Evidence Programme.

215. Sport for Development and Peace International Working Group, *Literature reviews on sport for development and peace*, Sport for Development and Peace International Working Group, Editor 2007, University of Toronto.
216. Jewett, R., et al., *School sport participation during adolescence and mental health in early adulthood*. J Adolesc Health, 2014. **55**(5): p. 640-4.
217. SPSS Inc, *SPSS Statistics for Mac*, 2008, Chicago: SPSS Inc.
218. Glass, G.V. and K.D. Hopkins, *Statistical methods in psychology and education*, ed. N. Heights. 1996: Allyn & Bacon.
219. Cohen, L. and M. Holliday, *Practical statistics for students: an introductory text*. 1996, London: SAGE publishing.
220. World Health Organisation. *World Health Organisation Growth Reference Tool*. 2007 [cited 2015 June 2015]; Available from: <http://www.who.int/growthref/tools/en/>.
221. Ramanayaka, R., et al., *Rate of stunting among a sample of postwar resettled families in the Vanni Region: a study from the Mullaitive District*. Ceylon Med J, 2013. **58**(2): p. 1.
222. Hettiarachchi, M., et al., *Prevalence and severity of micronutrient deficiency: a cross-sectional study among adolescents in Sri Lanka*. Asia Pac J Clin Nutr, 2006. **15**(1): p. 56-63.
223. Jayatissa, R. and R.M. Ranbanda, *Prevalence of challenging nutritional problems among adolescents in Sri Lanka*. Food Nutr Bull, 2006. **27**(2): p. 153-60.
224. Haider, R., *Adolescent Nutrition: A Review of the Situation in Selected South-East Asian Countries*, W.H. Organisation, Editor 2006, World Health Organization: New Delhi.
225. Rahman, M.A. and R. Karim, *Prevalence of stunting and thinness among adolescents in a rural area of Bangladesh*. Journal of Asian Scientific Research, 2014. **4**(1): p. 39-46.
226. Akresh, R., P. Verwimp, and T. Bundervoet, *Civil war, crop failure and child stunting in Rwanda*. Econ Dev Cult Change, 2011. **59**(4): p. 161-181.
227. de Wilde, J.A., et al., *Trends in body mass index distribution and prevalence of thinness, overweight and obesity in two cohorts of Surinamese South Asian children in The Netherlands*. Arch Dis Child, 2013. **98**(4): p. 280-5.
228. Rannan-Eliya, R.P., et al., *Trends and determinants of childhood stunting and underweight in Sri Lanka*. Ceylon Med J, 2013. **58**(1): p. 10-8.
229. Levermore, R., *Sport in International Development: Time to Treat it Seriously?* Brown J. World Aff, 2008. **14**(2): p. 11.
230. Coalter, F., *Sport for development : what game are we playing?* 2013, New York: Routledge.
231. Schulenkorf, N. and J. Sugden, *Sport for Development and Peace in Divided Societies: Copoperating for Inter Community Empowerment in Israel*. Eu J Sport Soc, 2011. **8**(4): p. 235-256.
232. Tol, W.A., et al., *School-based mental health intervention for children affected by political violence in Indonesia: a cluster randomized trial*. JAMA, 2008. **300**(6): p. 655-62.
233. Coalter, F., *Sport for Development Impact Study-A research initiative funded by Comic Relief and UK Sport and managed by international development through sport*, Comic Relief, Editor 2010, UK Sport,,: University of Stirling.
234. Haartman, D. and C. Kwauk, *Sport and Development : An Overview, Critique, and Reconstruction*. J Sport Soc Issues, 2011. **35**(284).
235. Kay, T., *Development through sport? Sport in support of female empowerment in Delhi, India*, in *Routledge Handbook of Sports Development*, B. Houlihan and M. Green, Editors. 2013, Routledge.

236. Craig, P., et al., *Using natural experiments to evaluate population health interventions: guidance for producers and users of evidence*. J Epidemiol Commun H, 2012. **66**(12): p. 1182 - 6.
237. Abu-Nimer, M. and M.C. Hallward, *Peacebuilding and Development in Transitional and Post-agreement Contexts: Improving Macro- and Micro- level Interventions*. J Peaceb and Dev, 2007. **3**(3).
238. Strier, R., *Gendered realities of poverty: Men and Womens views of poverty in Jerusalem*. Soc Sci Rev, 2005. **79**(2): p. 344-367.
239. Sugden, J., *The Ripple Effect: Critical Pragmatism, Conflict Resolution and Peacebuilding through Sport in Deeply Divided Societies*, in *Global Sport-for-Development: Critical Perspectives*, N. Schulenkorf and N. Adair, Editors. 2014, Palgrave Macmillan: Basingstoke.
240. Wholey, J.S., H.P. Hatry, and K.E. Newcomer, *Handbook of Practical Programme Evaluation*. 2010, SA: Jossey-Bass.
241. Tol, W.A., et al., *Mental Health and Psychosocial Support in Humanitarian Settings*, in *Global Mental Health: Principles and Practice*, V. Patel, et al., Editors. 2013, OUP: Oxford.
242. Wallis, J. and J. Lambert, *Football for Peace - Reflections from the Field*, in *Global Sport-for-Development: Critical Perspectives*, N. Schulenkorf and N. Adair, Editors. 2014, Palgrave Macmillan: Basingstoke.
243. Guilianotti, R. and G. Armstrong, *The Sport for Development and Peace Sector: A Critical Sociological Analysis*, in *Global Sport-for-Development: Critical Perspectives*, N. Schulenkorf and N. Adair, Editors. 2014, AIAA.
244. Sugden, J., *Teaching and Playing Sport for Conflict Resolution and Co-Existence in Israel*. JOHLSTE, 2006. **41**(2): p. 221-240.
245. Welty Peachey, J.W. and A. Cohen, *Research Partnerships in Sport for Development and Peace*. J Sport Manag, Accepted for publication August 27 2015.
246. Bandura, A., *Social learning theory*. 1977, Prentice Hall: Englewood Cliffs.
247. Briggs, L.R., *An Evaluation of 'Process as Product' in OTI Sri Lanka: A Typology of Community-level Peacebuilding Activities and Programming Recommendations*. J Peaceb Dev, 2007. **3**(3).
248. Gilbert, K. and W. Bennett, *Restructuring Lives through Sport*, in *Sport, Peace and Development*, K. Gilbert and W. Bennett, Editors. 2012, Common Ground Publishing: Champaign, Illinois.
249. Kaufman, Z.A., et al., *Baseline Results of a Cluster-Randomised Trial Assessing the Effectiveness of Sport-Based HIV Prevention in South African Schools*. Sex Transm Infect, 2013. **89**(SUPP1).
250. Clay, R., *More than one way to measure*. Mon Psych, 2010. **41**(8): p. 52.
251. Saxena, S., et al., *The 10/90 divide in mental health research: trends over a 10-year period*. Br J Psychiatry, 2006. **188**: p. 81-2.
252. Ley, C. and M. Rato-Barrio, *Movement, games and sport in psychosocial intervention: a critical discussion of its potential and limitations within cooperation for development*. Intervention, 2010. **8**(2): p. 14.
253. Tol, W.A., et al., *School-based mental health intervention for children in war-affected Burundi: a cluster randomized trial*. BMC Med, 2014. **12**: p. 56.
254. Jordans, M.J., et al., *Evaluation of a classroom-based psychosocial intervention in conflict-affected Nepal: a cluster randomized controlled trial*. J Child Psychol Psychiatry, 2010. **51**(7): p. 818-826.

255. Richards, J. and C. Foster, *Sport-for-development interventions: whom do they reach and what is their potential for impact on physical and mental health in low-income countries?* J Phys Act Health, 2013. **10**(7): p. 929-31.
256. Wickramage, K., et al., *Risk of mental health and nutritional problems for left-behind children of international labor migrants.* BMC Psychiatry, 2015. **15**: p. 39.
257. Medical Research Council, *A framework for the development and evaluation of RCTs for complex intervention to improve health*, MRC Health Services and Public Health Research Board, Editor 2000.
258. Siefken, K., G. Schofield, and N. Schulenkorf, *Process evaluation of a walking programme delivered through the workplace in the South Pacific island Vanuatu.* Glob Health Promot, 2015. **22**(2): p. 53-64.
259. Rasmussen, A., et al., *Comparing the validity of the self reporting questionnaire and the Afghan symptom checklist: dysphoria, aggression, and gender in transcultural assessment of mental health.* BMC Psychiatry, 2014. **14**: p. 206.
260. Kaufman, Z., B.P. Rosenbaur, and G. Moore, *Lessons learned from Monitoring and Evaluating Sport-for-Development Programmes in the Caribbean*, in *Global Sport-for-Development: Critical Perspectives*, N. Schulenkorf and N. Adair, Editors. 2014, AIAA.
261. Turner, R.M., S.M. Bird, and J.P. Higgins, *The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews.* PLoS One, 2013. **8**(3): p. e59202.
262. Lilford, R. and A.J. Stevens, *Underpowered studies.* Br J Surg, 2002. **89**(2): p. 129-31.
263. Desia, V., *NGOs, gender mainstreaming and urban poor communities in Mumbai.* Gend Dev, 2005. **13**: p. 90-98.
264. Ioannidis, J.P., *Why most published research findings are false.* PLoS Med, 2005. **2**(8): p. e124.
265. Feise, R.J., *Do multiple outcome measures require p-value adjustment?* BMC Med Res Methodol, 2002. **2**: p. 8.
266. Zhang, J., et al., *Some statistical methods for multiple endpoints in clinical trials.* Control Clin Trials, 1997. **18**(3): p. 204-21.
267. H, B.-H., *Transporting evidence-based interventions across cultures: using focus groups with teachers and parents of preschool children to inform the implementation of the Incredible Years Teacher Training Programme in Jamaica.* Child Care Health Dev, 2011. **37**: p. 649-661.
268. Schulenkorf, N. and N. Adair, *Global Sport-For-Development: Critical Perspectives.* 2014, Basingstoke, Hampshire: Palgrave Macmillan.
269. Shediak-Rizkallah, M.C. and L.R. Bone, *Planning for the sustainability of community-based health programs: conceptual frameworks and future directions for research, practice and policy.* Health Educ Res, 1998. **13**(1): p. 87-108.
270. Moher, D., et al., *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.* PLoS Med, 2009. **6**(6).
271. Summerfield, D., *How scientifically valid is the knowledge base of global mental health.* BMJ, 2008. **336**: p. 992-994.

APPENDIX A

Appendix A.1 Systematic review protocol and precis

Database Search:

The following databases were searched: "Ovid", "EBSCOHost", "Dissertation Abstracts", "Applied Social Sciences Index and Abstracts" (ASSIA), "Cochrane Controlled Trial Register", "Education Resources Information Centre" (ERIC), "Institute of Development Studies" (accessed via OpenDocs), "International Bibliography of the Social Sciences", "Latin American and Caribbean Health Sciences Literature" (LILACS), "Medicins Sans Frontieres Field Research", "Web of Science".

Manual searches:

"African Journal of Physical, Health Education, Recreation and Dance", "European Journal of Sport and Society", "International Journal of Sport and Society", "International Review of the Sociology of Sport", "Journal of Sport for Development", "Journal of Sport and Health Research", "Journal of Sport and Social Issues", "South African Journal for Research in Sport, Physical Education and Recreation", "Sport, Education and Society", "Sport in Society: culture, comment, media and politics", "Intervention", "Journal of Health and Sports Science and Child and Adolescent Mental Health".

List of Humanitarian and Grey Literature websites searched

<http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters>

<http://www.opengrey.eu/> (SIGLE)

www.opendoar.org

<http://ssrn.com/>

Web Of Science Conference Proceedings Search

National Registry of Evidence based programmes and practices

PROSPERO

WHO ICTR

Forced Migration Online

WHO

UNICEF

[UNHCR]

International Platform of Sport for Development and Peace

ReliefWeb

United Nations Development Program (UNDP)

United Nations Office on Sport for Development and Peace (UNOSDP)

Right to Play

Swiss Academy For Development

Systematic review search terms

TABLE A 1: Systematic review search terms

POP	[ADOLESCENT]	adolescen*, teen*, youth*, young people, boy, child*, college*, girl*, juvenile*, kid, kids, pube*, school*, student*, young male*, young female*, young person, young men, young people, youngster*, young women
	[CONTEXT – POST CONFLICT]	conflict* OR post-conflict OR peace OR post-war OR conflict affected OR conflict induced OR internally displaced OR displaced persons
INT	[SPORT FOR PEACE]	sport* OR football OR hockey OR tennis OR cricket OR athletics OR rugby OR tag rugby OR swimming OR volleyball OR netball OR basketball OR dodgeball OR boxing OR martial arts OR dance OR baseball OR softball OR rounders OR physical activity OR physical exercise OR recreation OR play OR run OR running
COMP	Any comparator	No search terms included
OUT	Any mental health outcome	No search terms included

Systematic review precis

I conducted a systematic review on the mental health impacts of programs for adolescents in post-conflict settings, which had clearly defined intervention sport or PA elements. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist, was registered on Prospero (CRD42014010833) and was published in the Journal of Sport for Development [91, 270].

A total of 11,722 publications were initially identified, of which 3 met the inclusion criteria and were included in a narrative synthesis. The studies were heterogeneous according to population, outcome, context and intervention method. The principal findings of my review failed to support the broad claims of the positive effect of sport and physical activity programmes on the MH of conflict-affected youth. However, the paucity of available high quality information and the heterogeneity of intervention methods meant that it was not possible to draw a firm conclusion either way at the time of the review. A notable strength of two included studies was the use of a locally developed and validated MH assessment tool.

Despite the equivocal outcome of the review, I identified some crucial knowledge gaps and pieces of best practice advice in the field of SDP. Common limitations included:

- short study duration and follow up;
- poor or unreported adaptation of methods;
- a lack of treatment mechanisms research.

Best practice advice encourages both interventions and associated evaluations to be long-term to observe intervention effects over time. Nesting studies in more permanent structures, such as schools, enables the tracking of participants over time for repeat measures to strengthen study designs and the evidence base.

The review process indicated that participant recruitment, engagement, and tracking can be difficult in such settings. This can impede the development and management of evaluations and complicate long-term analyses. Formative and process evaluations are crucial as they help ensure the success of an intervention and provide valuable learning lessons for future practice. Formative

evaluation should involve the pre-testing and adaptation of methods to ensure their cultural appropriateness and feasibility in the specific research context.

The review highlighted weaknesses in the SDP field at large, in which the majority of evaluations are of poor design, and typically implement pre-post methodologies without control population. There is a shortage of contextually relevant evidence for the use of sport in the assistance of conflict-affected populations. The SDP evidence base would benefit from more detailed investigation of how intended intervention effects are achieved and maintained over time. Such evaluations will allow the identification of the 'active' elements of programmes that can be bolstered and developed. However, in order for this to happen, investment in evaluation is needed at all stages: formative, process and impact.

The review process was invaluable in guiding the design of my study; the selection, adaptation and refinement of my outcome measures; and is one of the major outputs of my thesis,

Appendix A.2 Lederach's Web-building theory

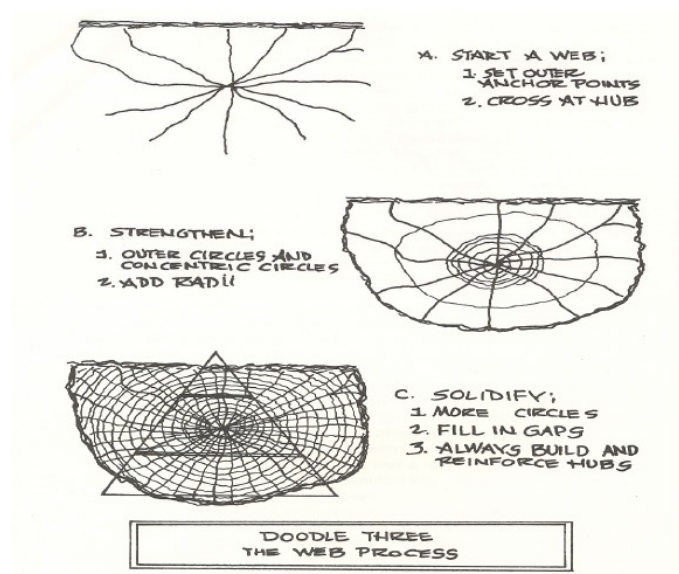


FIGURE A 1: Lederach's Web-building theory for peacebuilding

Lederach envisions a 'web-based' approach to peacebuilding. The above diagram represents a social web; the radiating strands represent relationships between individuals, social groups and other actors. The aim is to develop a network with as many anchorage points as possible, that can sustain damage without compromising the integrity of the web of relationships: social relationships must supersede the challenges of conflict and root causes of violence. This theory places social relationships as the centre of the peacebuilding agenda, and recognises the important role NGOs can play as mid-level actors.

Appendix A.3 Sugden's Ripple Effect Model

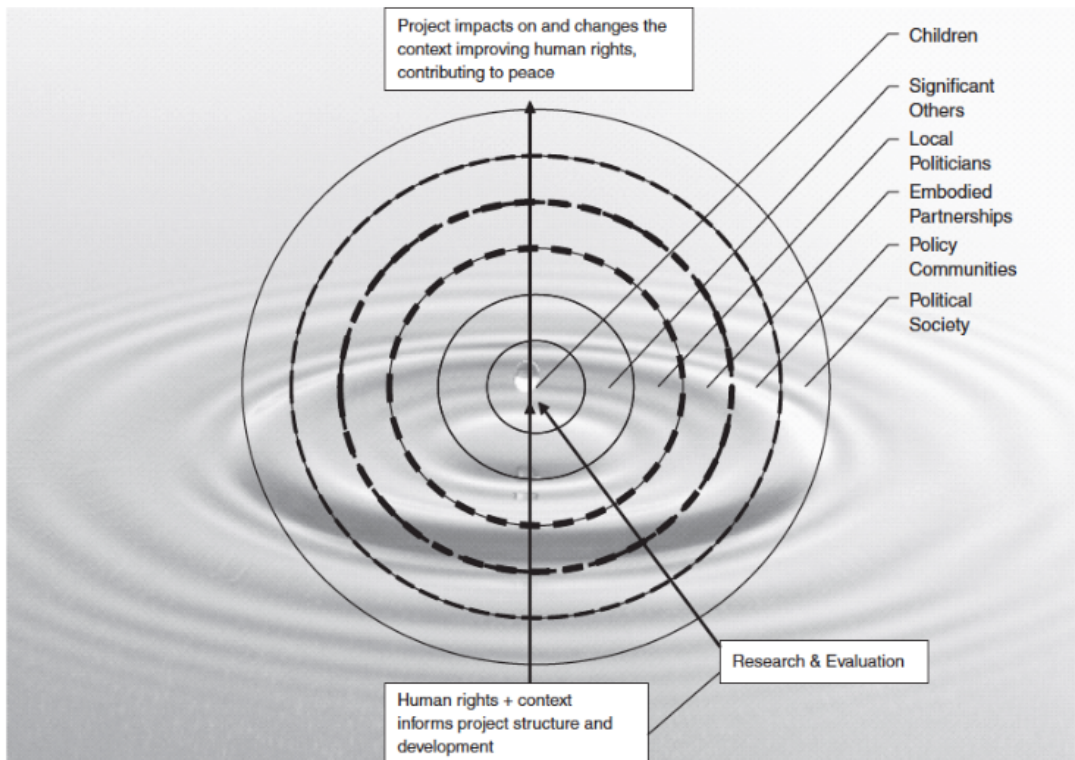


FIGURE A 2: Sugden's Ripple Effect Model

The Ripple Effect model identifies how peace activists can join up grassroots and civil society interventions and activities with the broader elements of political society. It is a structural figure that represents a fluid and dynamic process. However, it does not provide any advice on important actors or barriers in the peacebuilding process.

Appendix A.4 Coalter's logic model for SDP programmes

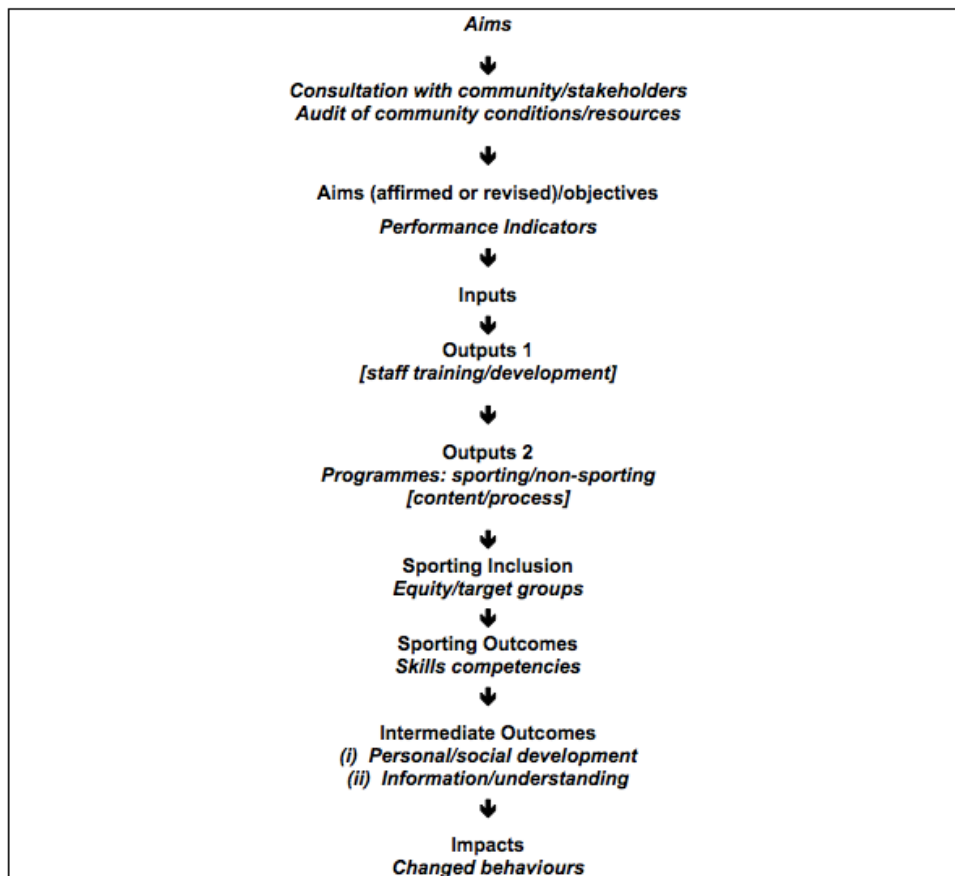


FIGURE A 3: Coalter's logic model for SDP programmes

Appendix A.5 A typical GFP programme

GFP provide sample games and a framework for how sessions should be planned and delivered, although encourage local adaptation of specific program elements. The learning materials provided at the basic GFP training in Amman cover 5 sports and provides sample games and learning objectives for several sessions. GFP loosely encourage the following framework for their programmes:

- General discussion about the session
- Warm up
- Sports based games with modified rules
- Cool down
- Discussion of specific aim of session highlighting learning points

An example game is a basketball game 'catch the dragon's tail'.

Learning objective: to understand that it is important to know your teammates well enough to complement one another

Activity goal: to communicate as a team and protect the tail while taking the tail of other teams

Activity description: players are divided into three (or more) teams and the players line up one behind the other and hold onto the shoulders of the individual in front of them. A strip of cloth is placed in the pocket of the individual at the back of the line. Make sure enough of the tail is exposed to grab the tail.

Players cannot hold their tail or push away people trying to grab their tail. All dragons must stay connected. If a team breaks apart or if a team's tail is taken, then they leave the play area and perform an exercise, such as tuck jumps or one lap around the area.

Load the activity:

- Make the dragon's longer requiring better communication to ensure it doesn't break apart
- Add an element in which people can shoot for hoops for extra points – encourage identifying strong shooters/grabbers

Debrief:

Did any of the team break apart in the activity?

Why might this have happened?

How important is it to know your teammates? How does knowing your teammates help your team?

- When we know each other and we know our strengths and weaknesses and we can communicate more effectively.

APPENDIX B

Appendix B.1 Map of Sri Lanka's Northern Province

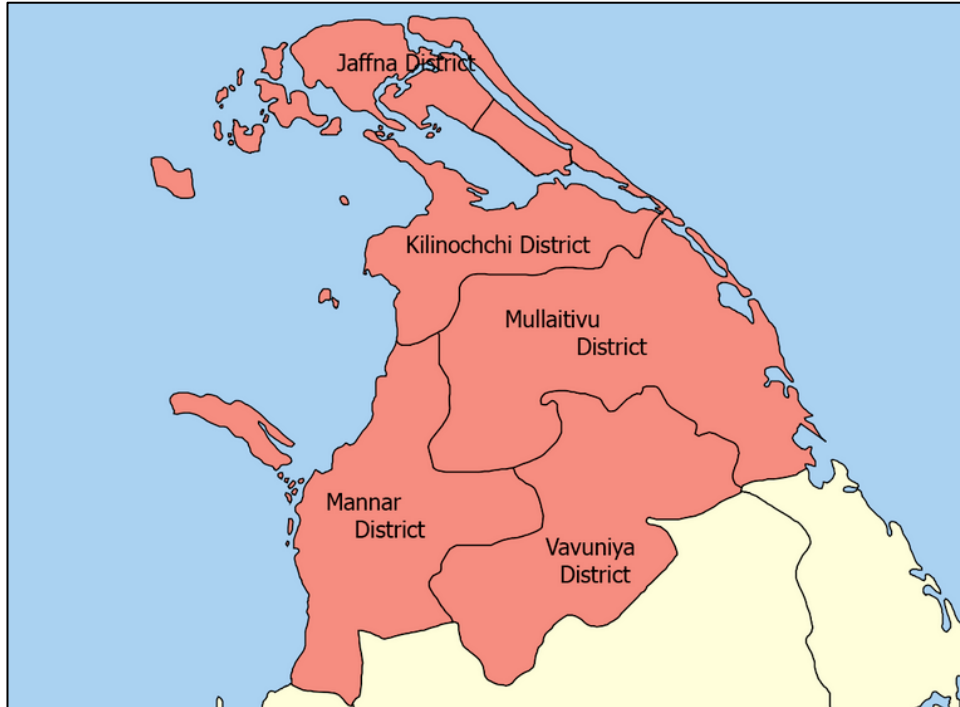


FIGURE A 4: Map of Northern Province

Appendix B.2 Mullaitivu district with each division highlighted.

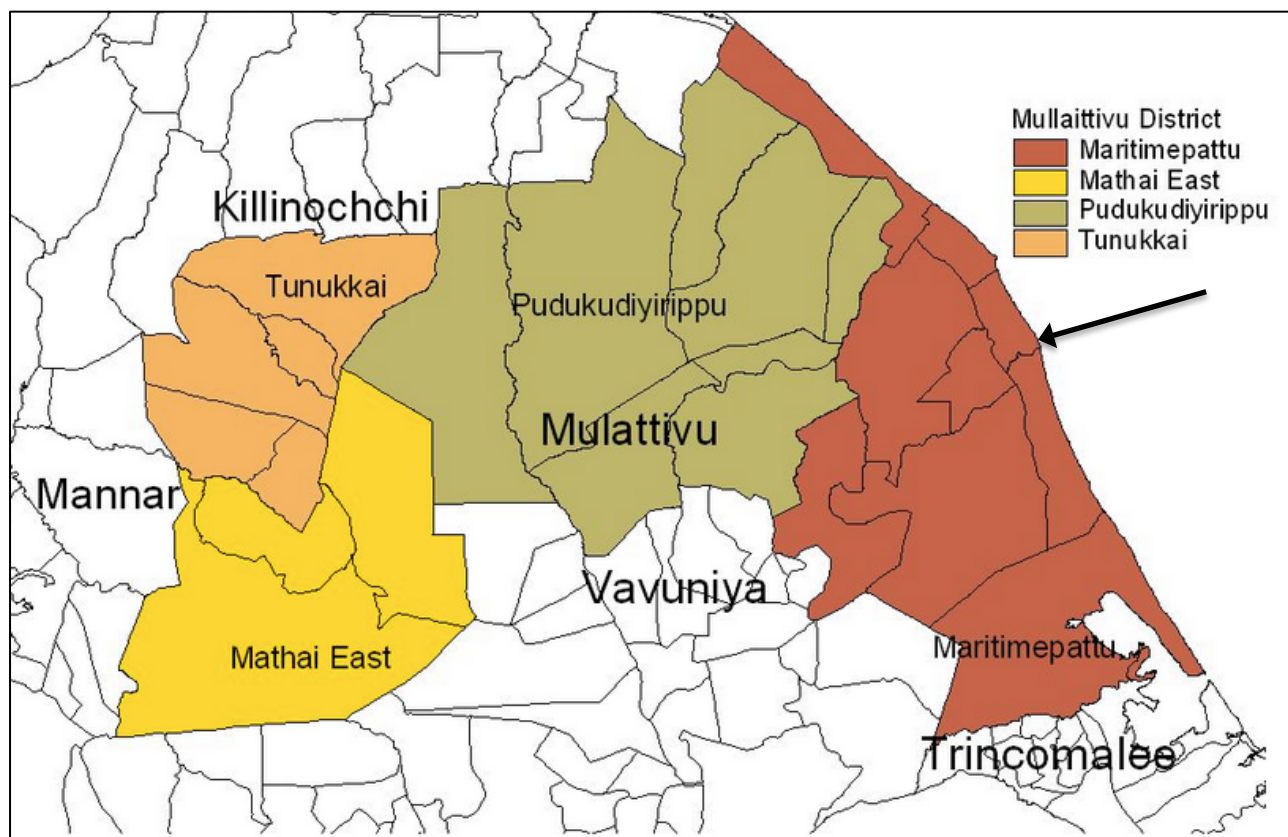


FIGURE A 5: Map of Mullaitivu district. Maritime Pattu is highlighted in red. The black arrow indicates the location of Mullaitivu town.

APPENDIX C

Appendix C.1 SLIPS-C ORIGINAL VERSION (MH-1)

My Current Feelings, Thoughts and Behaviors

Think about what you have been feeling, what you have been thinking, and how you have been behaving for the past TWO WEEKS. Then, for each of the items below, please indicate how often you have experienced each event in the last TWO WEEKS. There are no right or wrong answers, only your own feelings, thoughts, and experiences.

Think about what it means to say “a little” or “often” – different children might have a different view of it. So it might be a good idea to think about a calendar, and see what “a little” or “almost always” might mean.

Never (0 days a week)							
	Mon	Tue	Wed	Thurs	Fri	Sat	Sun
Week1							
Week2							

Rarely (1 day a week)							
	Mon	Tue	Wed	Thurs	Fri	Sat	Sun
Week1	X						
Week2			X				

Sometimes (2-3-days a week)							
	Mon	Tue	Wed	Thurs	Fri	Sat	Sun
Week1	X			X			
Week2	X	X				X	

Often (4-5 days a week)							
	Mon	Tue	Wed	Thurs	Fri	Sat	Sun
Week1	X	X		X		X	
Week2	X		X		X	X	X

Almost all the time (6-7 days a week)							
	Mon	Tue	Wed	Thurs	Fri	Sat	Sun
Week1	X	X	X	X	X	X	X
Week2	X	X		X	X	X	X

Or, you could think of it as a basket of fruit:



Now, how would you answer this question?

In the last two weeks, I smiled:

0-days a week 1 day 2-3 days 4-5 days a week 6-7 days a week
 (Never) (Rarely) (Sometimes) (Often) (Almost always)

If you smiled every day, circle "Almost always", but if you smiled only once in those TWO WEEKS, you would circle "Rarely".

Now, answer the following questions. Remember, there are no right or wrong answers, we only want to know how YOU felt, thought, and behaved in the last TWO WEEKS. If you forget what words like "little" and "a lot" mean, you can always go back and check the 'calendar' above. "X"s are drawn over each word as well, to help you remember what each word means (e.g. there are 4 to 5 "X"s over the words "a lot", to remind you that "a lot" means 4-5 times a week)

1. I saw birds

Never Rarely Sometimes Often Almost always
 x xx-x xxxx-x xxxxxx-x

2. I enjoyed life.

Never Rarely Sometimes Often Almost always
 x xx-x xxxx-x xxxxxx-x

3. I thought I was just as good as other children.

Never Rarely Sometimes Often Almost always
 x xx-x xxxx-x xxxxxx-x

4. I told lies.

Never Rarely Sometimes Often Almost always
 x xx-x xxxx-x xxxxxx-x

5. I had trouble falling asleep.

Never Rarely Sometimes Often Almost always

6. I felt angry.

Never Rarely Sometimes Often Almost always
 x xx-x xxxx-x xxxxxx-x

7. I had trouble brushing my teeth, combing my hair, and getting ready for the day.

Never Rarely Sometimes Often Almost always
 x xx-x xxxx-x xxxxxx-x

8. I did the things that make me happy.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

9. I had an appetite

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

10. I felt sad.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

11. I had scary dreams.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

12. I had trouble waking up in the morning.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

13. I felt tension in my body.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

14. I performed my normal religious rituals (like praying, worshipping my god, etc.)

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

15. I worried about a lot of things.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

16. I saw the comforting image/spirit of a person who died.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

17. I played with my friends at school.

x	xx	xxxxx	xxxxxxx	
Never	Rarely	Sometimes	Often	Almost always

18. My friends said that when I played with them, I wanted to fight with them.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	a little	sometimes	a lot	Almost all the time

19. I had problems with my teachers.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	a little	sometimes	a lot	Almost all the time

20. My work at school is slow.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always

21. My teachers or parents said that I daydream a lot.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always

22. My teachers said that I have a hard time paying attention in school.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always

23. I did the homework my teachers gave me when the environment was right.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always

24. My teachers told me that I am doing worse and worse in my school work.

x	xx-x	xxxx-x	xxxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always

25. I do badly in my tests

x	xx-x	xxxx-x	xxxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always

26. People told me that I am irritable.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

27. I got the urge to run away from home.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

28. I felt better if I do some things over and over again (e.g. checking the locks on the door, washing my hands).

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

29. I talked to my parents about things that happen to me and things I do at school.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

30. I see the frightening image/spirit of a person who died.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

31. I hurt animals.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

32. I cried.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

33. I played with my friends in the neighbourhood.

	x	xx-x	xxxx-x	xxxxxxx-x
Never	a little	sometimes	a lot	almost all the time

34. I thought that life is not worth living.

	x	xx-x	xxxx-x	xxxxxxx-x
--	---	------	--------	-----------

Never	a little	sometimes	a lot	almost all the time
35. I felt frightened.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
36. I had aches and pains (e.g. stomach aches, headaches, etc.)				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
37. I don't feel like talking with my brothers and sisters.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
38. I felt hopeless.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
39. I didn't do my duties at home.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
40. My parents said that I don't obey them.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
41. I had thoughts of killing myself.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
42. I have trouble remembering things.				
	x	xx-x xxxx-x	xxxxxx-x	
Never	Rarely	Sometimes	Often	Almost always
43. I looked for love and affection from people who were not my relatives				
	x	xx-x xxxx-x	xxxxxx-x	

	Never	Rarely	Sometimes	Often	Almost always
44. I had someone to talk to about things that were important to me		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
45. I didn't talk to anyone about my feelings		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
46. I smoked cigarettes		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
47. When things got difficult, I gave up easily		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
48. I stole things from others		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
49. My parents said I didn't respect them		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
50. I used bad language		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always
51. I fought with my brothers and sisters		x	xx-x	xxxx-x	xxxxxxx-x
	Never	Rarely	Sometimes	Often	Almost always

Now, on a scale of 1 – 10, mark how good you think your life is right now:

1

5

10

☹

:-|

☺

My life is horrible

My life is not bad or good

My life is great



YOU DID IT! Thank you for your valuable time! We hope that the information you give us will help us to help children like you.

Appendix C.2 SLIPS-C MH-2 version

The instructions page was also included in version MH-2 and SLIPS-GFP. It is not provided here to save space. The RTL drew the bowls of fruit on the board to facilitate understanding in the participants.

PARTICIPANT DETAILS

Name _____ Date _____

Age _____ Gender *Male Female*

School _____

What ethnic group are you? M T S O

What religion are you? M Ch H B

PHYSICAL FITNESS DATA

Body weight [kg] _____ Body Height [cm] _____

Beep Test Score _____

On a scale of one to ten mark how good your life is right now:

1 **5** **10**
☹ :-| ☺
My life is horrible My life is not bad or good My life is great

My Current Feelings, Thoughts and Behaviours

Think about what you have been feeling, what you have been thinking and how you have been behaving for the past TWO WEEKS. Then for each question below, please indicate how often you have experienced each event in the last TWO WEEKS. There are no right or wrong answers, only your own feelings, thoughts and experiences.

How would you answer this question?

In the last two weeks, I smiled:

0-day a week 1 day a week 2-3 days a week 4-5days a week 6-7days a week
(NEVER) (RARELY) (SOMETIMES) (OFTEN) (ALMOST ALWAYS)

If you smiled every day, circle "Almost always", but if you smiled only once in the last TWO WEEKS, you would circle "Rarely". Over the page you will be asked to tick boxes, not circle the answers. Please listen carefully and follow the instructions.

		Never	Rarely x	Sometimes xx	Often xxxx	Almost Always xxxxxx
1	I saw birds					
2	I enjoyed life					
3	I thought I was as good as other children					
4	I told lies					
5	I had trouble falling asleep					
6	I felt angry					
7	I had trouble in preparing for my day-to-day activities					
8	I did the things that made me happy					
9	I had a good appetite					
10	I felt sad					
11	I had scary dreams					
12	I had trouble waking in the morning					
13	I felt tension in my body					
14	I performed my normal religious rituals					
15	I worried about a lot of things					
16	I saw an image/spirit of a person who died and it comforted me					
17	I played with my friends at school					
18	My friends said that when I play with them I try to fight with them					
19	I had problems with my teachers					
20	My schoolwork was slow					
21	My teachers and parents said that I daydreamed a lot					
22	My teachers said I had a hard time paying attention in school					
23	I did my homework well when the environment was right					
24	My teachers told me I was doing worse and worse in school					
25	I did well in my schools tests					
26	People told me that I was irritable					
27	I felt like running away from home					
28	I felt better when I did some things over and over again (like checking locks or washing my hands)					
29	I talked to my parents about things that happened to me and things I did in school					
30	I saw the frightening image/spirit of someone who died					
31	I hurt animas					
32	I cried					
33	I played a lot with my friends in the neighbourhood					
34	I thought that life was not worth living					
35	I felt afraid					
36	I had aches and pains (in my stomach, head etc).					
37	I cooperated with my brothers and sisters					
38	I felt hopeless					
39	I did my duties at home					
40	My parents said that I don't obey them					
41	I thought about killing myself					
42	I played with friends at school					
43	I had trouble remembering things					
44	I looked for love and affection from people who are not my relatives					
45	I had someone to talk to about things that were important to me					
46	I didn't talk to anyone about my feelings					
47	I smoked					
48	When things became difficult, I gave up easily					
49	I stole things from others					
50	I used bad language					
51	I fought with my brothers and sisters					

Appendix C.3 SLIPS-GFP Final Version

PARTICIPANT DETAILS				
Name _____	Date _____			
Age _____	Gender		Male	Female
School _____				
What ethnic group are you?	M	T	S	O
What religion are you?	M	Ch	H	B

PHYSICAL FITNESS DATA	
Body weight [kg] _____	Body Height [cm] _____
Beep Test Score _____	

On a scale of one to ten mark how good your life is right now:

1	5	10
☹	:-	☺
My life is horrible	My life is not bad or good	My life is great

My Current Feelings, Thoughts and Behaviours

Think about what you have been feeling, what you have been thinking and how you have been behaving for the past TWO WEEKS. Then for each question below, please indicate how often you have experienced each event in the last TWO WEEKS. There are no right or wrong answers, only your own feelings, thoughts and experiences.

How would you answer this question?

In the last two weeks, I smiled:

0-day a week (NEVER)	1 day a week (RARELY)	2-3 days a week (SOMETIMES)	4-5days a week (OFTEN)	6-7days a week (ALMOST ALWAYS)
-------------------------	--------------------------	--------------------------------	---------------------------	-----------------------------------

If you smiled every day, circle "Almost always", but if you smiled only once in the last TWO WEEKS, you would circle "Rarely". Over the page you will be asked to tick boxes, not circle the answers. Please listen carefully and follow the instructions.

1. I saw birds				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
2. I enjoyed life.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
3. I had trouble falling asleep.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
4. I felt angry.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
5. I felt sad.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
6. I had scary dreams.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
7. I felt tension in my body.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
8. I worried about a lot of things.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always
9. I played with my friends at school.				
	x	xx-x	xxxx-x	xxxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

10. My friends said that when I played with them, I wanted to fight with them.

	x	xx-x	xxxx-x	xxxxxx-x
Never	a little	sometimes	a lot	Almost all the time

11. I had problems with my teachers.

	x	xx-x	xxxx-x	xxxxxx-x
Never	a little	sometimes	a lot	Almost all the time

12. My work at school is slow.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

13. My teachers or parents said that I daydream a lot.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

14. My teachers said that I have a hard time paying attention in school.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

15. My teachers told me that I am doing worse and worse in my school work.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

16. People told me that I am irritable.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

17. I cried.

	x	xx-x x	xxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

18. I played with my friends in the neighbourhood.

	x	xx-x	xxxx-x	xxxxxx-x
Never	A little	Sometimes	A lot	Almost all the time

19. I thought that life is not worth living.

	x	xx-x	xxxx-x	xxxxxx-x
Never	a little	sometimes	a lot	almost all the time

20. I had aches and pains (e.g. stomach aches, headaches, etc.)

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

21. I felt hopeless.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

22. My parents said that I don't obey them.

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

23. I looked for love and affection from people who were not my relatives

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

24. I didn't talk to anyone about my feelings

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

25. I fought with my brothers and sisters

	x	xx-x	xxxx-x	xxxxxx-x
Never	Rarely	Sometimes	Often	Almost always

Appendix C.4 BMI Measurement Protocol

Calculation of BMI requires a measure of height and weight. The following protocol outlines the equipment needed and the instructions for measuring height and weight.

Height measurement

Equipment needed:

- Tape measure
- Duct Tape
- Ruler
- Pen

Set up

- Look for a suitable area with space for the test and a flat wall.
- Attach tape measure to wall with tape.
- Ensure the tape measure is straight.

Instructions

Ask the participants to stand in a straight line and wait patiently.

Researcher should tell the participants to do the following:

- Take off their shoes
- Let down their hair
- Stand with their feet together against the tape measure
- Heels against the wall
- Relax their shoulders and look straight ahead

Researcher then does the following:

- Check participants feet are correct
- Place ruler on top of head

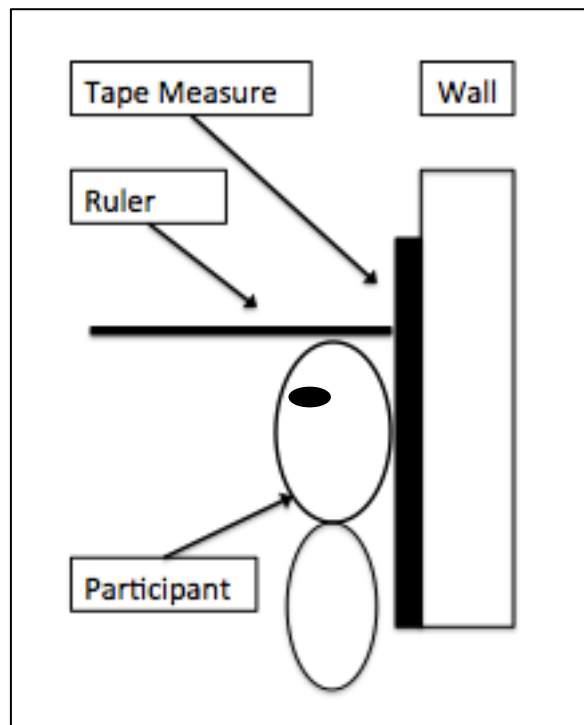


FIGURE A 6: Height measurement diagram

- Read the participant height on the tape measure
- Record height to the nearest 0.5 cm
- Write down height on participant sheet

Ask participant to stand clear and out the way and start next measurement

Weight measurement

Equipment needed

- Weighing scales
- Pen
- Chair (for researcher)

Set up

- Find a suitable area to set up.
- Must be solid, flat ground like concrete.
- Place weighing scales down
- 'Zero' the weighing scales
- Place chair in front of scales

Instructions

Ask the participants to wait in a line quietly.

Researcher sits on chair in front of weighing scales.

Researcher asks the participants to do the following:

- Empty their pockets.
- Take off their shoes and any heavy clothing.
- Stand still on the scales facing forward.

Researcher then

- Measures the participant weight to the nearest 0.5kg.
- Writes weight down on participant sheet.

Ask participant to stand out of the way and wait quietly for their next instruction.

BMI is calculated using the following, standard formula:

$$\mathbf{BMI = (weight\ in\ kilograms) / (height\ in\ meters)}$$

Appendix C.5 MFT measurement protocol

CF will be estimated using the MFT according to the following standard protocol.

- A testing grid is measured using a 20m measuring tape and marked out with cones, according to the diagram in FIGURE 9.5 below.
- The participants line up along the cones on one side of the testing grid.
- The participants are read the following instructions, in local language to ensure the protocol is fully understood.

“This test is called the beep test. You are required to run between the markers in time with the ‘beeps’ that you will hear. Each ‘beep’ is your signal to start running towards the opposite marker. You must run one whole length before you hear the ‘beep’ again.

You must not set off before you hear the ‘beep’. After about a minute, you will hear the audio recording advance to the next level. When this happens, the interval between the ‘beeps’ will get shorter and you will have to run a little bit faster. Each minute you will progress to the next level. The aim of the test is to get to as high a level as possible, so try as hard as you can to beat the ‘beep’. If you do not make your marker, try as hard as you can to make it across before the next ‘beep’. If you miss two beeps in a row you will have finished the test. Remember, the test starts very slowly, so you won’t have to run very fast.” (Richards: 2011: 309)⁴

- All subjects are given the chance to ask any questions.
- When all participants are ready, the audio recording is started.
- The subjects begin running between the markers, in time with the ‘beeps’ on the audio recording.
- If a participant misses a ‘beep’ the principal researcher encourages them to attempt to make the next line in time for the ‘beep’. They must succeed in catching up the line within two ‘beeps’.
- If the participant misses the ‘beep’ on two consecutive times they are stopped. The researcher notes down their score.
- Scoring is based on the level and stage on which they stopped. So level 6 stage 7 is scored as 6.7.

⁴ Richards, J. *Evaluating the impact of a sport-for-development intervention on the physical and mental health of young adolescents in Gulu, Uganda – a post conflict setting within a low income country.* 2011.

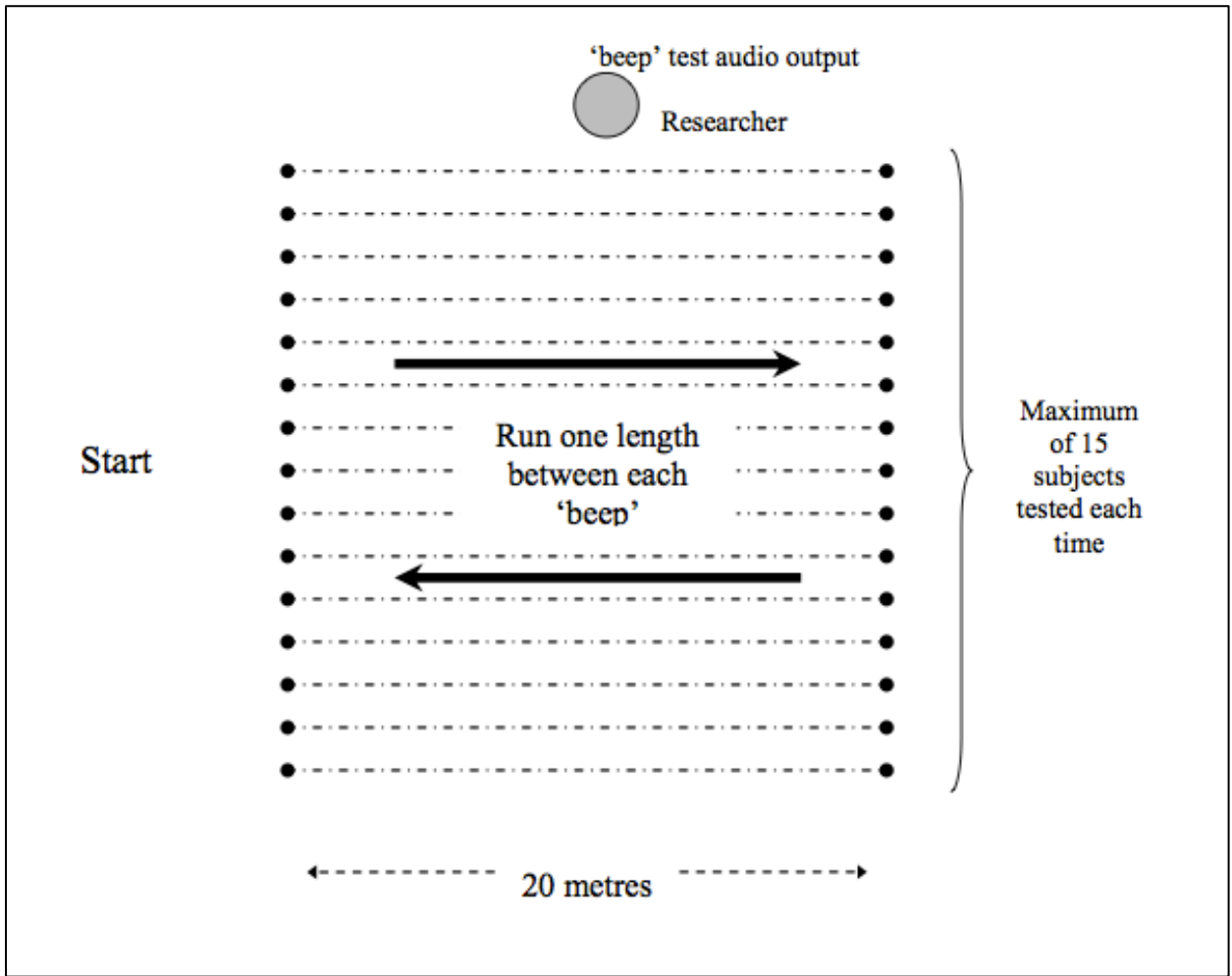


FIGURE A 7: MFT test diagram (Richards 2011)

Appendix C.6 Information sheet for parents

Re: Assessing the impact of an intervention on the psychosocial well-being of an adolescent population in Mullaitivu, Sri Lanka.

This is an information sheet for a study in schools in Mullaitivu district. It is important for you to understand the purpose of the study and what it will involve. Please take time to read this sheet carefully. Everyone is free to ask questions at any time.

Purpose of the study?

This project will test the physical fitness and well-being of adolescents in northern province. We hope that this will help us understand the impact of a Sports For Peace programme on young adolescents, and help make interventions more effective in the future.

How have adolescents been chosen?

Several schools have been selected for the study. All adolescents aged 13-15 can choose to take part in the study. Participation is voluntary. Attached is a consent form, please complete this and send it back to school if you want your child to participate in this study.

If you wish to remove your child from the study, contact a member of the research team. Contact details are below. Should your child choose to drop out of the study at any time, they can do so without giving a reason. They will not be penalised in any way for dropping out.

What will happen to adolescents in the study?

Three times over a 12 month period adolescents will take part in a series of tests, during their normal sports and physical education lessons in school:

- (1) We will measure the height and weight of participants
- (2) Participants will be asked to walk and run for 15 minutes and have their fitness measured.
- (3) Participants will complete a Sri Lankan well-being questionnaire asking about their feelings and emotions over the last two weeks.

What are the benefits in taking part in the study?

Taking part in this study will provide your child an opportunity to contribute to scientific research in your district. It will encourage the improvement of future sports-based interventions. Equally,

you can find out how physically fit and healthy your child is, using an international test.

Participants will not be reimbursed for their participation.

What are the risks to participants in the study?

There are no known risks for taking part in this study. There is minimal psychological risk from the mental health questionnaire and similarly minimal physical risk from the fitness tests. Trained personnel will deliver the questionnaire and appropriate follow up will be available, if required.

The research personnel are from an international NGO and trained to provide the necessary child support. A trained counsellor will be present at each data collection session.

If you wish to complain about any aspect of the way in which you have been approached or treated during the course of this study, you should contact Alexander Hamilton (details below) or you may contact the University of Oxford Clinical Trials and Research Governance (CTRG) email: ctrg@admin.ox.ac.uk

The University of Oxford, as Sponsor, has appropriate insurance in place in the unlikely event that you suffer any harm arising from the negligence of the University, or that of a partner, in this research, and with that harm resulting as a direct consequence of your participation in this study.

What will happen to the information?

All information is strictly confidential and only the researchers will have access to the data. All participants are identified by a code number and identifying codes will be kept in a separate location. All paper copies will be kept in a locked filing cabinet. All computer files will be kept on an encrypted laptop and will be password secured.

At the end of the study the results will be available for the school and their students. Results on individual participants will not be made available. We also hope to publish the results for the public once the study is finished.

Who will be doing the research?

The research is being organised by Alexander Hamilton from the University of Oxford, a PhD student, under the supervision of Dr Charlie Foster. The research is in collaboration with academics associated with the University of Colombo.

The study is funded by the personal scholarship of Alexander Hamilton, the Department of Population Health at the University of Oxford and Generations For Peace, a leading global NGO.

The University of Oxford and the University of Colombo ERC have given research and ethics approval.

If you have any questions or would like further information please contact:

Alexander Hamilton

BHFHPRG, Rosemary Rue Building, Department of Population Health, Old Road Campus, Headington, OX3 7LF

Tel: +447795576097

Tel: +946719467911

Email: alexander.hamilton@stx.ox.ac.uk

Dr Asela Olupeliyawa

Medical Education Development and Research Centre, Faculty of Medicine, University of Colombo, 25 Kynsey Road, Colombo 08, PO Box 271

Tel: +94718329300

Email: asela_o@yahoo.com

Appendix C.7 Consent forms

Assessing the impact of an intervention on the psychosocial well-being of an adolescent population in Mullaitivu, Sri Lanka.

Principal Investigator: Mr Alexander Hamilton
University of Oxford
PhD candidate – Population Health
Email: alexander.hamilton@stx.ox.ac.uk
Phone: 00447795576097
Sri Lanka mob: +946719467911

Your child has agreed to take part in a study run by the University of Oxford researching psychosocial wellbeing and physical fitness.

Please fill out and return this form if you do not want your child to take part in the study.

1) Have you read the information sheet? **YES/NO**

Name of child:

I do not want my child to take part in this study: **YES/NO**

Name of parent/guardian:

Signature/thumbprint:

Date:

Appendix C.8 Normality tests for MH results in Mullaitivu and Kilinochchi

MH

C.8.1 Kilinochchi ICC assumptions testing

To assess test-retest reliability the ICC was used. The use of the ICC requires that certain assumptions be fulfilled. Normality of SLIPS-C (MH-2) was initially assessed using histograms. The initial results warranted further assessment.

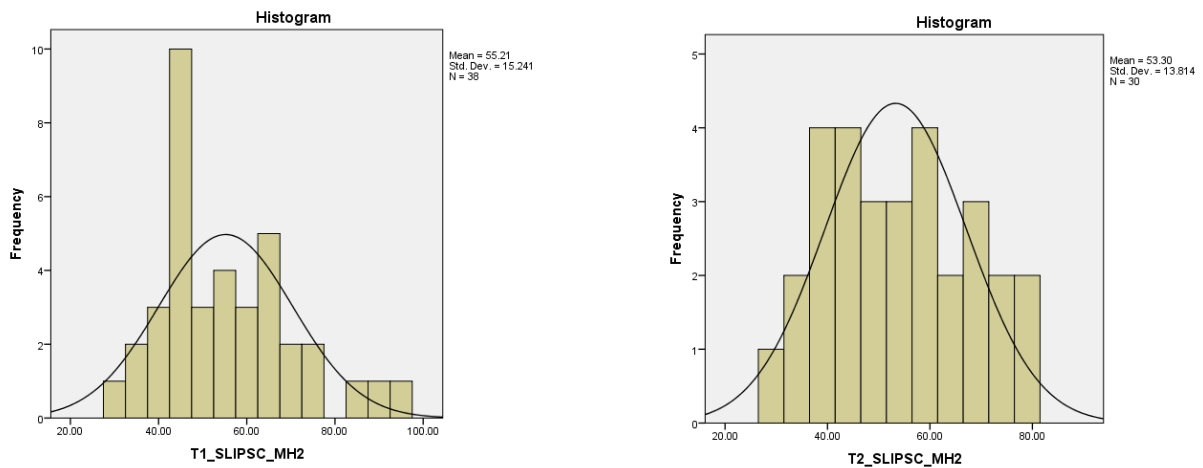


FIGURE A 8: Histograms of SLIPS-C (MH-2) in Kilinochchi, 2014

The Kolmogorov-Smirnoff (KS) test was used to assess normality of the SLIPS-C (MH-2) in the Kilinochchi study.

TABLE A 2: Kolmogorov Smirnoff test for normality in the Kilinochchi test-retest study, 2014

		n	Skew	S.E.S	Kurtosis	S.E.K	K-S
Kili	T1	38	0.790	0.383	0.274	0.750	0.021*
	T2	30	0.194	0.427	-0.924	0.833	0.200

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

S.E.S: Standard error of the skew

S.E.K: Standard error of the kurtosis

K.S: Kolmogorov-Smirnoff test

Bold and italic indicates a significant result.

An approximate measure of normality is whether the skew or kurtosis lies within 2 SES or 2 SEK respectively. The SLIPS-C distributed significantly differed from the standard normal distribution, and the was also significantly skewed at T1.

Despite SLIPS-C (MH-2) at T1 differing from the standard normal, the ICC can still be used. This is because the ICC requires that the *population* distribution is normal, not the sample. In this case, the sample distribution differs from the standard normal.

An additional assumption for the use of the ICC is to check population HOV. This assumption can be verified through a Levene's test.

TABLE A 3: Levene's test for HOV, Kilinochchi 2014

Outcome measure (T1 - T2 comparison)	Levene's statistic	Sig.
SLIPS-C (MH-2)	0.088	0.767

The data did not seriously violate any of the assumptions required for the use of the ICC

C.8.2 Mullaitivu MH normality tests

Normality of SLIPS-GFP was initially assessed using histograms. The initial results warranted further assessment.

The KS test was used to assess normality of the SLIPS-GFP in the Mullaitivu study.

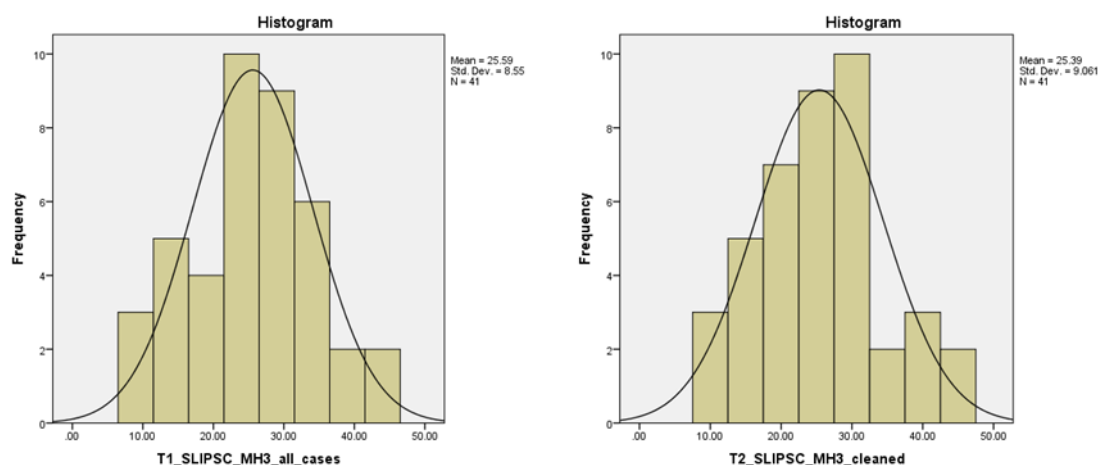


FIGURE A 9: Histograms of SLIPS-GFP score at T1 and T2

TABLE A 4: Kolmogorov-Smirnoff tests for SLIPS-GFP in Mullaitivu, 2015

		n	Skew	S.E.S	Kurtosis	S.E.K	K-S
Mull	T1	41	0.034	0.369	-0.416	0.724	0.2
	T2	41	0.232	0.369	-0.466	0.724	0.2

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

TABLE A 5: Levene's test for SLIPS-GFP between T1 and T2.

Outcome measure	Levene's statistic	Sig.
SLIPS-GFP	0.178	0.674

A Levene's test confirmed HOV for SLIPS-GFP at T1 and T2.

C.8.3 Kilinochchi PH

The distribution of the PH data was initially assessed using histograms. The initial results warranted further assessment.

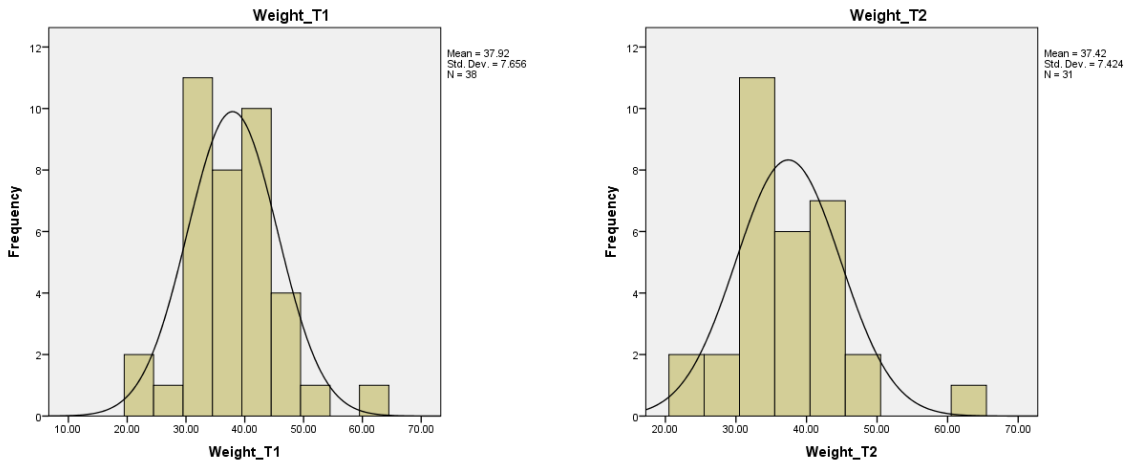


FIGURE A 10: Weight at T1 and T2 in Kilinochchi, 2014

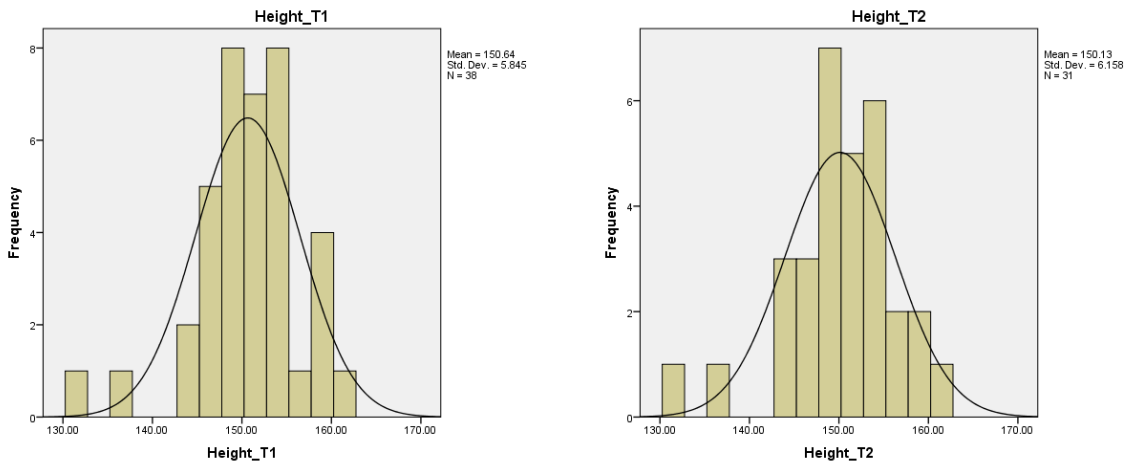


FIGURE A 11: Height at T1 and T2 in Kilinochchi, 2014

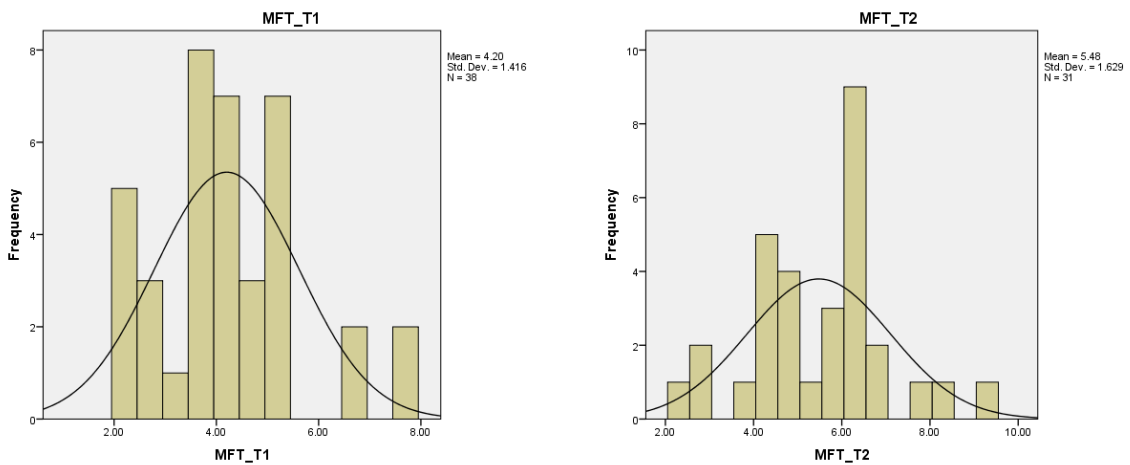


FIGURE A 12: MFT score at T1 and T2 in Kilinochchi, 2014

The KS test was used to assess statistically assess whether the distribution of the Kilinochchi PH data differed significantly from the standard normal.

TABLE A 6: Normality tests for the PH variables in Kilinochchi, 2014

Location	Outcome	Time	n	Skew	S.E.S	Kurtosis	S.E.K	K-S
Kili	Height	T1	38	-0.971	0.383	2.621	0.750	0.200
		T2	31	-0.972	0.421	2.324	0.821	0.024*
	Weight	T1	38	0.576	0.382	1.219	0.750	0.200
		T2	31	0.816	0.421	2.226	0.821	0.200
	MFT (S.L)	T1	38	0.652	0.383	0.51	0.75	0.076
		T2	31	0.102	0.421	0.327	0.821	0.200

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

There were several violations of normality in the height and weight measurements. Skewness and kurtosis of height at T1 and T2, and weight at T2, exceeded the minimal acceptable value of 2 x SEK/SES. The only significant K-S test result was height at T2. Distribution of height at T2 significantly differed from the standard normal distribution.

As previously stated, the ICC requires that the distribution of the population from which the sample is drawn be normally distributed. Multiple studies have indicated that height is normally distributed on a population level. The violation of the normality assumption is not a problem in this instance.

A Levene's test verified HOV between the PH variables at T1 and T2 in Kilinochchi.

TABLE A 7: Levene's test for HOV for the PH data, Kilinochchi, 2014

Location	Outcome measure	Levene's statistic	Sig.
Kili	Weight	0.013	0.909
	Height	0.016	0.899
	MFT score (S.L)	0.713	0.402

An ICC can be used to assess the inter-rater agreement for the Kilinochchi study PH outcome variables.

C.8.4 Mullaitivu PH normality tests

Histograms were initially used to assess normality of MFT scores in Mullaitivu. The initial results required further assessment.

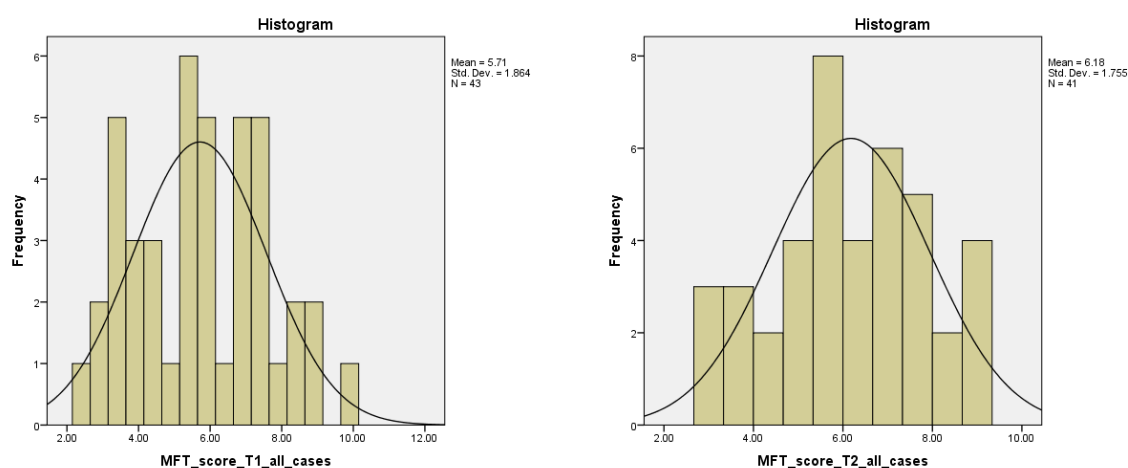


FIGURE A 13: Histograms of MFT score in Mullaitivu, 2015

Normality of MFT scores in Mullaitivu was assessed using the KS test. The KS test verified that the distribution of the MFT score did not differ significantly from the standard normal distribution.

TABLE A 8: KS tests for normality of MFT scores in Mullaitivu, 2015

Outcome	Time	n	Skew	S.E.S	Kurtosis	S.E.K	K.S
MFT	T1	43	0.206	0.361	-0.700	0.709	0.200*
	T2	41	-0.138	0.369	0.568	0.724	0.200*

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

A Levene's test verified homogeneity of variance between MFT at T1 and T2 in Mullaitivu.

TABLE A 9: Levene's test for HOV of MFT score in Mullaitivu, 2015

Location	Outcome	Levene's statistic	Sig.
Mullaitivu	MFT score	1.195	0.279

Appendix C.9 Ethics Approval for studies

C.9.1 Oxford Ethical approval

Oxford Tropical Research Ethics Committee

University of Oxford

Joint Research Office

Churchill Hospital, Oxford OX3 7LE

Tel. +44 (0) 1865 (5)72346 fax +44 (0) 1865 (5)72228

E-mail: jacqueline.gerencser@admin.ox.ac.uk



Alexander Hamilton

07 April 2014

Nuffield Department of Population Health

University of Oxford

2nd Floor, Rosemary Rue Building

Old Road Campus

Headington

Oxford

OX3 7LF

Dear Alexander

Full Title of Study: Assessing the impact of a sports peace program on the psychosocial well-being, cardiovascular fitness and the body composition on an adolescent population in post-conflict Sri Lanka

OXTREC Reference: 523-14

Thank you for your letters of the 28 November 2013 and 02 April 2014, and for your minimal risk application form. Other documents reviewed were:

Documentation	Version	Date
Protocol	V2.0	02/04/14
Participant Information Sheet		27/02/14
Assent Form		27/02/14
Appendix B – Fieldwork Risk Assessment	V1.0	27/02/14
SLIPS-C Questionnaire (child version)		27/02/14
Psychosocial Protocol		27/02/14
Beep Test Protocol		27/02/14
BMI Protocol		27/02/14
Class Teacher (letter)		27/02/14
Head Teacher (letter)	V2.0	27/02/14
Parents Opt-Out (letter)	V2.0	27/02/14

The OXTREC executive team reviewed the above application on the 06 April 2014 and gave approval for this study.

Approval is given for the first five years and subject to receiving the local ethical approval.

This approval will be fully implemented upon final consent of the full committee at the meeting on Thursday 22 May 2014.

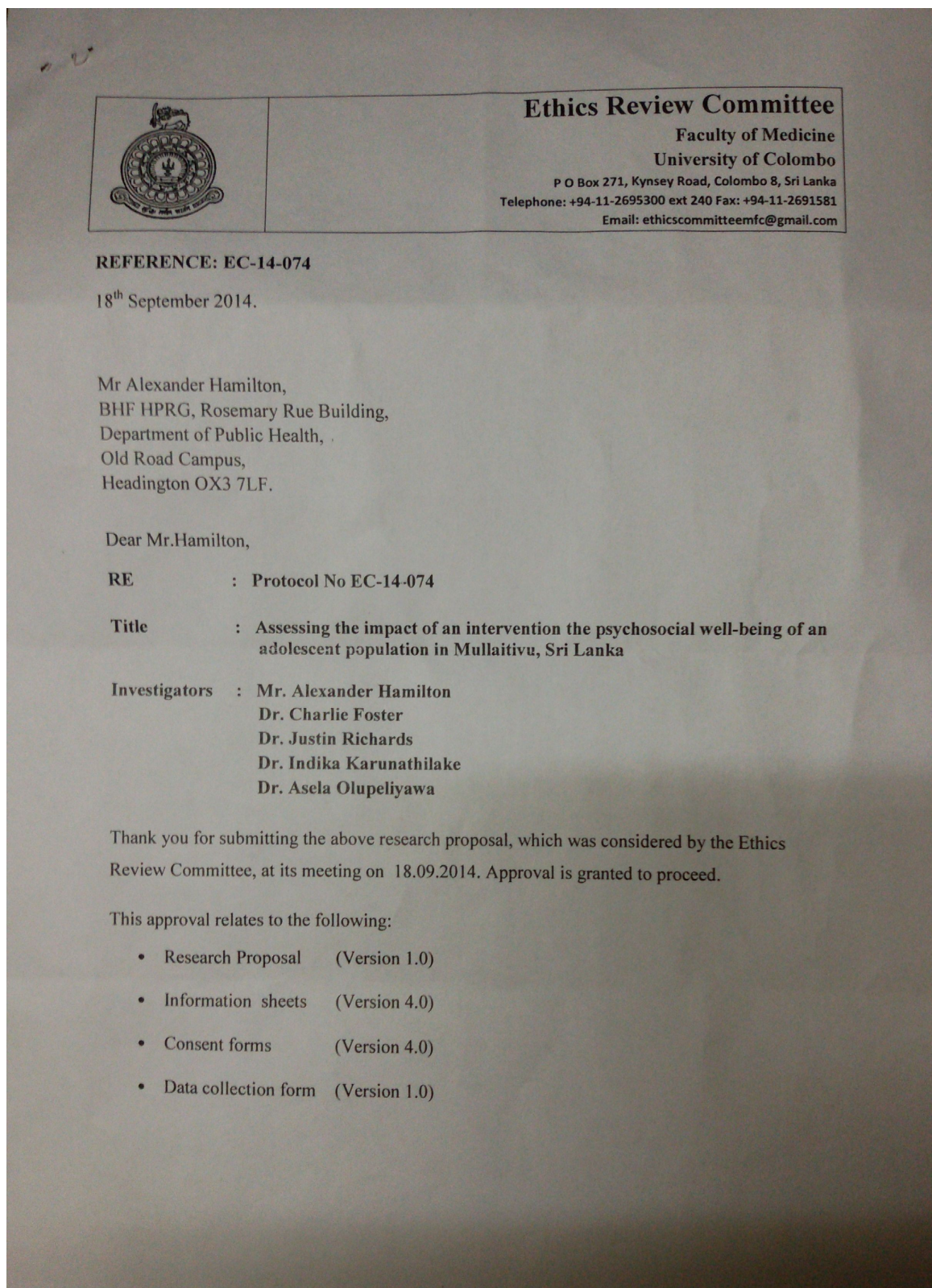
Yours sincerely



Dr Mary Warrell

OXTREC Chairman

C.9.2 Sri Lankan Ethical approval





Ethics Review Committee

Faculty of Medicine

University of Colombo

P O Box 271, Kynsey Road, Colombo 8, Sri Lanka

Telephone: +94-11-2695300 ext 240 Fax: +94-11-2691581

Email: ethicscommitteemfc@gmail.com

REFERENCE: EC-14-074

18th September 2014.

Mr Alexander Hamilton,
BHF HPRG, Rosemary Rue Building,
Department of Public Health,
Old Road Campus,
Headington OX3 7LF.

Dear Mr.Hamilton,

RE : Protocol No EC-14-074

Title : Assessing the impact of an intervention the psychosocial well-being of an adolescent population in Mullaitivu, Sri Lanka

Investigators : Mr. Alexander Hamilton
Dr. Charlie Foster
Dr. Justin Richards
Dr. Indika Karunathilake
Dr. Asela Olupeliyawa

Thank you for submitting the above research proposal, which was considered by the Ethics Review Committee, at its meeting on 18.09.2014. Approval is granted to proceed.

This approval relates to the following:

- Research Proposal (Version 1.0)
- Information sheets (Version 4.0)
- Consent forms (Version 4.0)
- Data collection form (Version 1.0)

APPENDIX D

Appendix D.1 Normality tests and comparison of sub-groups

I checked if there were any significant variable scores between subgroups. The statistical tests for comparing performance at M1 assume that the samples are approximately normally distributed and have HOV.

Scores were plotted as histograms and normality was assessed visually. Initial inspection of normality warranted further assessment. It was common for the data to violate normality assumptions, due to the stringent nature of normality tests. I tried standard transformation methods, largely had a marginal effect on improving the outcomes of the normality tests.

Violations of normality were common. Despite this, I analysed the differences using the ANOVA as:

- The normality tests are stringent
- The data approximated the normal distribution when plotted
- The sample sizes are large enough to use a conservative statistical test such as ANOVA
- Violations of HOV are of more concern when comparing means.

In those instances where normality assumptions were violated, non-parametric tests were used to verify the findings of the parametric tests. In the event that there was disagreement between the non-parametric and parametric tests, I have discussed each case in isolation.

TABLE A 10: ANOVA for main outcome variables assessing differences between girls and boys (df = 2)

BOYS - GIRLS	Levene's	Sig.	F	Sig.
SLIPS-GFP	1.754	0.186	9.63	0.002**
INT	0.141	0.707	5.45	0.02*
EXT	1.626	0.203	15.36	<0.001**
WTH	0.103	0.748	0.002	0.962
MFT	23.695	>0.001	163.36	<0.001**
BFA	1.018	0.314	33.94	<0.001**
HFA	0.574	0.449	43.51	<0.001**

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level, ** to the 0.01 significance level

A Mann Whitney test verified that boys outperform girls in the MFT (Z = -10.589, p 0.05).

TABLE A 11: ANOVA for main outcome variables comparing 13 and 14 year olds (df=2)

13 - 14	Levene's	Sig.	F	Sig.
SLIPS-GFP	0.59	.808	0.532	0.466
INT	2.327	0.128	1	0.318
EXT	2.477	0.116	3.291	0.071
WTH	0.086	0.770	4.53	0.034*
MFT	0.823	0.365	0.118	0.732
BFA	0.001	0.969	2.854	0.092
HFA	1.022	0.313	1.613	0.205

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

Non-parametric tests confirmed the results shown of the ANOVA.

TABLE A 12: Mean and 95% confidence intervals for the main MH and PH variables stratified by ethnic group.

Ethnicity	n	SLIPS-C	MFT	BFA	HFA
Tamil	244	25.34 (24.20-26.49)	4.88 (4.67-5.09)	-1.57 (-1.91--1.65)	-1.73 (-1.89--1.57)
Sinhala	66	23.14 (21.30-24.97)	4.92 (4.51-5.33)	-1.55 (-1.78--1.31)	-1.93 (-2.26--1.61)
Muslim	57	24.77 (22.41-27.14)	3.97 (3.62-4.32)	-2.35 (-2.60--2.10)	-1.70 (-1.97--1.42)

Initial analysis suggests that there is a difference in the BFA and MFT scores of the different ethnic groups. The Muslim groups significantly underperform when compared to the Tamil and Sinhala groups (TABLE 11.12)

TABLE A 13: ANOVA of all outcome variables comparing ethnic groups (d.f = 3)

13 - 14	Levene's	Sig.	F	Sig.
SLIPS-GFP	2.028	0.133	1.65	0.192
MFT	4.135	0.17*	7.77	<0.001*
BFA	0.742	0.477	10.64	<0.001*
HFA	1.437	0.239	0.777	0.460

NB: BOLD indicates a significant result, * indicates significance to the 0.05 level

ANOVA verified that there are significant between group differences MFT and BFA scores. Further analyses must control for ethnicity. I further stratified the sample to see if there was a school-level effect on the primary outcome variables, independent of ethnicity.

TABLE A 14: Mean and 95% confidence intervals for the main MH and PH outcome variable stratified by school

Ethnic group	School	n	SLIPS-C	MFT	BFA	HFA
Tamil	AMB	39	25.90 (23.15-28.65)	4.83 (4.37-5.27)	-1.57 (-1.92--1.22)	-1.90 (-2.37--1.43)
	MVK	30	23.53 (20.34-26.72)	4.77 (4.16-5.37)	-1.76 (-2.10--1.42)	-1.67 (-2.01--1.33)
	VATT	69	25.43 (23.32-27.55)	5.01 (4.62-5.39)	-1.87 (-2.10--1.66)	-1.67 (-1.97--1.37)
	CHM	41	20.63 (17.57-23.70)	5.68 (5.09-6.28)	-2.10 (-2.40--1.81)	-1.91 (-2.31--1.50)
	ALMP	65	28.74 (26.72-30.76)	4.34 (3.93-4.75)	-1.59 (-1.88--1.30)	-1.61 (-1.93--1.28)
Sinhala	SAMP	54	23.90 (21.77-26.04)	4.86 (4.40-5.31)	-1.55 (-1.82--1.28)	-1.88 (-2.24--1.51)
	KIR	12	19.67 (17.02-22.31)	5.22 (4.15-6.29)	-1.53 (-2.01--1.05)	-2.18 (2.99--1.37)
Muslim	THAN	57	24.77 (22.41-27.14)	3.97 (3.61-4.32)	-2.35 (-2.60--2.10)	-1.70 (-1.98--1.42)

It is evident from TABLE 11.13 that there are some school-level effects on the main MH and PH variables. FIGURES A14-A16 display these results graphically.

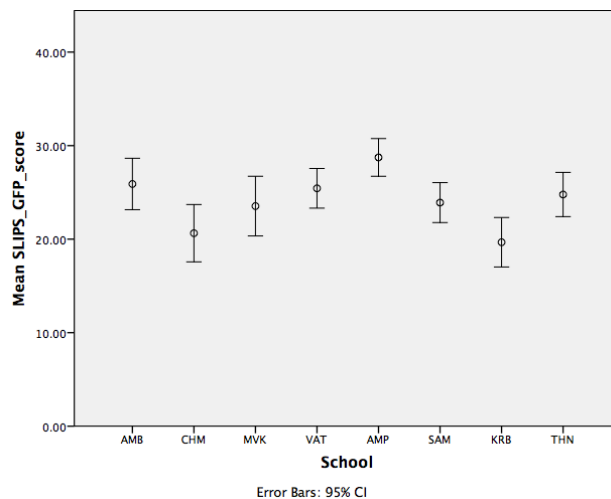


FIGURE A 14: Mean SLIPS-GFP scores stratified by school.

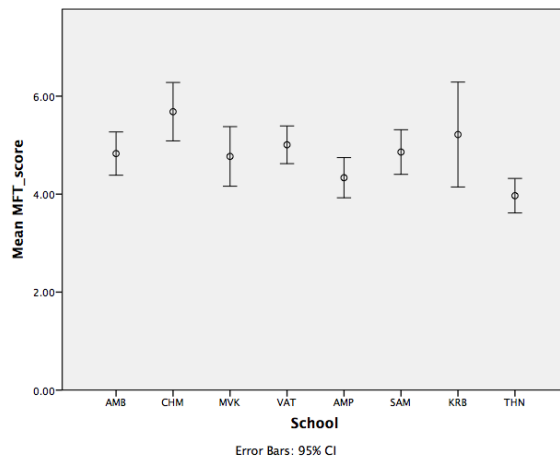


FIGURE A 15: Mean and 95% confidence intervals for MFT score stratified by school

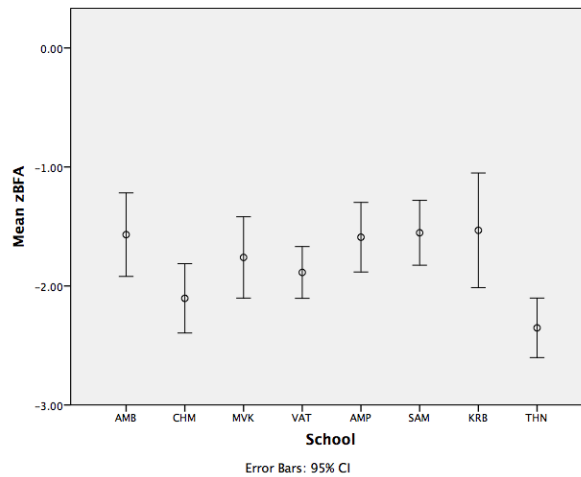


FIGURE A 16: Mean and 95% confidence intervals of BFA stratified by school

There are some between school differences in performance. The small sample size means that any statistical analyses at the school level are underpowered. AMP school had the worst average MH score, significantly worse than KRB, SAM and CHM (FIGURE A.14). KRB had the best average MH, but the sample size is so small ($n = 9$) that is difficult to draw any conclusions at this point.

CHM School performed the best in the MFT test; it also had the best sports ground, as it shared playing fields with a local football club. This is one possible explanation for their performance.

THN School performed very poorly on the MFT test; based on the 95% CIs, significantly worse than SAM, CHM and VAT schools. THN also had significantly lower BFA than all the other schools, except CHM. It is difficult to identify any consistent trends, except that the Muslim school, THN, performs poorly when compared to other local schools.

Appendix D.2 Comparison to global normative values

ANOVA was used to compare the performance of adolescents in Mullaitivu with global normative values. Certain assumptions must be verified before the implementation of ANOVA. The sample data must be drawn from a population in which the distribution of the means approximates the standard normal distribution. An additional assumption is that the samples have homogeneity of variance.

I stratified the sample by gender and age to compare with the global normative scores.

Normality of the fitness data was assessed visually using histograms. The initial assessments warranted further statistical tests.

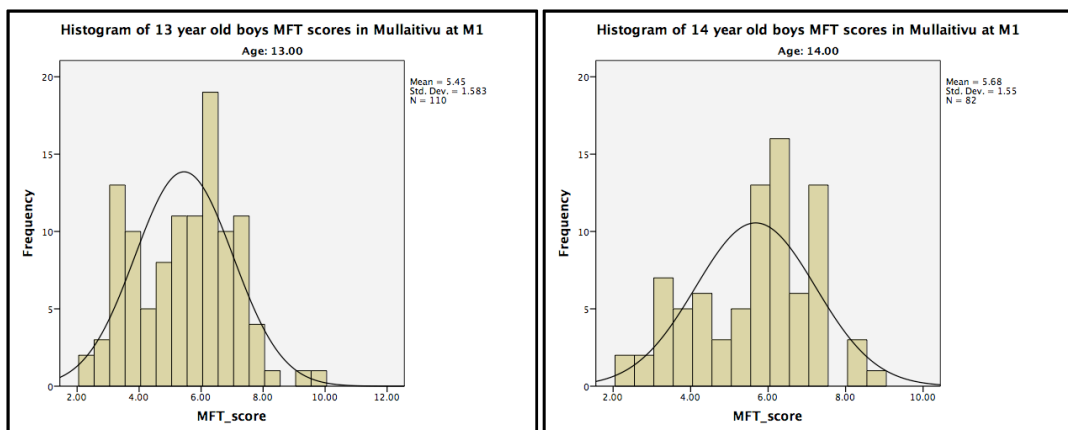


FIGURE A 17: Histogram of boys MFT scores in Mullaitivu at M1

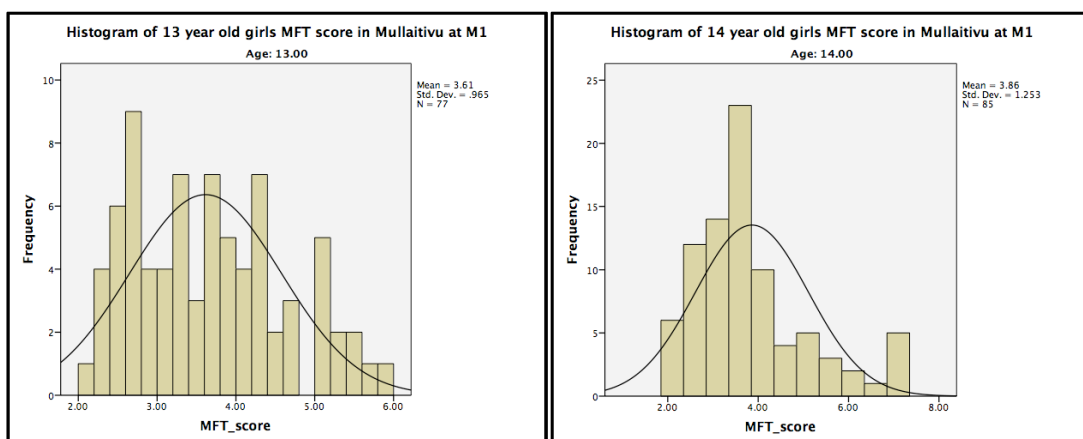


FIGURE A 18: Histograms of girls MFT scores in Mullaitivu at M1

TABLE A 15: Mean, standard deviation and normality test results of MFT scores in Mullaitivu at M1 stratified by sex and age

Sex	Age	N	Mean	SD	Skew	SES	K	SEK	S-W
Boys	13	110	5.45	1.58	0.012	0.230	-0.509	0.457	<i>0.036</i>
	14	82	5.68	1.55	-0.378	0.266	-0.626	0.526	<i>0.013</i>
Girls	13	77	3.61	0.965	0.451	0.274	-0.726	0.541	<i>0.05</i>
	14	85	3.86	1.25	<i>1.03</i>	0.261	0.611	0.517	<i><0.01</i>

NB: Bold and italic identifies a significant result

All groups failed the Shapiro-Wilk test. Standard transformations were applied to the data (square root and \log_{10}). Transformations made marginal improvements to the distribution of the data. In order for the Mullaitivu MFT scores to be compared MFT scores to the global normative values, they had to be converted into speed at last complete level of the MFT; this may have affected the distribution of the data. Significant difference was estimated using 95% confidence intervals, rather than ANOVA.

Appendix D.3 Comparison with Kilinochchi

The 22 item SLIPS-GFP was derived from the original 49 item SLIPS-C questionnaire I used to assess MH in Kilinochchi. The SLIPS-GFP has not been implemented in conjunction with a validated MH assessment tool. As a result, it is not possible to compare MH at M1 with any international population. Furthermore, there are those that question the cross-cultural validity of MH constructs, which clearly obstructs comparisons of this nature [271].

A comparison between the MH scores of adolescents in Kilinochchi is possible if the 22 SLIPS-GFP items are extracted from the 49-item version. This assumes that the excluded 27 items did not affect the remaining 22 items, which is a considerable assumption. The statistical test for comparing the performance in Kilinochchi and Mullaitivu assumes the samples are normally distributed and HOV between samples. Histograms of SLIPS-GFP and MFT scores were plotted and normality was assessed visually. After initial assessment further tests of normality were warranted.

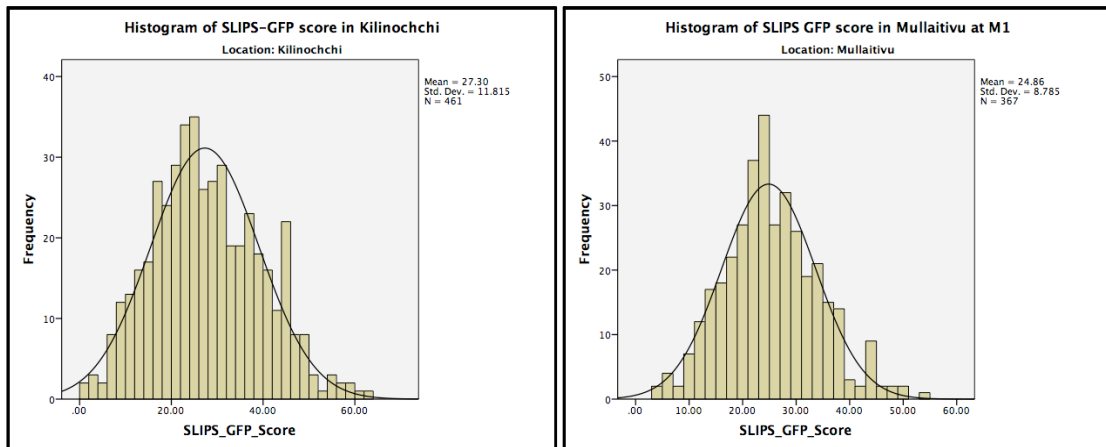


FIGURE A 19: Histograms of SLIPS-GFP score in Mullaitivu at M1 and Kilinochchi

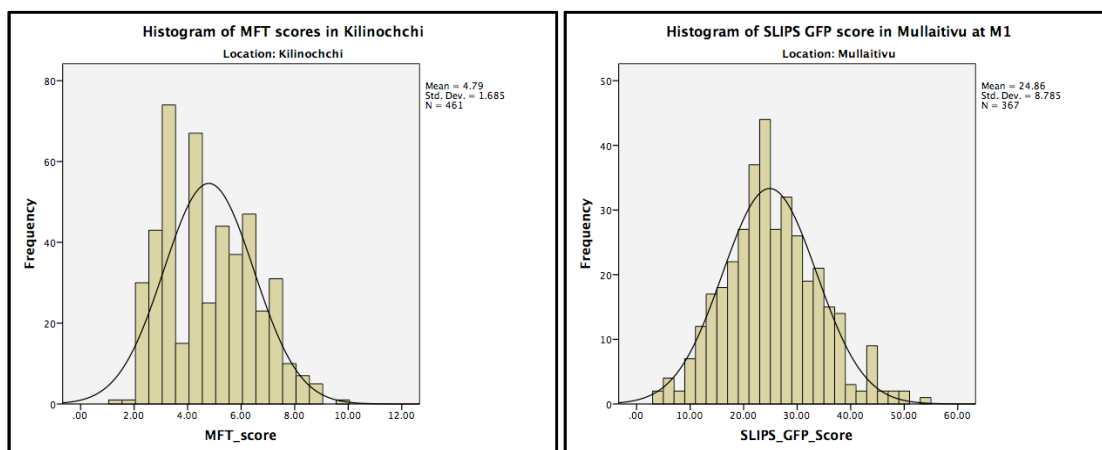


FIGURE A 20: Histograms of MFT scores in Mullaitivu at M1 and Kilinochchi

TABLE A 16: Mean, standard deviation and normality tests of SLIPS-GFP and MFT scores in Kilinochchi and Mullaitivu.

Location	N	Outcome	Mean	SD	Skew	SES	Kurt	SEK	S-W
Mullaitivu	367	SLIPS-GFP	24.86	8.79	0.289	0.127	0.083	0.254	0.042
			23.96 – 25.76						
		MFT	4.74	1.66	0.363	0.127	-0.820	0.254	
			4.59 – 4.92						
Kilinochchi	461	SLIPS-GFP	27.30	11.81	0.283	0.114	-0.352	0.227	0.002
			26.22 – 28.38						
		MFT	4.79	1.68	0.323	0.114	-0.8	0.227	
			4.64-4.95						

The distribution of the SLIPS-GFP and the MFT in Kilinochchi and Mullaitivu deviated from the standard normal, as is evidence in TABLE 11.16, in which all fail the S-W test. Similarly, the MFT is non-normal with regards to kurtosis in Kilinochchi and Mullaitivu. Standard transformations (natural logarithm and square root) had marginal impact on the distribution of the data. The SLIPS-GFP violated the assumption for HOV. Despite violation of the assumptions, I decided to use an ANOVA to test the differences between performance in Mullaitivu and Kilinochchi. This decision was based on the following reasons:

- There is a large sample size in both groups
- Statistical tests for normality are highly sensitive
- The data appear normally distributed on inspection

TABLE A 17: Levene's test for homogeneity of variance and ANOVA results for Mullaitivu and Kilinochchi

Outcome	Levene's	sig.	F	sig.
SLIPS-GFP	38.24	>0.001	10.85	0.01
MFT	0.002	0.962	0.132	0.717

As the SLIPS-GFP data in Kilinochchi and Mullaitivu do not have HOV, I have also performed a non-parametric test to certify the results of the ANOVA test. The Mann Whitney U test assumes that the samples have similar distributions, which can be assessed using the histograms in FIGURES A.19 AND A.20

TABLE A 18: Mann Whitney U test for mean rank scores in Kilinochchi and Mullaitivu

Outcome	Z statistic	Sig.
SLIPS-GFP	-2.764	0.06

The results of the non-parametric test support those of the parametric test.

Appendix D.4 Association between MH and PH variables

To assess the relationship between the PH and MH variables visually, the data were plotted as scatterplots.

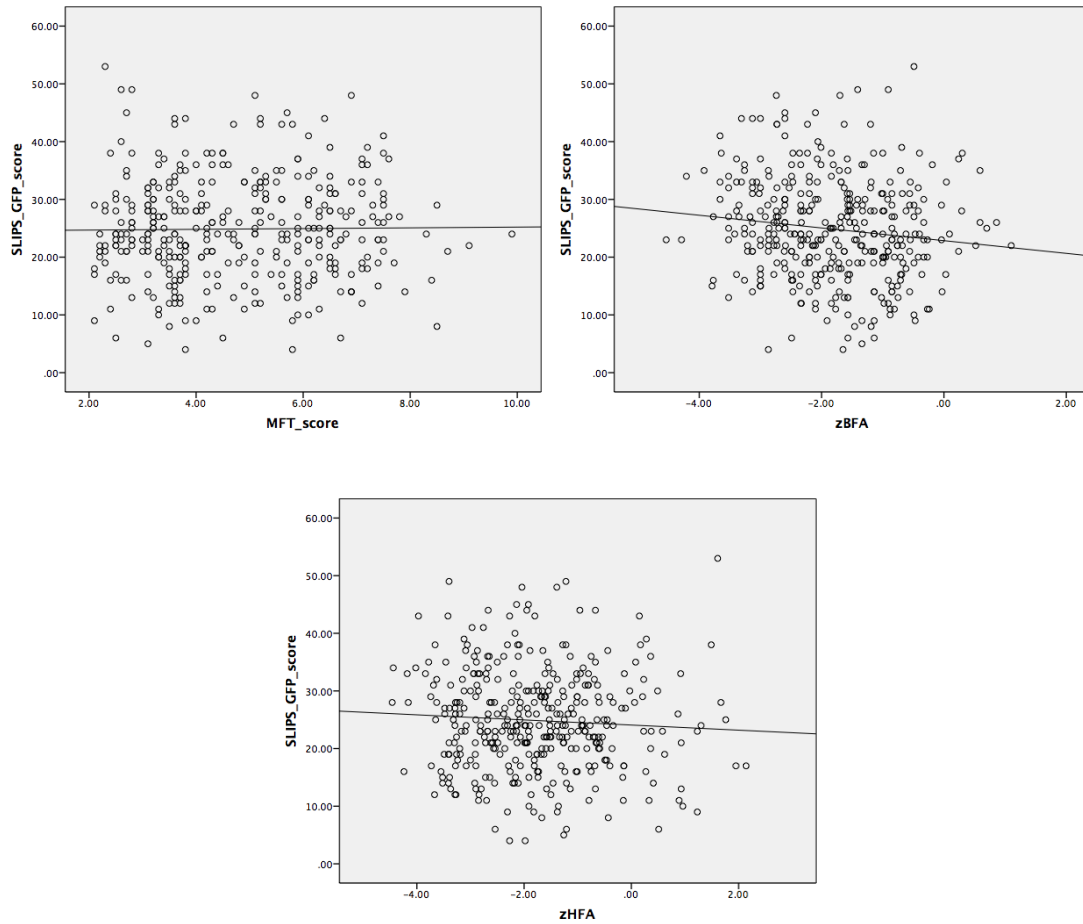


FIGURE A 21: Scatterplots of PH outcome variables against SLIPS-GFP

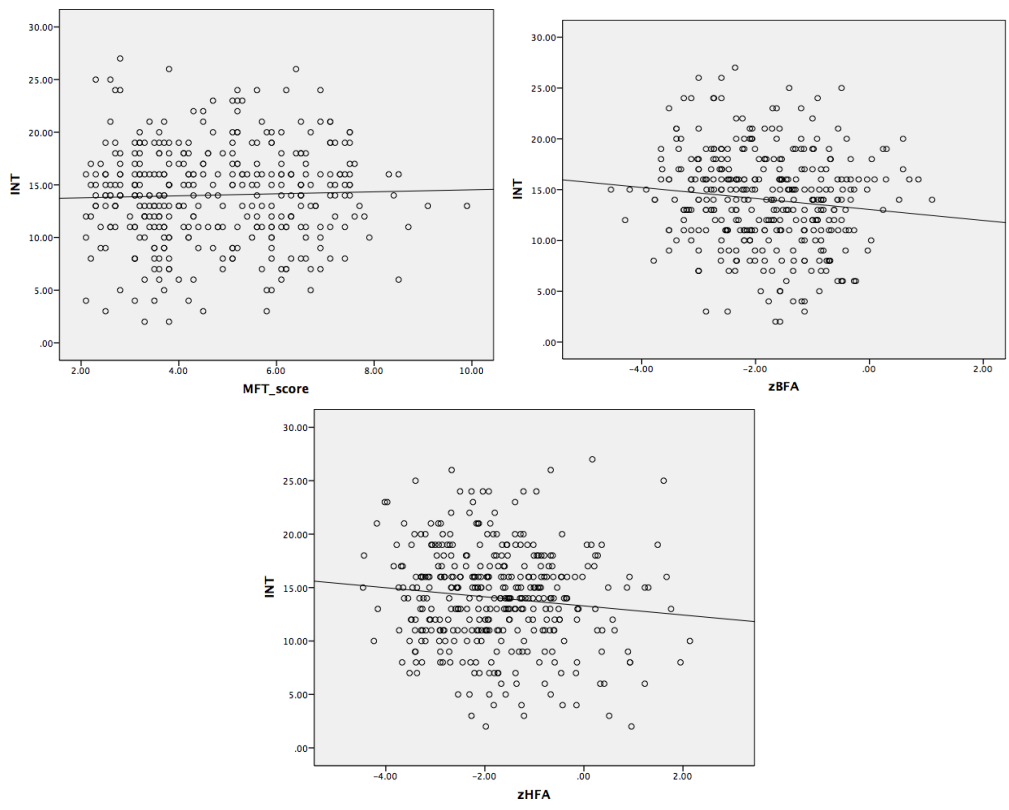


FIGURE A 22: Scatterplots of PH outcome variables against INT score

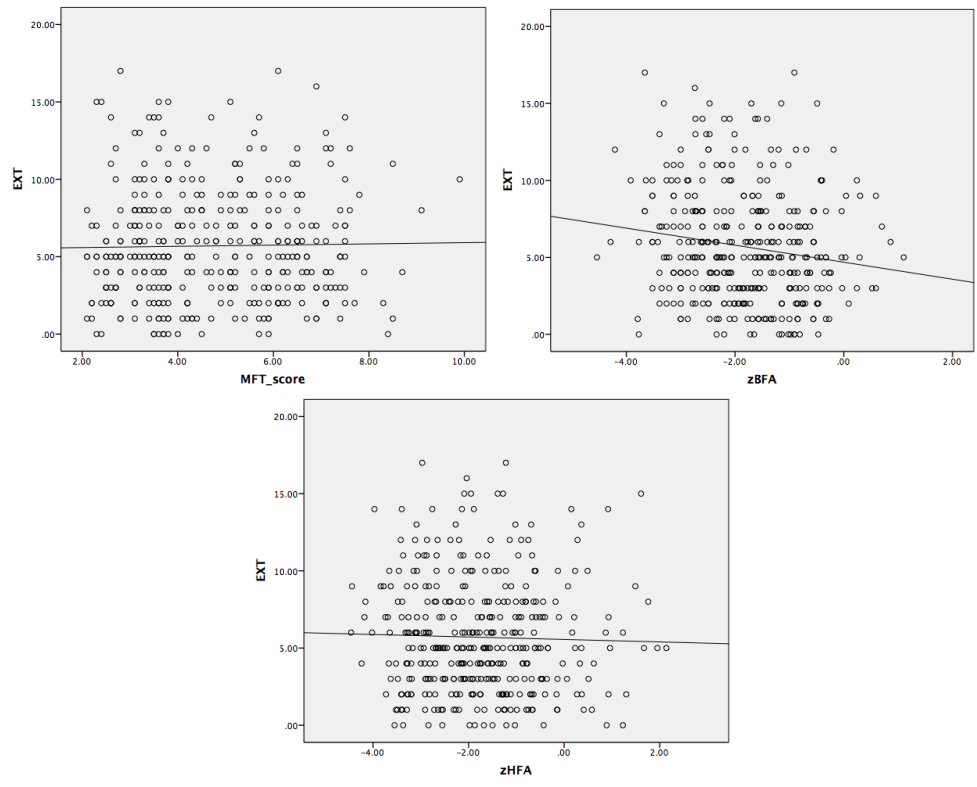


FIGURE A 23: Scatterplots of PH outcome variables against EXT score in Mullaitivu at M1

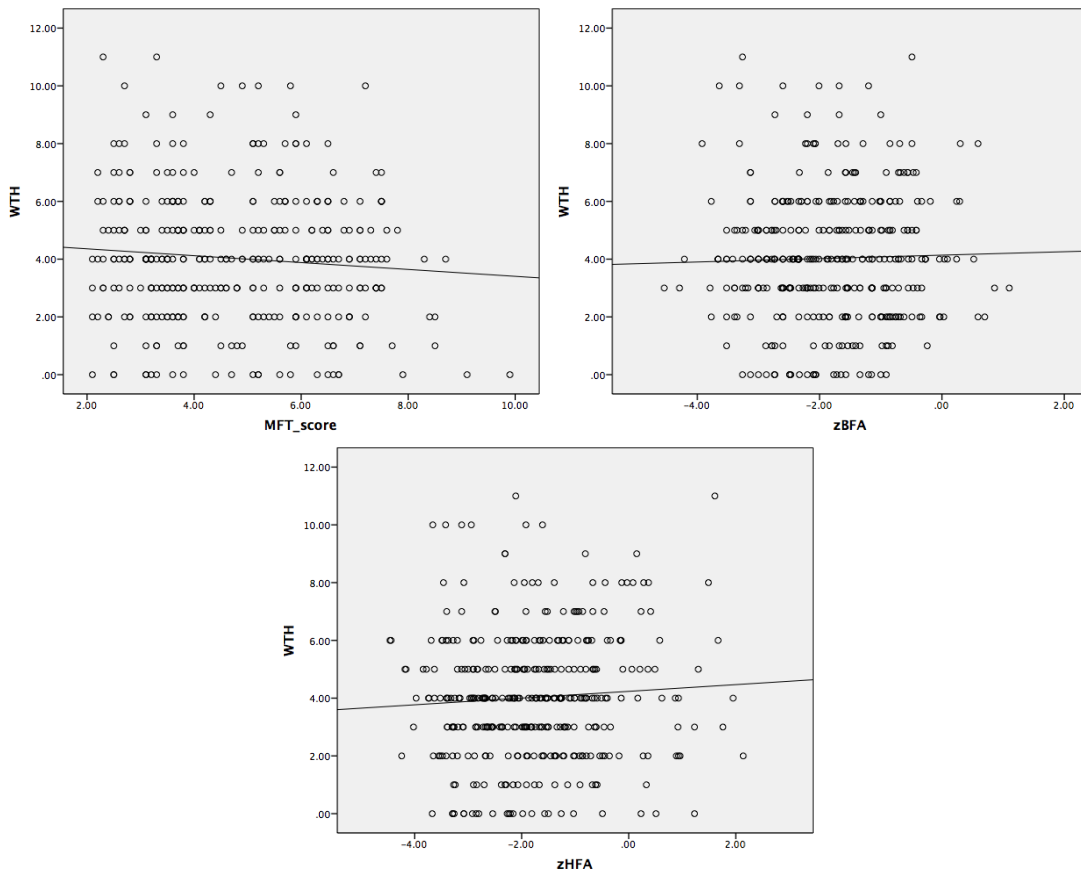


FIGURE A 24: Scatterplots of PH outcome variables plotted against WTH score in Mullaitivu at M1

The scatterplots identified no linear association between the physical and MH outcomes in the Mullaitivu sample. I stratified the sample by gender and also by ethnic group and re-plotted the scatterplots. There were no gender or ethnicity-specific relationships between the PH variable and MH variables. The scatterplots presented no linear, or non-linear, relationships between the outcome variables.

When the sample is stratified by intervention/control status, there are no associations between the PH and MH outcome variables.

D.4.1 Association between PH variables.

Scatterplots identified a linear relationship between HFA and MFT score. The use of Pearson's R requires the data to fulfil certain assumptions. Importantly, the data must show homoscedasticity and the residuals must be approximately normally distributed.

To assess homoscedasticity, the standardised residuals of the HFA are plotted against MFT score.

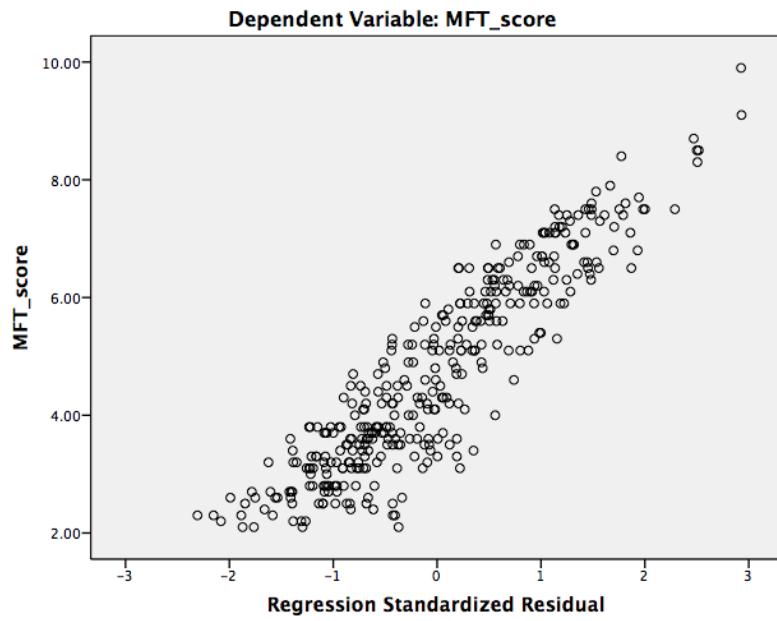


FIGURE A 25: Scatterplot of residuals against MFT score in Mullaitivu at M1

The scatterplot indicates that the samples have homoscedasticity, as there is no increase in the distance of the residuals from the line of best fit as the dependent variable increases.

To verify that the residuals of the regression line are approximately normally distributed, I plotted a PP plot of the regression standardised residuals.

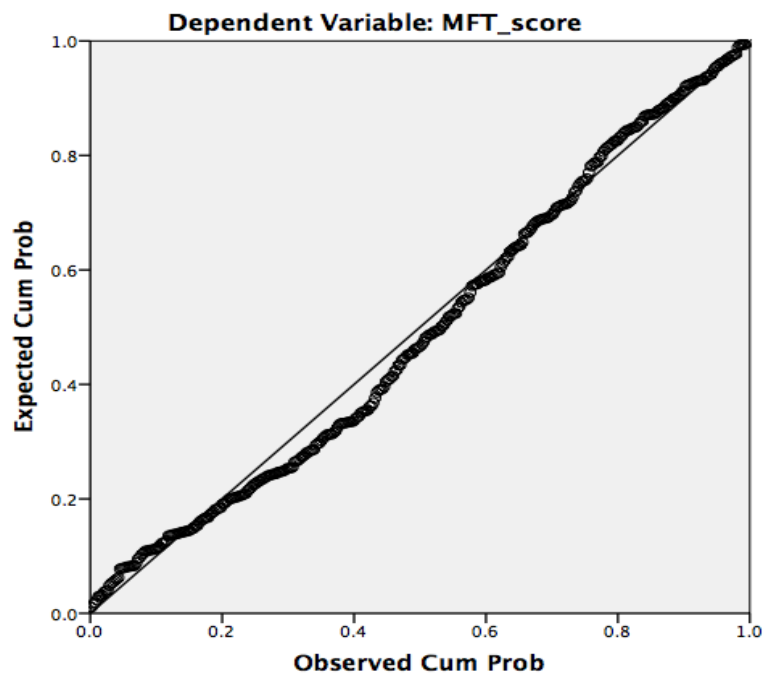


FIGURE A 26: PP plot of the regression-standardised residuals for HFA and MFT score

APPENDIX E

Appendix E.1 Paired t test assumptions

The use of the paired t test on the crude mean scores to identify any significant within group changes requires that difference between the scores at M1 and M2 are approximately normally distributed. Normality was initially assessed visually using histograms, which warranted further analysis.

TABLE A 19: Mean and 95% confidence intervals of the differences between scores at M1 and M2

		SLIPS-GFP	INT	EXT	WTH	MFT	BFA
Intervention (V)	Diff M1- M2	2.12	1.11	1.02	-0.40	0.64	-0.15
		-0.41<- >3.84	0.029<- >2.20	0.28<- >1.75	-0.84<- >0.03	0.26<- >1.01	-0.22<->- 0.09
Control (C)		4.37	2.49	1.82	-0.6	0.087	-0.09
		2.65<- >6.10	1.46<- >3.52	1.04<- >3.52	-0.61<- >0.50	-0.40<- >0.58	-0.18<->- 0.001

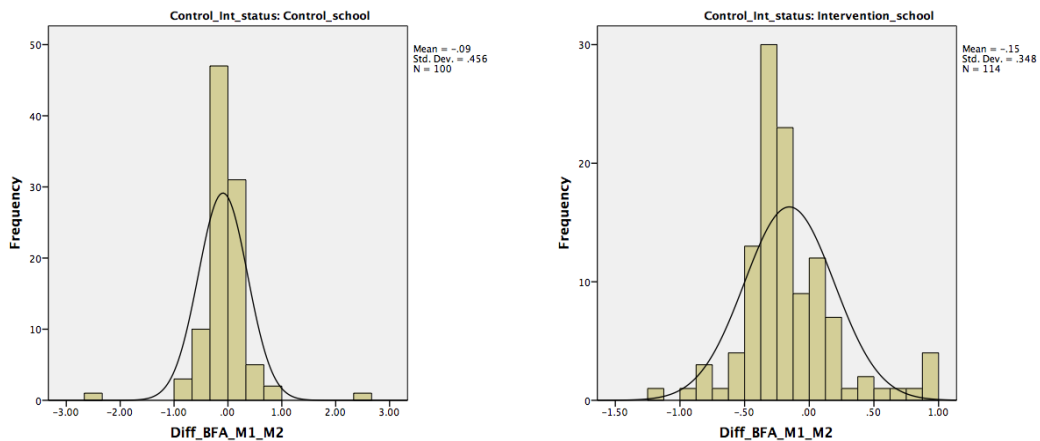


FIGURE A 27: Mean difference between BFA at M1 and M2 between intervention and controls

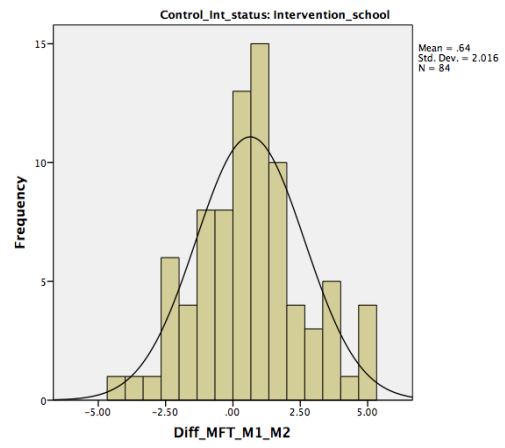
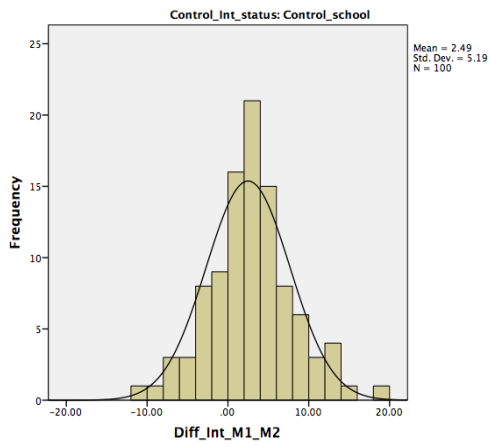


FIGURE A 28: Mean difference between MFT at M1 and M2 in intervention and control groups

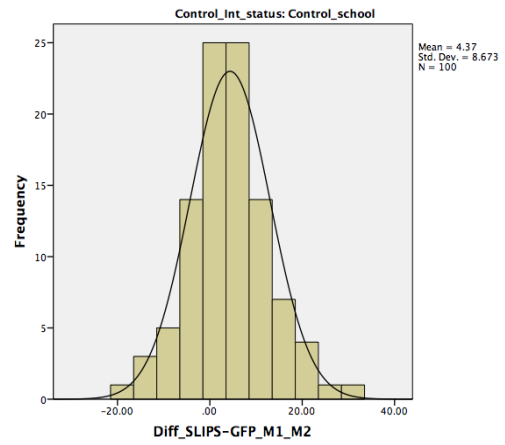
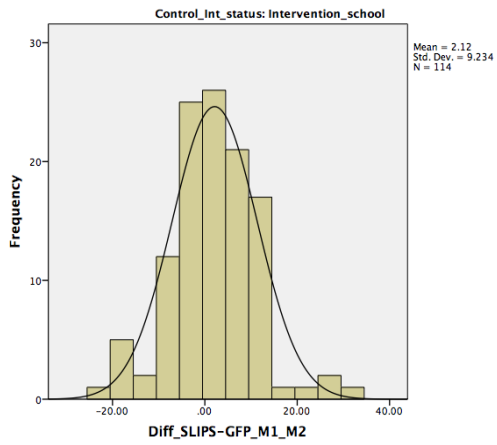


FIGURE A 29: Mean difference between SLIPS-GFP score at M1 and M2 in intervention and control groups

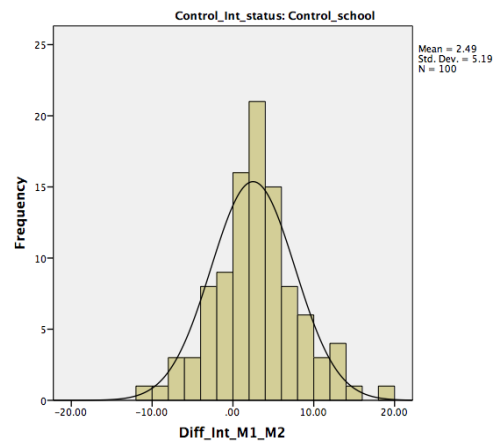
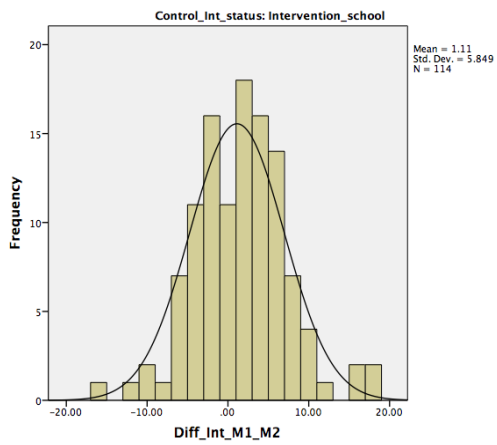


FIGURE A 30: Mean difference between Internalising subscale score at M1 and M2 in intervention and control groups

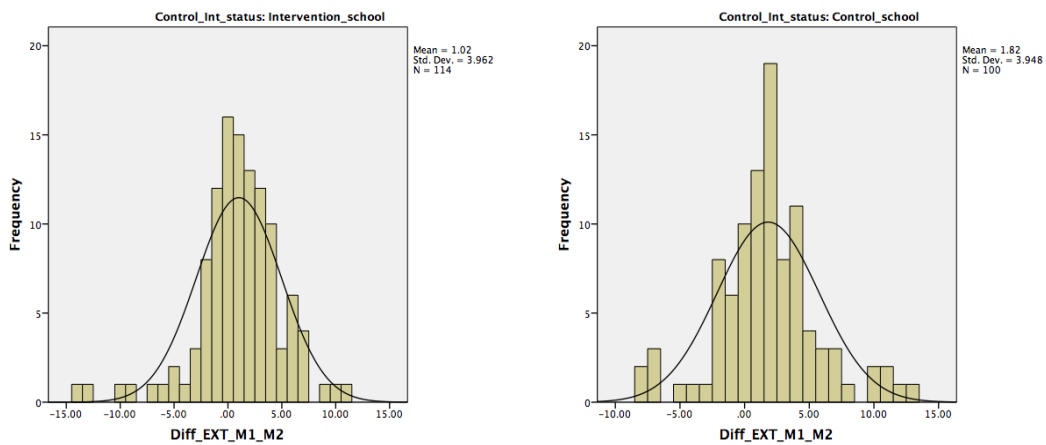


FIGURE A 31: Mean difference between externalising subscale score at M1 and M2 for intervention and control groups

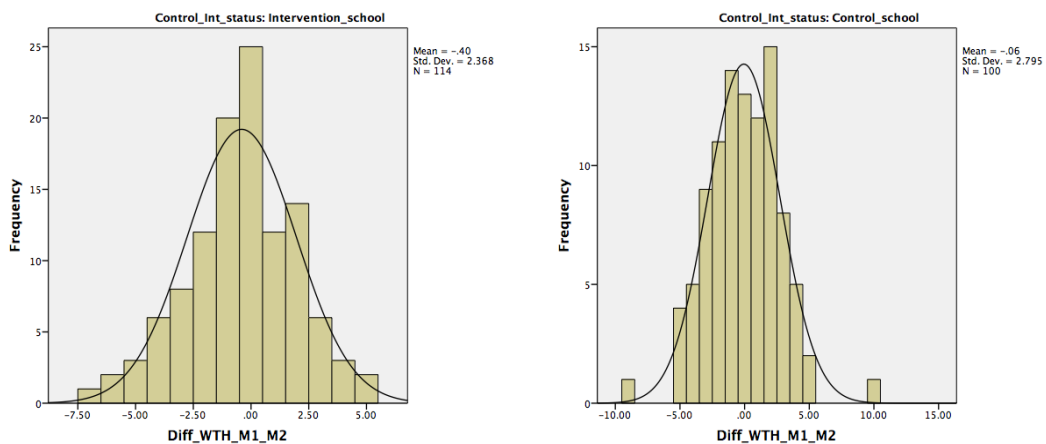


FIGURE A 32: Mean difference between Withdrawal subscale scores at M1 and M2 in intervention and control groups

Normality of the distribution of the mean differences between M1 and M2 in [V] and [C] groups was assessed using the Shapiro Wilk test.

TABLE A 20: Results of Shapiro Wilk test to assess the normality of mean difference between M1 and M2 for all outcome variables, stratified by intervention status

Independent variable	Dependent variable	SW score	Sig.
Intervention [V]	BFA	0.892	<0.001
	MFT	0.986	0.518
	SLIPS-GFP	0.977	0.05
	INT	0.981	0.098
	EXT	0.934	<0.001
	WTH	0.977	0.048

Control [C]	BFA	0.756	<0.001
	MFT	0.986	0.478
	SLIPS-GFP	0.989	0.604
	INT	0.985	0.341
	EXT	0.955	0.002
	WTH	0.973	0.038

The main outcome variables (SLIPS-GFP and MFT score) did not violate normality assumptions.

Normality violations are not of major concern, because the Shapiro Wilk test is particularly stringent and the histograms appear to be normally distributed. The use of a paired t-test is warranted in this example.

Appendix E.2 ANOVA for M2 comparisons (between [V] and [C] groups)

ANOVA was used to assess the difference between [V] and [C] scores at M2. A Levene's test was used to assess whether the samples had HOV and Shapiro Wilk test was used to assess if both samples were approximately normally distributed.

TABLE A 21: Shapiro Wilk test results for the main outcome variables at M2 stratified by intervention and control status

Independent variable	Dependent variable	Skew	S.E.S	Kurtosis	S.E.K	SW score	Sig.
Intervention [V]	BFA	-0.070	0.226	-0.728	0.449	0.985	0.226
	MFT	2.33	0.226	3.486	0.449	0.388	<0.001
	SLIPS-GFP	0.434	0.226	0.261	0.449	0.985	0.215
Control [C]	BFA	0.138	0.241	-0.686	0.478	0.982	0.203
	MFT	2.107	0.241	2.488	0.478	0.955	0.005
	SLIPS-GFP	0.337	0.241	0.283	0.478	0.986	0.377

TABLE A 22: Levene's test for HOV at M2 between intervention and control groups

	Levene Statistic	df1	df2	Sig.
BFA_M2	.831	1	212	.363
MFT_M2	1.45	1	168	0.23
SLIPS_M2	.003	1	212	.956
INT_M2	1.192	1	212	.276
EXT_M2	.117	1	212	.732
WTH_M2	1.678	1	212	.197

All cases had HOV. Normality is not of a major concern, provided the samples have HOV.

Appendix E.3 Stratified M1 and M2 scores for all the major outcome variables in the Mullaitivu study

TABLE A 23: Mean and 95% confidence intervals for all the major outcome variables at M1 and M2, stratified by intervention/control status and gender.

Independent variable		n	Dependent variable											
			SLIPS-GFP		INT		EXT		WTH		MFT*		BFA	
			M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
[V]	Boys	60	23.83 21.62-26.04	21.35 19.21-23.49	13.33 12.12-14.55	11.18 10.04-12.33	5.62 4.80--6.44	4.88 3.92-5.85	3.58 3.04--4.12	4.53 3.94-5.13	5.40 5.01--5.79	4.52 4.10-4.95	-2.10 -2.35--1.85	-1.93 -2.15--1.70
	Girls	54	23.30 21.06-25.53	21.57 19.06-24.10	13.11 11.84-14.38	13.15 11.77-14.53	4.94 4.11--5.78	3.61 2.72-4.50	4.22 3.58--4.87	4.02 3.50-4.54	3.77 3.51--4.04	3.45 3.07-4.89	-1.48 -1.73--1.23	-1.35 -1.59--1.10
[C]	Boys	49	28.65 26.30-31.00	23.10 20.35-25.84	15.47 14.25-16.69	12.49 11.14-13.84	7.71 6.62--8.81	5.49 4.48-6.50	4.18 3.48--4.89	3.96 3.22-4.70	5.48 5.08--5.87	4.60 4.09-5.11	-1.99 -2.28--1.70	-1.86 -2.16--1.57
	Girls	51	23.57 21.06-26.08	20.33 18.09-22.58	13.37 12.07-14.67	11.35 10.20-12.50	5.33 4.12-6.54	3.90 3.05-4.75	4.12 3.54-4.70	4.45 3.78-5.12	3.65 3.27-4.03	4.56 4.04-5.07	-1.46 -1.73--1.19	-1.40 -1.67--1.13

*MFT values for n are lower than for other groups because of missing data.

- [V] boys: 57
- [V] girls: 53
- [C] boys: 48
- [C] girls: 38

TABLE A 24: Mean and 95% confidence intervals for all the major outcome variables at M1 and M2, stratified by intervention/control status and age

Independent variable		n	Dependent variable											
			SLIPS-GFP		INT		EXT		WTH		MFT*		BFA	
			M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
[V]	13	58	23.66 21.44-25.87	20.62 18.30-22.94	13.83 12.69-14.96	12.29 10.95-13.64	5.03 4.15-5.91	3.67 2.84-4.50	3.60 3.04-4.16	4.10 3.58-4.63	4.81 4.36-5.27	4.14 3.74-4.54	-1.89 -2.16--1.62	-1.71 -1.98--1.45
	14	54	23.76 21.47-26.05	22.56 20.20-24.91	12.76 11.40-14.11	11.96 10.71-13.21	5.57 4.78-6.37	5.02 3.96-6.07	4.26 3.63-4.88	4.54 3.91-5.15	4.58 4.18-4.98	3.96 3.46-4.45	-1.70 -1.97--1.43	-1.58 -1.81--1.35
[C]	13	52	27.58 25.22-29.93	22.10 19.28-24.91	14.87 13.79-15.94	11.94 10.59-13.30	7.85 6.73-8.96	5.17 4.14-6.20	3.79 3.12-4.46	3.94 3.27-4.62	4.70 4.26-5.13	4.82 4.33-5.31	-1.81 -2.09--1.53	-1.73 -2.02--1.44
	14	43	23.56 20.89-26.23	20.81 18.56-23.06	13.51 11.92-15.10	11.60 10.40-12.80	4.67 3.55-5.80	3.98 3.09-4.86	4.44 3.81-5.07	4.51 3.72-5.30	4.62 4.06-5.18	4.20 3.65-4.74	-1.69 -2.01--1.37	-1.58 -1.89--1.27

*MFT values for n are lower than for other groups because of missing data.

- [V]13: 41
- [V] 14: 41
- [C] 13: 47
- [C] 14: 35

TABLE A 25: Mean and 95% confidence intervals for all the major outcome variables at M1 and M2, stratified by intervention/control status and ethnicity

Independent variable		n	Dependent variable											
			SLIPS-GFP		INT		EXT		WTH		MFT*		BFA	
			M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
[V]	Tamil	50	23.38 20.95-25.81	20.70 18.20-23.20	13.36 11.87-14.85	12.34 10.82-13.86	5.46 4.59-6.33	3.96 3.10-4.82	3.72 3.08-4.36	4.02 3.51-4.53	4.73 4.33-5.13	4.39 3.94-4.84	-1.64 -1.91--1.37	-1.53 -1.78--1.28
	Sinhala	37	23.32 20.66-25.99	23.51 21.02-26.01	12.59 11.30-13.89	12.4 11.09-13.89	4.95 3.83-6.06	4.91 3.52-6.32	4.16 3.43-4.89	4.62 3.96-5.29	5.36 4.53-6.19	3.34 2.87-3.81	-1.63 -1.97--1.29	-1.51 -1.83--1.20
	Muslim	27	24.30 20.83-27.76	20.04 16.15-23.92	13.85 12.06-15.64	11.19 9.35-13.02	5.48 4.30-6.66	4.00 2.60-5.40	3.81 2.90-4.73	4.33 3.21-5.45	4.86 4.15-5.57	4.88 4.00-5.77	-2.35 -2.71--1.99	-2.08 -2.44--1.71
[C]s	Tamil	95	26.48 24.67-28.30	21.77 19.94-23.60	14.63 13.71-15.56	11.89 10.99-12.80	6.67 5.80-7.52	4.79 4.10-5.48	4.24 3.79-4.69	4.20 3.69-4.71	4.59 4.25-4.94	4.65 4.28-5.02	-1.71 -1.92--1.50	-1.61 -1.82--1.40
	Sinhala	5	18.00 13.88-22.12	20.2 11.99-28.41	10.00 8.04-11.96	12.20 7.44-16.96	3.4 -0.18-6.98	2.60 0.72-4.48	2.40 -0.45-5.26	4.40 1.54-7.26	5.88 4.10-7.67	3.46 2.01-4.91	-1.86 -3.11--0.58	-1.91 -2.92--0.91

*MFT values for n are lower than for other groups because of missing data.

- [V] Tamil: 47
- [V] Sinhala: 35
- [V] Muslim: 18
- [C] Tamil: 81

Appendix E.4 Intention to treat analysis

The use of the statistical tests to assess whether there were any significant differences between participants with and without follow up required the verification of certain assumptions. The assumptions are that the outcome variable scores are normally distributed in the population that the samples are drawn from, and that the two distributions being compared have HOV.

Initially normality was assessed visually using histograms. Initial assessment warranted further analysis. The Shapiro Wilk test was used to assess the normality of the distributions.

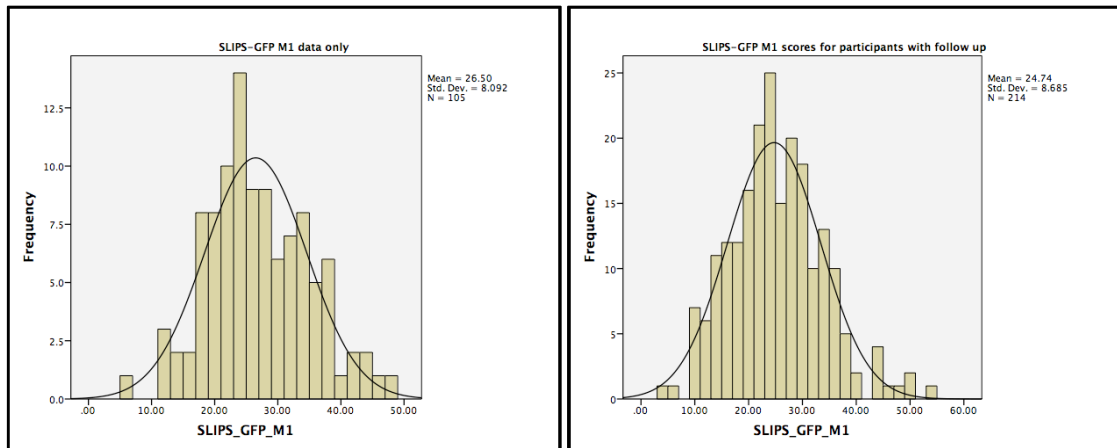


FIGURE A 33: Histograms of SLIPS-GFP score at M1 for participants with and without follow up data

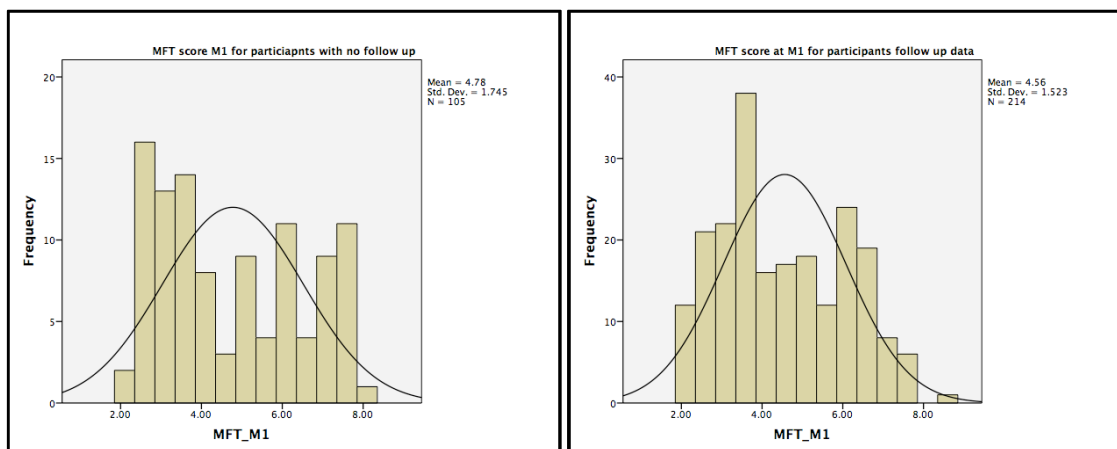


FIGURE A 34: Histogram of MFT scores at M1 for participants with and without follow up data

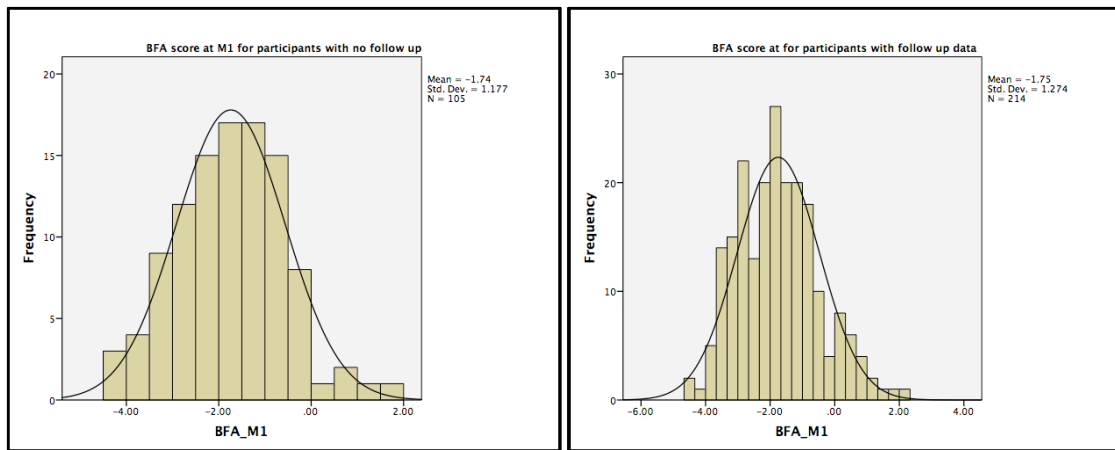


FIGURE A 35: Histograms of BFA score at M1 for participants with and without follow up score

The Shapiro Wilk test was used to verify whether the samples were normally distributed.

TABLE A 26: Table of Shapiro Wilk test results for primary outcome variables at M1 for participants with and without follow up

Group	N	Dependent variable	Skew	S.E.S	Kurtosis	S.E.K	SW score	Sig.
Participants without follow up	105	SLIPS-GFP	0.267	0.236	-0.141	0.467	0.989	0.578
		MFT	0.247	0.236	-1.313	0.467	0.923	<0.001
		BFA	0.280	0.236	0.217	0.467	0.987	0.370
Participants with follow up	214	SLIPS-GFP	0.390	0.166	0.325	0.331	0.987	0.053
		MFT	0.285	0.166	-0.985	0.331	0.959	<0.001
		BFA	0.446	0.166	-0.032	0.331	0.983	0.01

The MFT scores violated normality assumptions, as did the BFA scores for participants with follow up. Standardized transformations had little impact on the distribution of the data.

Homogeneity of variance was assessed using Levene's test.

TABLE A 27: Levene's test statistic and ANOVA comparing participants with and without M2 data collection

Dependent variable	Levene's statistic	Sig	F	Sig
SLIPS-GFP	0.241	0.624	3.013	0.084
MFT	6.992	<i>0.009</i>	1.29	0.257
BFA	0.646	0.422	0.002	0.961

NB: Bold and italic indicates a significant result.

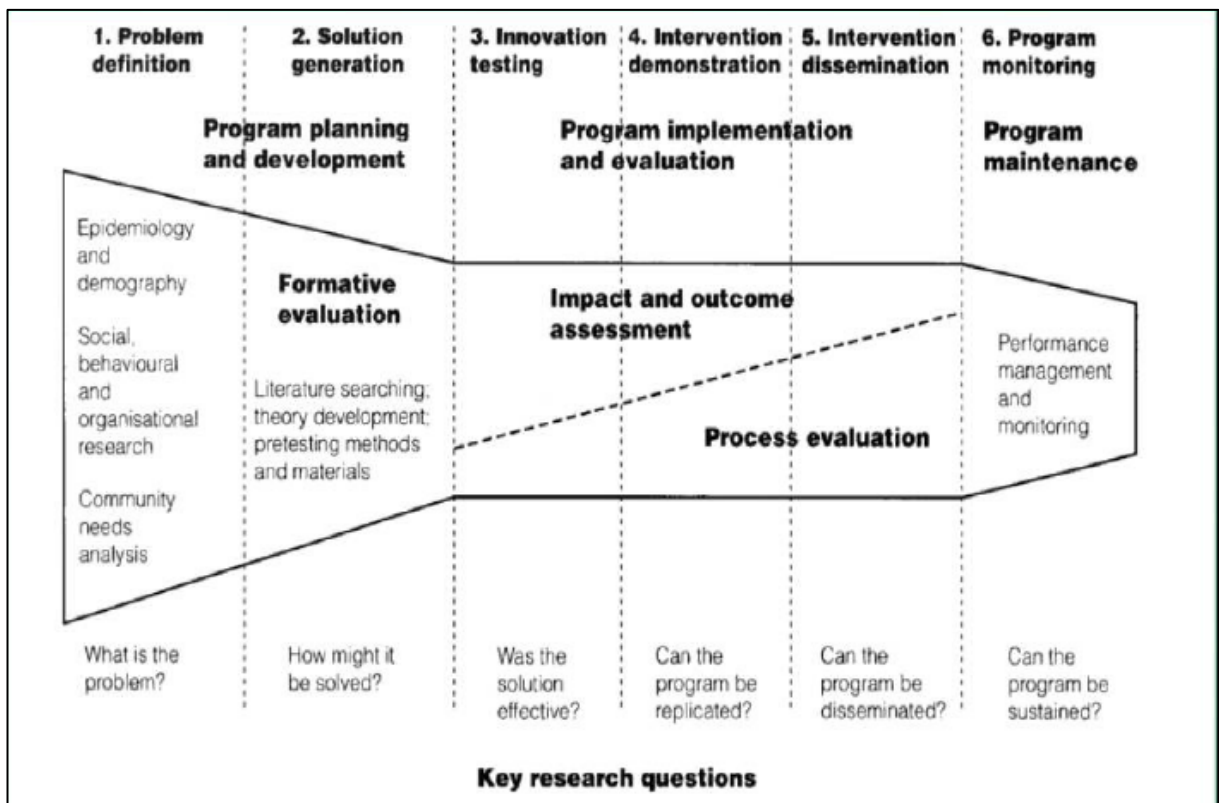
The results of the ANOVA indicate that there are no differences in the M1 scores of those individuals with and without follow up data. Violations of normality are not of major concern in this case so long as the samples have similar variances.

The MFT scores do not have homogeneity of variance and failed the normality tests. This, coupled with the fact that both samples are not normally distributed means that the result of the ANOVA should be read with caution. I used a non-parametric test to confirm the results of the MFT ANOVA. A Mann Whitney U test confirmed the results of the ANOVA ($U = 10599, p = 0.411$) and it can be concluded that there was no significant difference in the mean (and median) scores in any of the main outcome variables at M1 between participants with and without M2 data.

APPENDIX F

Appendix F.1 Bauman and Nutbeam (2014) 'Rocket Model'

The rocket model is a framework that outlines the crucial stages and actions of health promotion programmes and projects, from planning to dissemination. The early stages of the model refer to the development, planning and testing of a proposed health promotion intervention. The latter stages guide the dissemination, adoption and maintenance of effective programmes [179]. Not all health promotion actions will follow this model exactly, but it provides a useful framework by which to plan or retrospectively assess the steps of an evaluation.



Appendix F.2 Dissemination plan

There are several important stakeholders with whom I intend to share the results of the project.

Generations For Peace dissemination plan

- Full project report for internal circulation and publication on GFP website
- Two page report – due end of April 2014
- GFP blog pieces – 3 completed and published online
- Presentation of results to GFP headquarters staff – date TBC
- Teaching session at next GFP advanced training in Amman – date TBC

International dissemination plan

In order to ensure that the results of the project are accessible by all stakeholders, I have decided to pursue several dissemination options: publishing in academic journals, presenting at international conferences, publishing in grey literature and producing a short summary for circulation in Sri Lanka.

I have already made several steps towards disseminating the results of my project:

1. Publishing in academic peer-reviewed journals

Hamilton. A, Foster. C, Richards, J Systematic review of the impact of sport and physical activity programs on the mental health of adolescents in post-conflict settings (2014). *Journal of Sport for Development*

Hamilton. A, Foster. C, Richards, J, Surenthirakumaran, R. A cross sectional study on the relationship between mental and physical health in Kilinochchi, northern Sri Lanka. Article under review at *Conflict and Health*. 2015

2. Presentation of results at international conferences

Hamilton, A, Foster. C, Richards, J, Surenthirakumaran, R. A cross sectional study of mental and physical health in Kilinochchi, northern Sri Lanka. Poster at ISBNPA 2015

Hamilton, A.. The relationship between mental and physical health in northern Sri Lanka. Oral Presentation at International Sociology of Sport Association 2015 (ISSA), Paris

Hamilton, A. The impact of a sport for development and peace intervention on the mental and physical health of adolescents in northern Sri Lanka. Oral Presentation at ISSA 2016, Budapest

Hamilton, A, Foster. C, Richards. J. Cardiovascular fitness in northern Sri Lanka: a cross sectional comparison of rural and semi-urban adolescents. Abstract submitted to International conference on Physical Activity and Health (ISPAH) 2016, Bangkok

Hamilton, A, Foster. C, Richards. J. Assessing the impact of a sports-based intervention on the mental and physical health of adolescents in Sri Lanka (ISPAH) 2016, Bangkok

3. Publication of articles in grey literature websites

Hamilton, A. Sport for Development and Peace: the claims outreach the evidence base. Article for the International Platform on Sport for Development and Peace. September 2014.

4. Simple summary for circulation in Sri Lanka

See overleaf for simple summary of project for translation into Tamil and Sinhalese and circulation within northern Sri Lanka.

F.2.1 Short summary

Background:

Good mental and physical health are crucial for development in childhood and adolescence. Importantly, good physical health as an adolescent limits future risk of non-communicable diseases, such as heart disease and certain types of cancer. Research in northern Sri Lanka indicates that children exposed to conflict, and difficult environments after conflict, are more likely to have poor mental health. Mullaitivu and the surrounding area were some of the worst affected in the Sri Lankan civil conflict. Despite this, very little formal research has been conducted to identify the prevalence of mental and physical health problems in young people. Generations For Peace (GFP) and the Community Development Organisation (CDO) planned a programme using sport to promote peace and reconciliation between Tamils, Sinhalese and Muslim children in Mullaitivu schools. The aim of this research project was to assess mental and physical health in selected schools in Mullaitivu, and measure the impact of the GFP/CDO programmes. We also looked at the relationship between physical health and mental health, to identify potential areas for future interventions.

Methods:

Measurements were taken in schools in Mullaitivu town and in Welioya in March and June 2015. All participants were in Grade 8 or 9 and attending school. Mental health was assessed using a locally-devised mental health assessment tool. Participants also had their height, weight and physical fitness measured. Physical health results were compared with global averages to assess the performance of students in Mullaitivu.

Results:

There was considerable student absenteeism during the data collection process. For those present, the results indicate that young adolescent boys have worse mental health than girls, and that there are no differences in the mental health of the different ethnic groups. Young adolescents in Mullaitivu are less physically fit than similar populations around the world. There is a very high prevalence of stunting and thinness, particularly in young boys. Children in

Mullaitivu are thinner and more stunted than the Sri Lankan National average, and also many other South Asian populations.

It is too early to identify the impact of the intervention, as only very few sessions were delivered during the test. There was an improvement in mental health between March and June. There was a community wide decrease in boy's fitness. Fitness improved in girls not in the intervention. The intervention had no additional benefit on the mental health of participants compared to those not taking part.

Discussion:

There were several issues that affected the data collection and delivery of the intervention. The high rate of absenteeism means that the data may not represent all the students enrolled in schools in Mullaitivu. Adolescents in Mullaitivu are unfit and very many are malnourished. The Ministry of Health and Sport should promote physical activity and sports in schools to improve fitness as fitness is important for functioning and future health.

Very few intervention sessions were delivered because the CDO could not mobilise an adequate number of volunteers. The CDO should look to integrate programmes into permanent structures to ensure that they can be adequately staffed. More research is required to properly assess levels of mental health in Mullaitivu so that effective support can be provided.