

Measuring Impacts of Youth Empowerment:

An International Systematic Review and a
Randomized Controlled Trial with Out-of-
School Youth in Jordan

Matthew Morton

Green Templeton College, Trinity Term 2011

*Thesis submitted in partial fulfillment of the requirements of the
degree of D.Phil. in Evidence-Based Social Intervention in the
Social Sciences Division at the University of Oxford*

Word Count: 88,207

Abstract

Measuring impacts of youth empowerment: An international systematic review and a randomized controlled trial with out-of-school youth in Jordan

Youth empowerment programs (YEPs)—programs that build on young people’s strengths and engage them in decision-making—have gained global attention as a strategy to improve a range of emotional, social, and behavioral outcomes. Guided by the MRC Framework for Complex Interventions, this dissertation employed a mixed-methods approach to synthesize and contribute to the evidence base on impacts of YEPs on adolescent development.

This dissertation includes a systematic review of the effects of YEPs on self-efficacy and self-esteem and a randomized controlled trial (RCT) of a YEP in Jordan. Implementation and process research was also integrated to better understand impact study results and investigate issues for the dissemination of youth empowerment methodology.

In the review, three studies were included from 8,789 citations. The limited data meta-analyzed did not show an effect on self-efficacy (95% *CI* = -0.42 to 1.86; *z* = 1.23). None of the three studies independently showed significant intervention effects on the primary outcomes. Secondary outcomes showed mixed results.

The RCT assessed the effects of an empowerment-based non-formal education program for out-of-school Arab youth. The study included 127 participants, mean age of 15.9 (*SD* = 1.62), and data were collected at baseline and 4-month follow-up. No significant intervention effects were observed for developmental assets (e.g., self-efficacy or social skills). Analysis did show a positive intervention effect on SDQ conduct problems (95% *CI* = 0.13 to 1.48; *d* = .57); effects were mostly attributable to changes in the younger (13-15) age group.

Subgroup analysis with implementation study data indicated that a higher level of program empowerment appeared related to better outcomes. Differences in implementation and attendance across program sites may have diluted intervention effects; relationships between these variables and outcomes should be investigated in future studies with appropriate statistical power and research design. Implications for research and practice are discussed.

Acknowledgments

The work that went into this dissertation was the product of a great many supports and contributions from numerous institutions and individuals. Tremendous inputs were given by those who share a common interest in strengthening programs and policies for young people and the communities in which they live with the help of better research evidence.

In particular, this was all possible due to the financial investments in my personal studies and/or the research projects made generously by the Eckerd Family Foundation, Green Templeton College, and Foundation Open Society Institute (Zug.).

Additionally, staff at Questscope for Social Development in the Middle East and the Jordan Ministry of Education should be commended for their commitment to pursuing bold programs and rigorous evaluation in the spirit of improving opportunities for some of Jordan's most marginalized young people. I give special thanks to the staff, program facilitators, and volunteers in Jordan with whom I worked directly on the research; their hard efforts and local expertise were invaluable to the project. I also very much appreciate new friends in Jordan who showed me the best of the country, including its warm hospitality.

Special appreciation is due as well to the young people that participated in the study, as well as the Youth Advisory Committee (YAC) youth leaders who contributed to the initiative through participatory research. Their enormous value, both in terms of the present and their potential, was constantly obvious and constituted the underlying motivation for this research.

With regards to academic support and mentorship, I could not have asked for better supervision than that provided by Professor Paul Montgomery. His partnership throughout the research, and his constant push for me to perform at my best, made doctoral studies an especially enriching period for my professional development. Furthermore, I owe thanks to colleagues at the Centre for Evidence Based Intervention and Professor Sue Dopson at the Saïd Business School for important practical and scholarly counsel at various stages of the process.

Finally, I am deeply grateful to my friends and family. Through the trials inbred in work toward a doctorate, they have supplied all of the ingredients of perseverance: support, encouragement, humor, and patience. I cannot thank them enough.

Publications and presentations arising from this thesis

Papers published or submitted

Morton, M. & Montgomery, P. (2011). Youth empowerment programs for improving self-efficacy and self-esteem of adolescents. *Campbell Systematic Reviews*.

Morton, M. & Montgomery, P. (in submission). Empowerment-based non-formal education for Arab youth: A pilot randomized controlled trial. (peer-reviewed journal)

Morton, M. (2011). *Stakeholders' report: Strengthening youth opportunities: A pilot impact and process study of empowerment-based non-formal education for out-of-school youth in Amman, Jordan*. Oxford: University of Oxford.

Papers in preparation

Factors that support and impede youth empowerment: A qualitative process study in Jordan.

Applying the MRC Framework for Complex Interventions to evaluation of youth empowerment programs in developing countries.

Transportability of youth development outcome measures to different cultural contexts. Theoretical paper.

Presentations

Jordan stakeholders presentation on results. December 9, 2010. Amman: British Council.

The Office of Her Majesty Queen Rania Al Abdullah. Jordan stakeholders presentation on results. December 8, 2010. Amman, Jordan.

Conference on Out-of-School Youth. Jordan stakeholders presentation on results. May 21, 2011. Amman: Ministry of Education.

Graduate Research Students Conference. The impacts of youth empowerment: An international systematic review and a randomized controlled trial with out-of-school youth in Jordan. May 24, 2011. Oxford: University of Oxford.

Acronyms and abbreviations

Frequently used acronyms:

CBO – community-based organization

CONSORT – Consolidated Standards of Reporting Trials

GSE – general self-efficacy

ITT – intention-to-treat

JD – Jordanian dinar

MOE – (Jordan) Ministry of Education

MRC – Medical Research Council

NGO – non-governmental organization

QS NFE – Questscope non-formal education

RCT – randomized controlled trial

SDQ – Strengths and Difficulties Questionnaire

USD – United States dollar

YEP – youth empowerment program

Frequently used statistical abbreviations:

α – Cronbach's alpha

CI – confidence interval

d – Cohen's *d* (or mean difference where indicated)

M – mean

Mdn. – median

OR – odds ratio

p – significance level

SD – standard deviation

z – z-score

Table of Contents

Chapter 1: Background	14
Introduction	14
Adolescence.....	17
Background.....	17
Neurobiological perspective of adolescent development	21
Ecological perspective of adolescent development.....	28
Youth empowerment programs	33
Concepts and definitions.....	33
Empowerment to increase adolescent program participation	38
Multilevel implications of youth empowerment	41
Levels of participation	41
Examples of YEPs	46
Theory of change	50
Theories behind the theory of change	52
Intervention costs	70
Considering harm and null results	73
Conclusion.....	78
 Chapter 2: Secondary Research: The State of the Evidence from a Systematic	
Review	80
Introduction	80
Background	80
Description of the intervention.....	80
Outcomes of interest.....	82
Previous related reviews	88
Objectives of the review	92
Methodology.....	93
Criteria for considering studies for this review	93
Search methods for identification of studies	96
Data collection and analysis	99
Data synthesis.....	104
Results	106
Results of the search	106
Description of the studies.....	107
Risk of bias in included studies.....	119
Effects of the interventions.....	122
Discussion	126
Summary of the main results.....	126
Overall completeness and applicability of evidence.....	127
Quality of the evidence	127
Potential biases in the review process.....	128
Agreements and disagreements with other studies or reviews	129
Conclusions.....	130
Implications for practice	130
Implications for research.....	132
Summary conclusion	136
References for included studies	137

Chapter 3: Context for Primary Research.....	138
Introduction	138
Country profile	141
Demographics.....	141
Urbanization.....	144
Politics.....	144
Economy.....	146
Civil society.....	148
Youth challenges and opportunities in Jordan	149
Youth population growth and social exclusion.....	149
Education trends	152
Reasons for early school leaving	154
Reasons for youth unemployment.....	156
Educational reform, vocational training, and non-formal education	159
Youth empowerment in Jordan	165
Amenable conditions for youth empowerment.....	165
Challenging conditions for youth empowerment	168
Out-of-school Arab youth	174
Conclusion	179
Chapter 4: Randomized Controlled Trial: Intervention and Methods.....	183
Introduction	183
Intervention	185
Implementing organization	185
Questscope non-formal education.....	185
Program methodology	187
Location within the youth empowerment framework	189
Methods	197
Research coordinating team.....	197
Setting and population.....	199
Research design	199
Protocol	205
Ethics	205
Funding.....	206
Effectiveness versus efficacy.....	206
Multisite selection and preparation.....	207
Sample size calculation.....	209
Recruitment.....	210
Randomization	211
Data collection.....	212
Contamination	214
Primary analysis	216
Subgroup analysis and moderators and mediators	219
Questionnaire	224
Measures.....	227
Primary outcome.....	227
Secondary outcomes.....	230
Implementation measure	236
Chapter 5: Randomized Controlled Trial: Results.....	239
Introduction	239

Enrollment, follow-up, and dosage	240
Baseline comparisons	243
Internal consistency of measures	245
Primary analysis	247
Primary outcome	250
Secondary outcomes	250
Subgroup analysis	251
Age	251
Implementation	254
Attendance	258
Discussion	259
Outcomes	259
Subgroup variables	276
Study strengths	285
Study limitations	287
Generalizability	291
Summary and conclusions	295
Chapter 6: Implementation and Process Study	298
Introduction	298
Methods	299
Implementation study	299
Process study	303
Implementation study results	307
Adherence	309
Exposure	319
Quality of delivery	323
Participant responsiveness	328
Program differentiation	330
Discussion of implementation lessons	335
Heterogeneity	336
Attendance	336
Empowerment	337
Standardization-adaptation balance	340
Component analysis	341
Implementation assessment	341
Process study results	344
Process mechanisms for youth	345
Process promoters for staff	349
Individual and external factors	354
Discussion of process lessons	360
Process mechanisms for empowerment at the point of service	360
Process enablers of empowerment at the staff-level	362
Context factors for empowerment implementation	367
Scaling-up effects on empowerment implementation	369
Concluding remarks on implementation and process	372
Chapter 7: Discussion	376
Introduction	376
Summary of findings	377
Systematic review	377

Randomized controlled trial	378
Implementation study.....	379
Process study	379
Discussion of findings	380
Systematic review	380
Randomized controlled trial	383
Implementation study.....	385
Process study	388
Considering results in the context of extant literature	389
Youth empowerment and self-efficacy	392
Revisiting theory	395
Implications for practice.....	400
Is youth empowerment investment-worthy?	400
Addressing the empowerment-implementation problem	403
Cross-cultural applications of youth empowerment	407
Implications for research	409
Youth empowerment and the evidence-based framework	409
Incubation periods for research and development.....	413
Measurement issues for empowerment as an outcome.....	415
Component analysis.....	418
Dissemination research.....	419
Concluding remarks	420
References.....	424
Appendices.....	458

List of tables

Table 2.1. Characteristics of included studies.....	114
Table 2.2. Primary outcomes from included studies	124
Table 2.3. Secondary outcomes from included studies.....	125
Table 5.1. Comparison of groups by baseline demographics and outcome measure scores.....	244
Table 5.2. Cronbach's alphas for outcome measures.....	246
Table 5.3. RCT outcomes	249
Table 5.4. Comparing change scores between age groups in QS NFE	252
Table 5.5. Separated outcomes for male centers with highest & lowest empowerment scores	257
Table 6.1. Summary of QS NFE implementation fidelity	308

List of figures

Figure 1.1. Dissertation research flow	17
Figure 1.2. Triadic model of motivated behavior	24
Figure 1.3. Bronfenbrenner's Ecological Model	29
Figure 1.4. Model depicting hypothesized relations among community dimensions, social mediators, developmental processes, and desired outcomes in young adulthood	32
Figure 1.5. Hart's Ladder of Participation	42
Figure 1.6. TYPE Pyramid	44
Figure 1.7. Basic theory of change diagram for YEPs	51
Figure 1.8. Theoretical mapping behind youth empowerment programming's theory of change	55
Figure 1.9. Rates of return on human capital investment in disadvantaged children	73
Figure 2.1. Review flow diagram	106
Figure 2.2. Meta-analysis for general self-efficacy.....	123
Figure 3.1. Jordan map	143
Figure 3.2. Jordan population trends.....	143
Figure 3.3. Jordan growth rate trends.....	143
Figure 3.4. Adolescent share of Jordan population over time.....	150
Figure 3.5. Paradox of increased unemployment with investment and growth.....	151
Figure 3.6. Unemployment by education level (aged 15-29), 2005	157
Figure 3.7. Distribution of QS NFE/IFE centers in governorates in Jordan	164
Figure 4.1. QS NFE program cycles	188
Figure 4.2. Questscope non-formal education theory of change	191
Figure 4.3. Research coordinating team.....	198
Figure 4.4. Research process diagram.....	204
Figure 4.5. QS NFE study center locations.....	207
Figure 5.1. Recruitment summary	241
Figure 5.2. RCT flow diagram.....	242
Figure 5.3. Estimated marginal means* for pre-post age group differences.....	253
Figure 5.4. Social supports of friends by attendance	258

Figure 5.5. Box plot: Differences between groups at posttest on conduct problems*	266
Figure 6.1. Interviews diagram	305
Figure 6.2. Overall empowerment survey results (bar graph, left, and box plot, right)	313
Figure 6.3. Impact domain scores	315
Figure 6.4. Impact scores by age group	318
Figure 6.5. Attendance by program sites	320
Figure 6.6. Meaningfulness domain results (bar graph, left, and box plot, right)	329
Figure 6.7. Factors influencing implementation of youth empowerment programming	364
Figure 7.1. MRC Framework: Key elements of the development and evaluation process	411

List of appendices

Appendix 2.1. Databases and websites searched	459
Appendix 2.2. Typical professional outreach correspondence	460
Appendix 2.3. Screening guide.....	461
Appendix 2.4. Coding form	462
Appendix 2.5. Guide for appraising study quality	466
Appendix 2.6. Excluded studies.....	469
Appendix 2.7. Outcomes tables with excluded-school based studies.....	476
Appendix 4.1. Impact study informed consent form.....	478
Appendix 4.2. Distribution of age across RCT sample.....	481
Appendix 4.3. Impact study questionnaire	482
Appendix 4.4. Empowerment measure questionnaire	488
Appendix 5.1. Mann-Whitney U baseline statistics	491
Appendix 5.2. Confidence intervals for primary analysis	492
Appendix 5.3. Diagnostics and transformations of outcomes data	495
Appendix 5.4. Mann-Whitney U primary analysis statistics.....	501
Appendix 5.5. Alternative analysis with peer problems item dropped	502
Appendix 5.6. Full subgroup analysis statistics	504
Appendix 5.7. Moderator analysis results for age.....	514
Appendix 6.1. Qualitative interview informed consent form.....	516
Appendix 6.2. Qualitative interview guides	517

Chapter 1: Background

Introduction

The largest waves of young people in history will soon transition into adulthood. Adolescents (aged 10 to 19 by World Health Organization and United Nations definition), amounted to a population of 1.2 billion in 2003 (UNFPA, 2003). Understanding the most effective approaches for improving developmental outcomes for adolescents is a critical challenge that merits global attention. The period of adolescence is particularly important given its instrumental role in the development of habits and competencies that will affect young people's wellbeing and resilience throughout their lives. Adolescence is also a vulnerable time in which emotions and risk-taking tendencies are amplified (Call et al., 2002; Dahl, 2004; Rutter, 2001; World Bank, 2006b). This dissertation aims to increase empirical understanding of the use of youth empowerment as a strategy for developing psychosocial assets among adolescents.

Youth empowerment has been promoted internationally. The African Union, European Union, United Nations, World Bank, numerous national governments (e.g., United Kingdom's *Youth Matters*), and the philanthropic community are only a few examples of prominent institutions to have explicitly endorsed strategies to increase participation of young people in policy and programming (African Union, 2006; EU, 1999; Rosen & Maureen, 2001; UKDCSF, 2005; UN, 2005; World Bank, 2006b). To some extent, the argument for youth empowerment is based on rights (Freeman,

2005). The United Nations Convention on the Rights of the Child (UNCRC), which assures children the right to be heard and form their own views (article twelve), commonly encapsulates a ‘rights approach.’ This view emphasizes redistribution of power given a perceived injustice embedded in inabilities of young people to exercise their own voice and influence in matters that affect them.

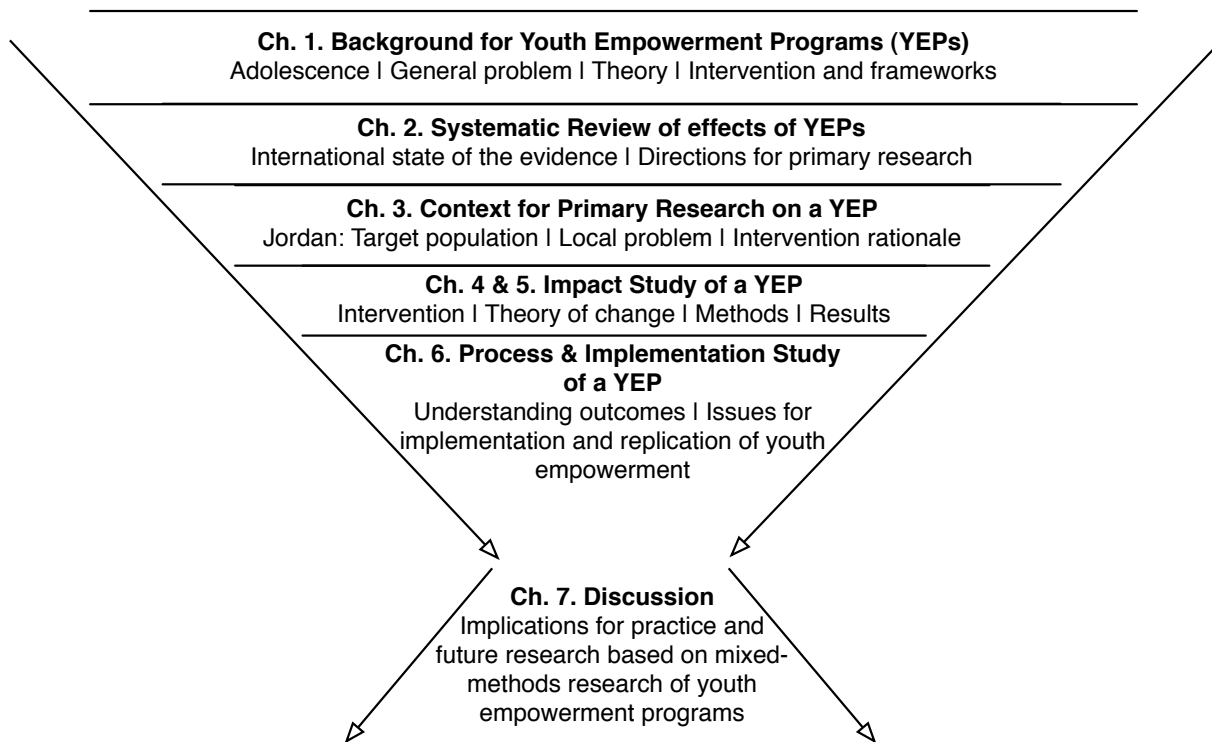
Another approach espouses an instrumental argument. This perspective frames youth empowerment not as a right to be protected, but as a modality for improving youths’ developmental outcomes and strengthening institutions and communities by way of young people’s contributions (Altman & Feighery, 2004; Jennings, 2006; Suleiman et al., 2006a). The rights-based argument for youth empowerment involves a philosophical and political debate. The instrumental argument—i.e., ‘empowerment leads to positive developmental outcomes’—is an evaluative question of causality, which this dissertation aims to assess.

This dissertation will progress from the general to the specific investigating youth empowerment programs (YEPs). The research flow is intended to follow a logical, evidence-based sequence as illustrated in figure 1.1, reflecting stages outlined in the Medical Research Council’s updated framework for the development and evaluation of complex interventions (Craig et al., 2008). The dissertation begins with a broad overview of the problems and theory in which adolescent development and youth empowerment are situated. The background section provides the context for the systematic review reported in chapter six, which followed rigorous Campbell Collaboration standards, of the effects of YEPs on adolescent development. The

review provides an important precursor to the primary trial, as it identifies gaps in existing research and facilitates more informed strategies for how best to conduct primary studies. Chapter three reviews the context for primary research of youth empowerment programming in the Hashemite Kingdom of Jordan. Following from this context, chapters four and five describe the methods and results of a randomized controlled trial of a youth empowerment program, Questscope non-formal education (QS NFE), for dropout youths conducted in Jordan.

Chapter six investigates further into youth empowerment programming by supplementing the impact evaluation with process and implementation study of the same intervention. The research examines the quality of implementation of an empowering program process and environment, and it explores factors that support and/or hinder youth empowerment. Findings from the process and implementation study help to contextualize the results of the impact study and have relevance for existing scholarship on organizational behavior pertaining to the nonprofit sector. The final chapter provides a concluding discussion based on the overall findings of the dissertation. Lessons for evidence-based practice and youth empowerment, including directions for future programming and research, are considered.

Figure 1.1. Dissertation research flow



Adolescence

Background

The range of time in the lifecycle attributed to adolescence has been historically, physically, and culturally malleable. Dahl (2004) has defined adolescence somewhat loosely as “that awkward period between sexual maturation and the attainment of adult roles and responsibilities.” The definition includes both physiological and social parameters of the adolescent experience. As studies indicate increasingly early onset of puberty—especially among African American girls—and delayed initiation of adult social roles, such as marriage and employment, the adolescent period by Dahl’s definition has stretched considerably in modern history (Dahl, 2004; Euling et al., 2008; Kaplowitz et al., 2001). As such, Dahl contends that although the adolescent experience typically lasted 2 to 4 years

through most of human history, modern biological and social circumstances translate to an adolescent stage that lasts 8 to 15 years, rendering common interchangeable application of the terms “adolescents” and “teenagers” somewhat misleading.

This increased period of adolescence carries both opportunity and vulnerability. On one hand, the majority survives adolescence without detrimental calamity (Wald, 2003), and the protracted period enables more time for skill development, learning, exploration, and relationship formation. On the other hand, it means expanded duration of heightened susceptibility to behavioral and health risks associated with adolescence. Hall (1904), one of the earliest researchers of modern adolescent psychology, coined the description of adolescence as a period of “storm and stress,” which came to encapsulate common conceptualization of adolescence as a particularly turbulent period for many.

As young people experience adolescence, some confront particularly difficult struggles. Recent cross-national self-report data demonstrates especially high youth delinquency rates in Western European and Anglo-Saxon countries (Enzmann et al., 2010). In the United States (USA) during 2008, 2.11 million persons under the age of 18 were arrested (Puzzanchera, 2009). Nearly as many young people drop out of school in the USA (20-25%) as those that obtain bachelors degrees (28%) (Wald, 2003). International efforts to compile indicators on wellbeing of children and young people have only begun to analyze adolescence as a separate population group. Nevertheless, global statistics reported by UNICEF (2005) indicate that

adolescent substance abuse is rising and millions of adolescents are working in hazardous or abusive environments. In other words, while many young people will make at least minimally successful transitions to adulthood (Masten & Garmezy, 1985; Wald, 2003; Werner & Smith, 1992), a large number face problems that could jeopardize their future and have negative repercussions for broader society.

The price of neglecting young people's healthy development can be substantial. A ten-year longitudinal study of children in London found that costs for children with conduct disorder were ten times higher than costs associated with individuals without such problems by the time they reached age twenty-eight (Scott et al., 2001). The largest costs were incurred from crime, followed by extra educational provision, foster and residential care, and state benefits. Meanwhile, Rutter and Smith (1995) have documented sharp rises in psychosocial disorders among adolescents from mid to late 20th century for many developed countries. Although emotional and behavioral problem trends for several countries, such as the United States, the United Kingdom, and the Netherlands, have begun to plateau or modestly decline in more recent years, they remain at concerning levels (The Nuffield Foundation, 2009). If YEPs can make a contribution to circumventing the long-term consequences of social behavioral problems, the cost-savings to tax payers may be considerable.

More importantly, those promoting youth empowerment efforts are concerned predominately with more than expressions of antisocial behavior. As Pittman (1999) has popularly stated, "problem-free isn't fully prepared." Many

youths avoid the most nettlesome experiences associated with adolescence but still struggle to meet the increasingly diverse demands of a competitive global economy or integrate fully with civil society during adolescence and transitions to adulthood. These challenges are particularly pronounced, for example, in regions like the Middle East and North Africa where young people constitute the largest proportion of society (50-65% aged 24 and under), and, yet, working-age youth also have the highest unemployment rates (25-40%) (Fuller, 2003).

Proponents of empowering approaches to youth services contend that young people cannot be adequately prepared without a focus on psychosocial development and effectively capturing young people's interest (Kirby & Bryson, 2002; Larson, 2000). Indicators suggest that society has fallen short to this end for sizeable numbers of youth. For example, in a cross-temporal meta-analysis of 72 samples of American college students (total $n = 13,732$), Konrath and colleagues (2010) found a 40% reduction in empathy among American young people since 1979, with the greatest reductions occurring over the last decade. An earlier study that compiled a random sample of 16,000 electronically recorded moments in the daily experiences of 392 middle school youths found that youths reported being bored for more than 27% of those random moments (Larson & Richards, 1991). A longitudinal study ($n = 354$) of rural middle school students in the USA by Sharp and colleagues (2006) indicates that adolescents' self-regulated motivation and boredom during free-time increases with age, adolescents with less parental involvement tend to report higher levels boredom. A longitudinal study of a more diverse, national, American sample ($n = 1,215$) found that adolescents reporting chronic boredom showed significantly

lower psychological wellbeing in terms of self-esteem, locus of control, and emotions about future prospects, than adolescents reporting chronic interest (Hunter & Csikszentmihalyi, 2003). Such indicators could highlight a need for youth interventions to achieve more than just reducing delinquent behavior. Programs need to challenge, engage, and equip young people to develop personal assets to succeed and contribute meaningfully.

Neurobiological perspective of adolescent development

This section turns to the knowledge base concerning adolescent development from a natural sciences perspective and its relationship to intervention strategies. Sociological and psychological explanations of youth development by themselves provide essential but partial understanding of adolescent behavior. Important physiological changes are also taking place and interacting with environmental factors to stimulate behavior (Petersen, 1988), and much of this activity takes place in the brain. The nature of neurobiological changes and implications during adolescence are not fully understood, and researchers are careful to caution against overly simplified explanations of very complex processes (Cauffman, 2004). Nevertheless, research continues to advance and yield valuable insights for adolescent intervention and evaluation.

Much of the attention to neuroscience for adolescence has focused on risky behavior. While a lot of emphasis is placed on the formative period of early childhood (infancy to preschool age), Patterson and colleagues (2000) suggest adolescence as a second critical interval in which problem behaviors develop. From

a physiological perspective, risks may be particularly pronounced for youths with early onset of puberty. A Finnish survey involving 36,549 respondents (aged 14-16) found significant associations between earlier puberty and negative internalizing mental health outcomes such as depression and anxiety (girls) and externalizing symptoms such as substance abuse, bullying, and truancy (girls and boys) (Kaltiala-Heino et al., 2003).

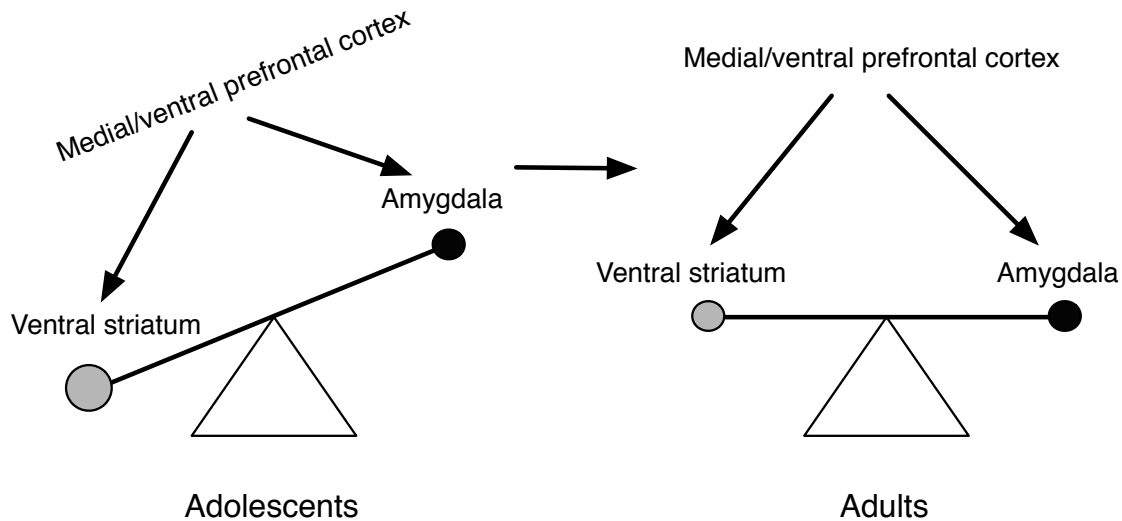
As a premise for this research, studies on adolescent risk behavior have indicated three general points: (1) adolescents have a greater propensity for risk-taking behavior than children or adults; (2) adolescents are more likely to engage in risk-taking behavior while in a group than while alone; and (3) heightened risk-taking tendencies are biologically and socially driven, not cognitively driven (Steinberg, 2004, 2008). With respect to the last point, it is, in other words, not as if teenagers simply lack understanding of the consequences of risk-taking behavior. Research has found the logical reasoning and basic information-processing abilities of teenagers to be comparable with adults (Steinberg, 2008). Studies, in fact, indicate that adolescents have *higher* risk perceptions than young adults (Haase & Silbereisen, 2011). After reviewing the topical research, Boyer concluded that, if anything, adolescents may “overestimate their personal vulnerability” (Boyer, 2006). As such, it is perhaps unsurprising that systematic reviews have shown interventions based on the supposition that adolescents can be educated into avoiding risky behavior such as sexual activity (Underhill et al., 2007) and drug use (Gates et al., 2006; White & Pitts, 1998) to be largely ineffective strategies for changing behavior.

Ernst and colleagues (2006) have found a basis for risk-taking behavior in the unique characteristics of brain development during adolescence, represented by what they describe as the triadic model of motivated behavior. Their research has synthesized evidence from animal and human behavioral observations as well as neuroimaging with youths and adults. The model suggests that there are three interacting circuits in the adolescent brain that direct young people's proclivity to particular behaviors. The *ventral striatum circuits* support sensation-seeking tendencies—reward processes and approach behavior. Conversely, the *amygdala circuits*, sometimes called the 'behavioral brake,' foster avoidance responses. Believed to be responsible for regulating between the ventral striatum and amygdala, the *prefrontal cortex circuits* enable cognitive control facilitating the relative contributions of reward and avoidance signals to behavior.

The existence of these systems is not unique to adolescents, but their relative contributions are believed to differ between adolescence and adulthood. In adulthood, the ventral striatum and amygdala are more balanced. In adolescents, however, the amygdala plays a weaker role, and the prefrontal cortex is believed to be underdeveloped and thus less effective in its regulatory capacity. Moreover, the tilt toward sensation-seeking drivers in adolescence is associated with a period of pubertal maturation, which is linked to heightened emotional intensity (Dahl, 2004). Impulsiveness also has a biological foundation through low self-regulatory control, peaks during adolescence, and increases an individual's susceptibility to risk-taking behavior (Haase & Silbereisen, 2011). These conditions converge to form a compelling biological basis for behavior that is unique in its proportions to

adolescence. Ernst and colleague's (2006) depiction of the expected differences between how these circuits interact to contribute to behavior in adolescents and adults is shown in figure 1.2.

Figure 1.2. Triadic model of motivated behavior



Source: Ernst et al., 2006

The purpose of this section is not to focus on individual-level contributors to risky behavior but rather to posit that the same psychobiological characteristics of adolescence that heighten risk-taking propensities could also provide one rarely considered aspect of rationale for the youth empowerment intervention approach. Given the natural proclivities of adolescents toward risk-taking, it may be misguided for interventions to aim to persuade young people into avoiding risk-taking behavior. Rather, a programmatic orientation toward helping young people *channel* natural risk-taking tendencies (internally compelled by heightened impulsivity and reward-seeking biological dispositions) toward positive challenges and

opportunities may be better grounded in scientific understanding of the biological mechanisms that contribute to adolescent behavior.

The risk-avoidance approach to adolescent intervention is a socially constructed, negative orientation toward an otherwise neutral neurobiological phenomenon. Many interventions designed through this paradigm may be ineffective in changing youth outcomes in part because adolescents simply find other informal and unhealthy channels to exercise natural propensities toward reward-driven expression. Interventions that facilitate opportunities for young people to take positive, rewarding risks in leadership, decision-making, and contribution, on the other hand, may be more likely to help young people develop their strengths while curbing unhealthy expressions of risk-taking tendencies.

As Romer and colleagues (2010) suggest with respect to drug related risk-taking behavior, “preventing the use of drugs of dependence, perhaps by encouraging engagement with other activities that also stimulate the dopamine system...could also be a way to safely channel sensation seeking.” In other words, if a hypersensitive reward system is a natural aspect of the adolescent experience, then interventions would potentially need to incorporate strong, dynamic methods for stimulating that characteristic of young people in order to effectively engage and develop them, as well as deter them from unhealthy alternatives for risk-taking behavior.

While closer attention to the body of evidence concerning biological contributors to adolescent behavior adds a vital perspective into theory of change

development for interventions, there are important caveats. As with social and psychological explanations of youth behavior, physiological explanations attempt to make sense of what are in fact abundantly complex and dynamic processes that furthermore vary between individuals and interact with other influences, including multifarious social and ecological systems.

Additionally, one can deduce markedly different intervention implications from the same premise of adolescent neuroscience (that adolescents are biologically driven toward higher risk propensity). For instance, Romer and colleagues (2010) and this dissertation hypothesized that, because propensity toward risk-taking is ostensibly a biological ‘given’ during adolescence, positive activities that facilitate *rechanneling* of reward-seeking tendencies could improve developmental outcomes. Steinberg (2008), however, used the same premise of adolescent brain development to recommend interventions that create *limitations* on young people’s opportunities to indulge risk-taking tendencies in harmful ways—e.g., better enforcement of laws governing the sale of alcohol, raising the driving age, and expanding access to contraceptives.

The rechanneling posture is more strength-based, aiming to capitalize on young people’s reward-seeking tendencies as a programmatic opportunity, whereas Steinberg’s response better reflects a deficit-based model, essentially proposing to protect youth from themselves while focusing little on positive development. They are divergent approaches rooted in the same background logic. Nevertheless, it is also conceivable that both approaches may be legitimate responses to the nature of

adolescence and could play complementary roles. They may not be mutually exclusive suppositions.

Furthermore, it is also not clear that YEPs can provide sufficient reward or sensation-seeking stimuli to adequately engage vulnerable young people or replace unhealthy risk-taking outlets. Given the evidence suggesting that adolescent risk-taking tendencies are heightened in groups, it is further plausible that YEPs (interventions involving peer groups) could also unintentionally facilitate the aggregation of peer groups that foster unhealthy risk-taking behavior outside of the program. Such negative peer group effects have been documented by Dishion and colleagues (1999) with program involvement increasing delinquency, and by Lorente and colleagues (2004) with sports team involvement linked to increased alcohol abuse. As such, while YEPs provide a theory-driven approach that could effectively respond to the growing body of evidence for adolescent brain development, failure of important assumptions to materialize could jeopardize the hypothesis that YEPs will have a positive impact on developmental and behavioral outcomes.

Importantly, while various disciplines, such as psychobiological research, cognitive development research, emotional development research, and social development research, have largely been pursued independently, Boyer (2006), among others, argues that these different perspectives interact to influence young people's attitudes and behaviors. As mentioned above, for example, experimental research suggests that the mere presence of peers can increase adolescents'

propensity toward risk-taking (Gardner & Steinberg, 2005), indicating that social interactions with latent biological characteristics can affect behavior.

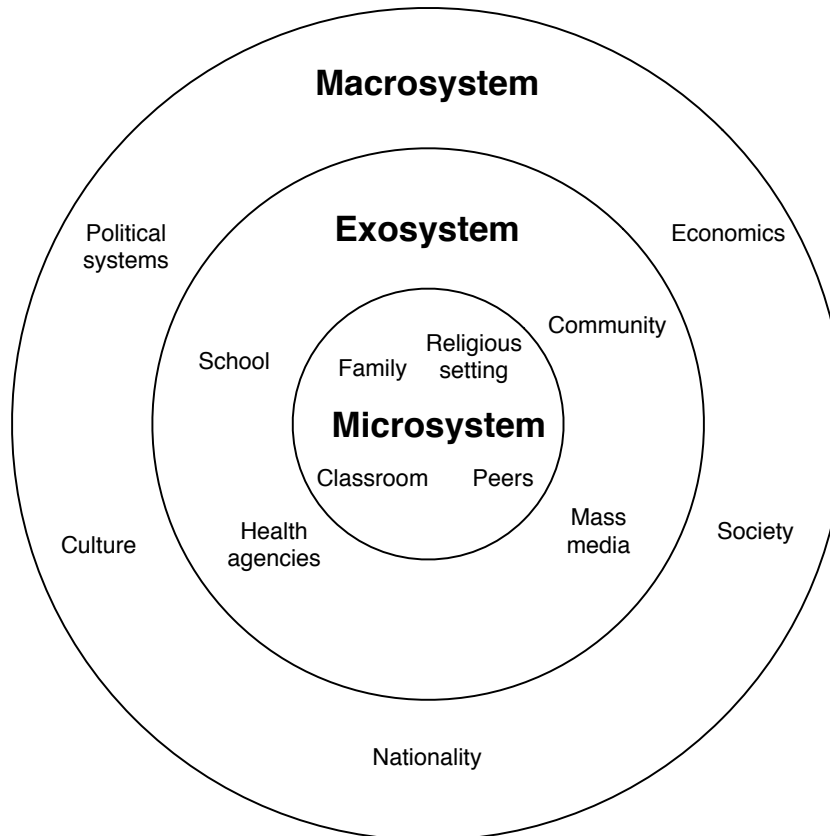
Experimental research by Haase and Silbereisen (2011) manipulated adolescents' affect by exposing them to positive, negative, or neutral sounds and images, according to the experimental arm to which they were assigned, and found that youth exposed to (and reporting) positive or negative affect (but especially positive) were statistically significantly more inclined to express willingness to participate in risky behaviors (e.g., riding with a drunk driver, engaging in a fight, or having unprotected sexual intercourse) than adolescents with neutral affect. In other words, immediate social stimuli can have an important influence on young people's attitudes and propensity toward certain kinds of behavior. The following section transitions to the socio-ecological factors expected to contribute to adolescent development and introduce opportunities for theory-driven intervention.

Ecological perspective of adolescent development

Socially, adolescence presents a new season for young people's interactions. As adolescents assume increasing levels of independence from the home, the interdependencies with peers, schools, and communities generally do not supplant family influences, but they do become more pronounced and competitive with them (Bogenschneider, 1996; Connell & Halpern-Felsher, 1997). With enlarged independence, adolescents also begin to explore new identities; they practice and start to assume some of the roles and responsibilities of adulthood (Eccles et al.,

1997). Because adolescence constitutes a very different part of the life cycle than childhood—with different developmental needs—distinctive kinds of supports and opportunities are required.

Figure 1.3. Bronfenbrenner’s Ecological Model



Although some have insisted that the responsibility for antisocial behavior belongs squarely to the individual youths themselves (Duke, 1978), the larger body of research evidence on risk and protective factors weighs more heavily in favor of the influence that environments have on young people’s behavior (Bishop, 2004; Blum, 1997; Brooks-Gunn, 1997; Leventhal & Brooks-Gunn, 2003a; Rutter, 2001). The literature does not dismiss the role of young people as capable agents within

their own lives and the influence of personal behavior, but it does suggest that environments tend to have powerful direct and indirect influences—though they are not fully understood—on youth behavior. The importance of context for children and youths is perhaps most eminently encapsulated by Bronfenbrenner's (1979) ecological model (see figure 1.3), which draws attention to multiple layers of environmental factors that ripple out from the individual.

Current evidence suggests that, while families have a dominant influence on younger children, as they age into adolescence, peer influence plays a much stronger role alongside the moderating effects of strong parent connections (Aseltine, 1995). For example, Mounts and Steinberg (1995) found in a one-year longitudinal study of 9th to 11th graders in the USA ($n = 500$) that changes in grade point average (GPA) and drug use were predicted by friends' grades and drug use; these effects were moderated by adolescents' reports of authoritative parenting. Although attention often concentrates on the negative impact problematic peer groups can have on young people, YEPs have an opportunity to facilitate *positive* peer associations around young people that could expand their social capital and networks in favor of more positive outcomes. Research further indicates that young people particularly from higher-risk family situations tend to associate themselves and their identities more closely with school and community ties, again suggesting a potentially powerful role to be had by positive interventions outside the household (Rutter, 2001).

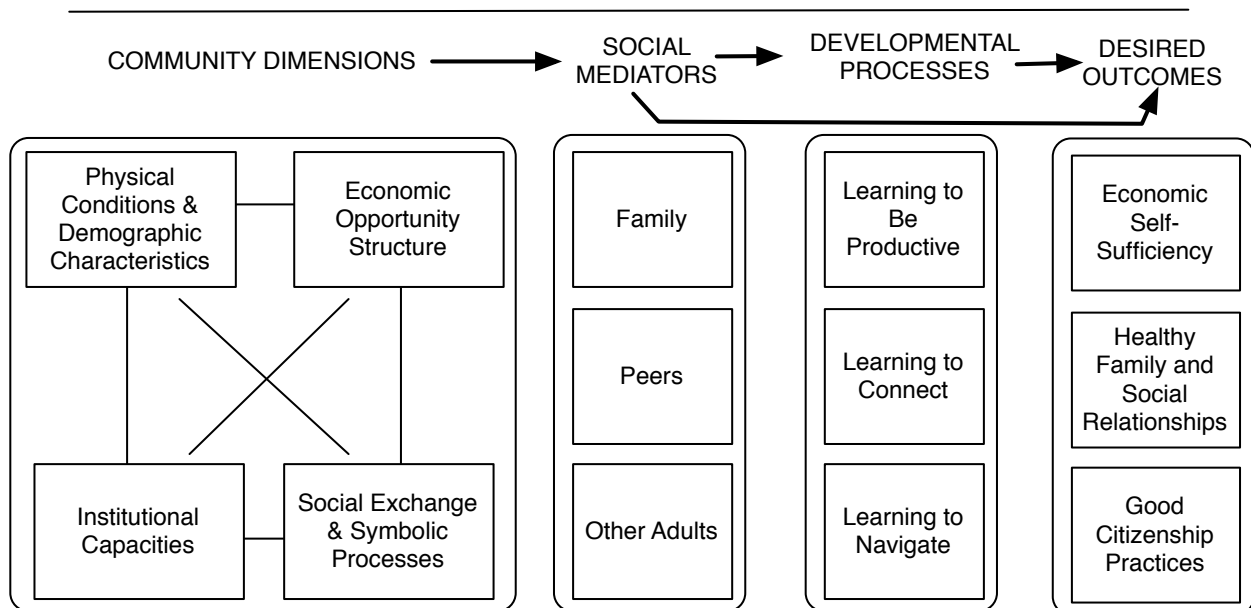
Building on ecological perspectives like Bronfenbrenner's, Connell and colleagues (1995) developed a conceptual model that hypothesizes change pathways between different aspects of a young person's ecological reality (see figure 1.4). Connell and colleagues note that the diagram is intentionally unidirectional because it tries to hypothesize outcomes, but influences could be reverse directional as well. Going by this framework, YEPs may have a particularly important role to play in community dimensions of 'institutional capacity,' by adding to young people's developmentally-focused institutional connections, and 'social exchange,' by fostering positive, asset-building interactions with supportive adults and other youths. As such, the model helps to articulate the potential role of YEPs' theory of change within an ecological context.

On the other hand, the framework also illustrates the challenging contextual realities in which YEPs' change processes are situated. The diagram reminds us that YEPs are still not likely to change all of the community dimensions that contribute to youth outcomes, and they constitute but one force among many in a young person's environment that support or compete with YEPs' influence on their healthy development. Put simply, there are more factors at work in an adolescents' development than their experience with a YEP, and so even the best of YEPs should be wary of overstating their expected impact.

A seeming paradox in the youth empowerment literature is a simultaneous recognition of the importance of individual agency—driven by beliefs of self-efficacy—and a theory of change that posits that a young person's individual agency

is influenced by an ecological environment often beyond his/her control, and therefore self-efficacy can be changed by intervention that changes the right aspects of a young person's social environment. Bandura (1986) applies *social cognitive theory* to offer some coherence to the interaction between two seemingly contradictory concepts: an ecological perspective and an individual agency perspective of youth behavior. Bandura suggests that human agency operates within a broad network of socio-structural influences, and that, "given the dynamic bi-directionality of influence, social cognitive theory rejects a dualism between human agency and a disembodied social structure" (Bandura, 2006a, p. 5). In other words, a youth's environment can influence his or her behavior, but the youth's behavior can also influence his or her environment.

Figure 1.4. Model depicting hypothesized relations among community dimensions, social mediators, developmental processes, and desired outcomes in young adulthood



Source: Connell, Aber, and Walker (1995)

Youth empowerment programs

Concepts and definitions

Defining YEPs. This dissertation defines youth empowerment programs (YEPs) as interventions that regularly involve young people as partners and participants in the decision-making processes that determine program goals, planning and/or implementation. With the support of caring adults, YEPs engage young people in program leadership as a characteristic of their involvement in safe, positive, and structured activities.

Common examples of YEPs are found in *particular* youth councils, teen centers, community-based participatory research programs, social action and advocacy groups, peer education models, and informal and non-formal education programs *that* regularly integrate youth participation in program decision-making, as stated above. As Wallerstein (2006) suggests, because by definition the participants of empowerment programs determine the direction and characteristics of the program, the aims of empowerment programs may change from one social or cultural context to the next. The primary durable characteristic of an empowerment program is the intervention's *process*, by which participants are regularly involved in decision-making or problem-solving in their immediate environment (Zimmerman, 2000).

Structurally, this participation within programs usually takes the form of advisory councils, committees, youth on boards, workgroups, or staff positions. Sometimes, young people and adults serve together in formal leadership capacities

such as committees; other times, membership is reserved exclusively for youths with adults acting in more of a supportive role. Youth empowerment involves a collective, democratic, and prosocial process of engagement, which implies group interaction (Cargo et al., 2003; Jennings, 2006). Consequently, exclusively one-to-one youth development interventions, such as most mentoring schemes, would not meet this dissertation's criteria.

Like YEPs, many non-empowerment-based out-of-school programs involve structured activities and safe spaces during hours in which adolescents are generally more susceptible to unproductive and delinquent use of time. They do not qualify as YEPs, however, if youths are not systematically involved with program decision-making. Some peer education models, for example, may only activate adolescents in content delivery rather than shaping program planning and implementation (Shiner, 1999). Often, youth center and after-school program schedules, objectives, and activities are adult-driven. Youths may be asked for their input or sporadically involved in programmatic decision-making, but if their involvement is not structured so as to ensure opportunity for real influence and regular participation in programmatic decision-making processes, the intervention is not empowerment-based.

Positive youth development. Much of the youth empowerment literature in North America has been linked to the positive youth development (PYD) movement (Holden et al., 2004). The general aims of PYD are encapsulated by the phrase coined by Pittman and colleagues and mentioned above, “problem free is not fully

prepared” (Pittman & Irby, 2000). In other words, PYD directs the strategy of interventions from problem-solving, prevention, and deterrence toward “youth preparation and development” (Roth & Brooks-Gunn, 2003a).

Exponents generally agree that PYD involves an asset-based approach to working with youth, which emphasizes the presence of positive environments, pro-social activities, and supportive relationships in young people’s lives. Positive youth development links the process of a youth’s involvement in the intervention with developmental outcomes (Roth & Brooks-Gunn, 2003b). While some definitions of PYD explicitly stress the participation of young people in organizational decision-making and program design as a critical feature in the PYD process (NCFY, 2007; Small & Memmo, 2004; USDHHS, 2001), others do not (Catalano et al., 2004; National Research Council, 2002; Roth, 2004). The PYD movement represents a broader trend toward strength-based approaches to youth and children of which YEPs represent one facet. In other words, while all YEPs are necessarily PYD, not all PYD programs are necessarily youth empowerment.

Empowerment-based programming. Youth empowerment represents a shift in youth services toward interventions that integrate young people as active partners in decision-making processes and power structures. The definition of youth empowerment is not entirely consensual. Interpretations of the term vary, for example, by the extent to which young people should be entitled to influence various levels of decision-making or the minimum supports and resources that characterize a genuine empowerment process.

Sometimes, the literature uses ‘empowerment’ to describe capacity-building with individuals or groups in order to prepare them for greater independence, but empowerment is treated as a set of competencies to be taught rather than practiced through the process of directly influencing the program itself. Additionally, other terms have been used to describe youth empowerment, including, *inter alia*, youth involvement, youth engagement, youth voice, youth participation, youth as resources, youth-adult partnerships, youth infusion, and, in some cases, Positive Youth Development or Community Youth Development (Gray & Hayes, 2008; Zeldin et al., 2000).

This review operates generally from the definition of youth empowerment used by the World Bank and UNICEF: "Empowerment is the expansion of assets and capabilities of young people to participate in, negotiate with, influence, control, and hold accountable institutions that affect their lives" (La Cava et al., 2004). The World Bank and UNICEF tie in self-power and independence as core features of empowerment and emphasize the role of youth as active participants rather than passive recipients in the processes that affect their lives.

A paper by the World Health Organization and United Nations described the “active involvement of young people” in planning, implementing, and evaluating as one of the most important principles of successful programming (WHO/UN, 1999). This involvement, the report suggests, ensures relevance of activities to young people’s needs, commitment to program objectives, and more positive contributions to the development of those young people actively involved. Similarly, quantitative

research on factors associated with young people's community participation is most significantly correlated with the extent to which they perceive the agenda as relevant or important to them (Altman, 1998). Youth empowerment programs aim to foster relevance by involving youth in shaping the program's activities and priorities.

Empowerment literature suggests that avoiding meaningful empowerment for young people limits the role of youths to passive recipients of services, at best, which perpetuates feelings of distrust, belittlement, and lack of control. Lofquist's 'Spectrum of Attitudes' describes a continuum of relationships with youths that includes attitudes of young people as Objects, Recipients, or Resources (roughly, things done 'to youth', 'for youth', and 'with youth', respectively) (Lofquist, 1989). The meaningful participation of young people positions them in the role of resources in organizational leadership.

It is the role of resource, empowerment advocates contend, that recognizes the agency of young people, which reinforces trust, ownership, independence, and self-efficacy. The Carnegie Trust UK further argues that the process helps young people develop social skills and capabilities including competencies in communications, negotiation, and teamwork (Carnegie, 2008). The perpetual, underlying constraint of young people in 'object' or 'recipient' capacities conceivably defies natural needs of adolescents to take risks, gain independence, and be challenged and supported in a critical period of development.

Terminology. Academics and practitioners use a number of terms to describe YEPs. Popular alternatives include variations of youth participation, youth-led, youth-driven, youth voice, youth engagement, and, in some cases, positive youth development. All of these terms, including youth empowerment, can be interpreted loosely if not carefully defined. Moreover, vocabulary invariably shifts across cultures and programming areas. The author of this dissertation elected to use *youth empowerment*, as *empowerment* indicates power transfer while leaving room for youth-adult partnerships, which can be highly participatory for adolescents.

Moreover, the term *youth empowerment*—at least in American literature—is arguably more closely associated with *process* constructs that better incorporate the supports, resources, and opportunities necessary to make participation meaningful rather than nominal (see, for example, Blanchard et al., 1996; Brown, 2009; Cargo et al., 2003; Chinman & Linnery, 1998). The author does not assert that youth empowerment should replace parallel vocabulary, and the review’s search terms described in chapter two were sensitive to the diverse intervention terms used to describe YEPs.

Empowerment to increase adolescent program participation

Importantly, YEPs may constitute a solution to one of the most fundamental challenges for interventions attempting to change adolescent outcomes: attracting them to attend. Afterschool programs are relatively commonplace in the USA with more than 6 million children enrolled. Problematically, however, only 8% of participants are adolescents in grades 9 to 12 (Hall & Gruber, 2007). Paradoxically,

however, an impact evaluation of fifteen afterschool programs in Maryland, USA ($n = 825$)—albeit using randomized controlled trials for only three sites versus nonrandomized controlled trials with the others—found that program participation was more effective in reducing delinquent behaviors for adolescents than for elementary school students (Gottfredson et al., 2004).

It remains unclear as to how much of the participation gap is attributable simply to a lack of age-appropriate afterschool programs for teens, but survey data from the national evaluation of the 21st Century Community Learning Centers programs suggests that boredom, disinterest, and lack of relevance are among the most prevalent themes that dissuade older youth from participating or lead them to stop coming shortly after trying it out (Durlak & DuPre, 2008). To this end, the process of empowerment may improve outcomes not only directly through the empowerment process, but also indirectly by increasing relevance, ownership, and, therefore, motivation for participation (higher exposure) in social programs. Indeed, a mixed-methods study of out-of-school time programs in six American cities involving program survey data (198 programs) and in-depth interview data (28 programs) found that providing many leadership opportunities to youth was identified as the program characteristic most associated with higher retention of older youth ($\beta = .25$, $p < .001$; Deschenes et al., 2010).

Yet, YEPs must not only overcome the challenge of engaging adolescents, but more specifically that of engaging groups of adolescents most in need of intervention. A study of YEPs in the USA involving youth-led tobacco prevention

advocacy, for example, showed participants ($n = 2,059$) to be majority White (61%), planning to attend college or already enrolled (91%), and reporting average or above average school performance (94%) (Holden et al., 2005). Because more advantaged and better performing groups of young people are likely to have experience with empowering opportunities in their homes, schools, and communities, they may be more inclined to seek out or have access to empowering programs. Adults involved in YEPs may more readily engage less vulnerable youth because of pre-existing associations (e.g., high-performing youths may be more likely than low-performing youths to have connections with adult community or school leaders helping to organize YEPs) or because adults perceive already high-performing youth leaders to be better leaders for the program or easier to work with.

Youths have a tendency to affiliate with peer groups with characteristics similar to their own (Cohen, 1977), and, therefore, more affluent youths who may have constituted early cohorts of YEPs may be less likely or less able to attract youths with different socioeconomic characteristics. Given that empowerment theory was largely engendered from intent to engage and advance marginalized populations, an important matter of evaluative inquiry for YEPs involves the extent to which they can attract and transform particularly vulnerable groups of adolescents.

Multilevel implications of youth empowerment

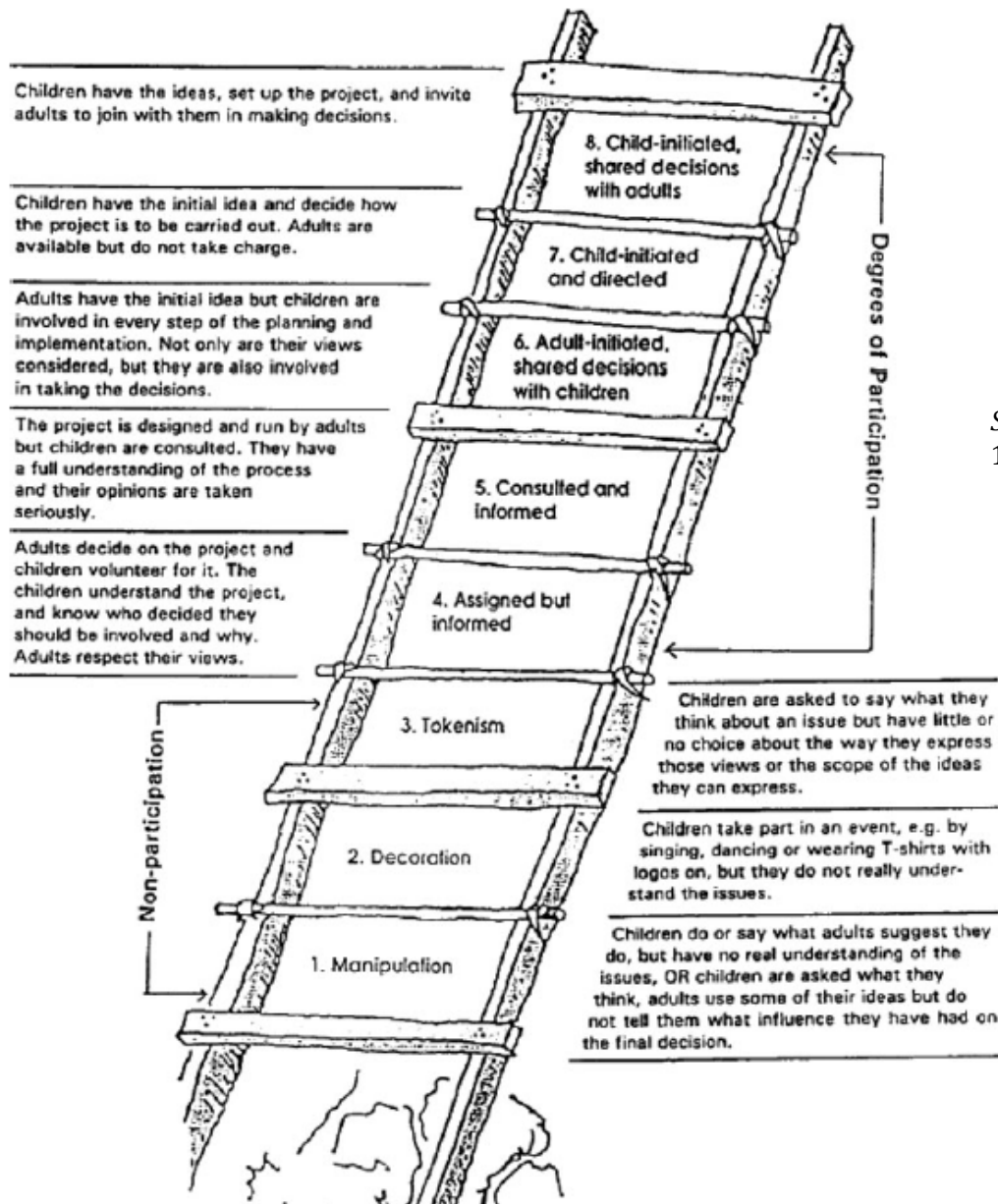
Although this review examines the possible impacts of YEPs on individual outcomes for youths, empowerment exponents draw attention to potential gains for a number of stakeholders. At a micro-level, when youths benefit, so too could their parents or guardians, as positive community participation and behavior replaces concern for their children's involvement in antisocial activities. At a meso-level, public services and programs for adolescents may stand a better chance of improving their 'products' and performance by better integrating the views of 'customers' (Carnegie, 2008). Educators could benefit from potentially disconnected students exhibiting better connections, behaviors, and motivation in school. Communities may benefit directly when youth empowerment efforts—as they often do—center on community service and advocacy initiatives (Jennings, 2006). On a macro-level, government and taxpayers could benefit from more skilled, competent, and compassionate young people entering adulthood putting less strain on public resources, contributing less to crime, engaging more with their communities, and being better equipped to compete in a global economy.

Levels of participation

A primary challenge for research on YEPs is to define what constitutes youth empowerment. The fact that empowerment can be viewed in different ways is in part a consequence of the nature of empowerment, which is a non-static process often characterized by different levels of participation at different levels of decision-making. Several typologies have been developed over approximately the last two

decades to try to create practical categories accounting for these variations within youth empowerment.

Figure 1.5. Hart's Ladder of Participation



Source: Hart, 1992

Lofquist's (1989) 'Spectrum of Attitudes' gave a basic typology of relationships with youths that classified attitudes toward young people as objects, recipients, or resources (roughly, things done 'to youth', 'for youth', and 'with youth',

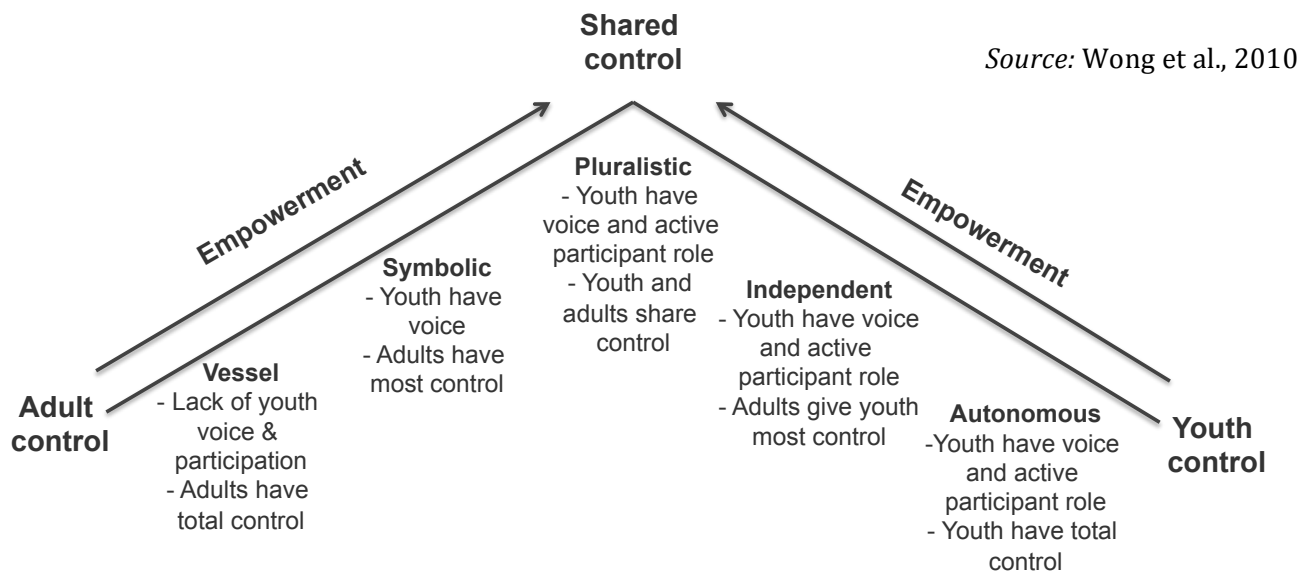
respectively). Hart's prominent 'ladder of participation' went further to delineate a continuum of eight levels at which young people can be engaged (see figure 1.5; Hart, 1992).

Hart's ladder made a particularly important contribution by illustrating what kinds of activities do *not* qualify as participation as well as those that do. Despite the ubiquitous references to the 'ladder of participation', however, Hart himself later recognized needs for updating the framework based on more current knowledge about youth development and cautioned readers against applying the framework too strictly (Hart, 2008). Hart suggested that the ladder was never intended as a "comprehensive tool for...measuring work with children," but rather as a "jumping-off point" for critical reflection. Interpretations of the ladder that Hart tried to counter included expectations that youth must always perform at the top of the ladder for full empowerment to exist as well as dismissal of the role of adults in power-sharing and helping youths to develop the competence and confidence to participate effectively.

Other authors subsequently developed frameworks as attempts to build or improve on Hart's ladder, including Treseder's degrees of participation (Treseder, 1997), Shier's pathways to participation (Shier, 2001), and, most recently, Wong and colleagues' Typology of Youth Participation and Empowerment (TYPE) Pyramid (see figure 1.6; Wong et al., 2010). Wong and colleagues proffered a typology of youth participation that values the role of adults in the empowerment process more explicitly by placing youth-adult shared control as the peak of youth empowerment.

“In co-learning with youth,” Wong and colleagues wrote, “adults can serve as resources and collaborators—versus being the experts—by facilitating critical dialogue, awareness, and building skills toward critical consciousness in partnership with young people” (2010, p. 105).

Figure 1.6. TYPE Pyramid



The TYPE Pyramid most clearly depicts the distinction between youth empowerment and youth anarchy implied by much of the youth empowerment literature. For some, emphases among YEPs on equipping young people with increased influence and control in decision-making processes might conjure unsettling images like the fictional novel by Golding (1954), *Lord of the Flies*—unsupervised youths left to govern themselves only to exploit an abrupt grant of autonomy to wild and destructive effect. The concern echoes Plato’s warning about democracy without order that “excess of liberty, whether in States or individuals,

seems only to pass into excess slavery,” which no doubt inspires all the more trepidation when those individuals happen to be pubescent (Plato, 2009, p. 250).

Yet, an anarchical interpretation differs from the most prominent youth empowerment literature—particularly in the context of intervention—which stresses a central and vital role for adults (Hart, 2008; Jennings, 2006). In fact, YEPs may in principle require an even more active adult role than youth programs in which the primary adult function is limited to activity implementation and supervision rather than development and support of youths’ skills and contributions. Youth empowerment programs do, however, change the *nature* of the relationships between youths and adults so that they become more horizontal with adults acting as facilitators and partners with youths, and with both youths and adults respecting the unique contributions that each other brings to the partnership (Wong et al., 2010).

The review reported in chapter two aims to capture evaluations of programs that fall within the top three rungs of participation on Hart’s ladder, which integrate adolescents into program decision-making. Consistent with the Wong and colleagues’ typology and much of the recent youth empowerment literature on youth-adult partnerships (Camino, 2000; Evans et al., 2004; Jones & Perkins, 2005b; Zeldin et al., 2008), this dissertation’s systematic review also incorporates the role of adults in the intervention inclusion criteria.

As the various typologies illustrate, youth can be engaged at different levels of participation and shared control with adults along the spectrum of youth

empowerment. The participatory criteria held by this dissertation's review—regular involvement of youth in program decision-making—sets a basic standard for youth empowerment programs so as to be inclusive of the range of programs based on empowerment models, but evaluation could show that different levels and characteristics of youth participation and adult involvement facilitate different program effects in general or for particular subgroups of young people.

Examples of YEPs

Because youth empowerment can take many forms and settings, examples of YEPs can help illustrate what the intervention looks like in practice. This section includes two cases of actual YEPs with intervention descriptions extracted and paraphrased from existing studies or reports. The YEPs exhibit some clear differences. The two examples are set in different geographical and cultural contexts (Cape Town, South Africa and California, USA). One program, YES!, is more structured and curriculum-based, whereas the other, Children's Clubs, tends to integrate many different kinds of programs and is less preplanned in terms of its process. Consistent with empowerment theory, the forms and content of activities vary between the two programs—for example, photographic research and reflection activities underlie one program whereas cultural activities are frequent in the other.

Despite differences, all of the defining elements of YEPs are identifiable from both program descriptions. Both, for instance, regularly engage youth in program decision-making processes as a characteristic of the intervention, and both involve trained adults as facilitators and supporters in the empowerment process. Notably,

the primary target groups in both examples are pre- and early-adolescents, but many other YEPs target mid- and older-adolescents. These example programs did involve older youths, but the particular intervention designs involved them in such a way that they were largely treated distinctly from younger youths and trained to contribute to facilitating participation of younger youths.

Example I. Youth Empowerment Strategies (YES!) Project – California, USA

The YES! project is an afterschool program and participatory research initiative for urban, disadvantaged early adolescents, aged 9 to 12. The program aims to reduce unhealthy risk behaviors such as substance use and bullying through an empowerment approach that engages youth as problem-solvers and enhances their skills and self-efficacy to become agents of change. YES! organizes youth groups in which participants learn and practice cognitive and social skills and participate in democratic decision-making processes to plan and implement group-designed social action projects in their schools and communities.

Participants meet for about 90 minutes weekly for approximately 25 sessions during the school year. Each group has a graduate student facilitator and a young person selected from a local high school serving as a co-facilitator. Both facilitators receive 30 to 60 hours of training in participatory education, facilitation, youth development, and social action from the program. The groups meet in school venues with room locations and times set by the school.

The program enacts youth empowerment within a fairly structured (albeit flexible in terms of its implementation) curriculum that helps participants move through progressive stages of preparation and development. Core activities include group formation activities (establishing group norms and teambuilding); training in photography; 'photovoice' activities in which participants take, analyze, and reflect on photos of school or community life; identification and planning of social projects; and recruitment of allies and implementation of social action projects. Social action projects covered a wide range of themes, such as awareness about gang pressure to use drugs, reducing litter in classrooms, improving the condition of school bathrooms, and awareness about alternatives to fighting.

YEP elements

Participation in decision-making – Youths participate in decision-making about group expectations, discussion topics, their own photography, types of projects pursued, and planning project activities. Older youths also facilitate activities with younger youth.

Supportive adults – Each group has a graduate student facilitator and older youth facilitator intended to support the empowerment process and the youths themselves.

Asset-building activities – Youths participate in teambuilding and skill-building training and activities around photographic expression and social action. Youths also participate in and organize projects that involve activities such as skits, campaigns, and coalition building, which involve process opportunities for capacity building.

Prosocial environment – Groups involve collective decision-making and participation, group rules and values, and adults that help inculcate positive group interactions.

Information source: Wilson et al. (2007)

Example II. Children's Resource Centre's Children's Groups – Cape Town, South Africa

With rooting in the anti-apartheid movement, the Children's Resource Centre (CRC) aims to help groups of children and youth in poor neighborhoods organize themselves to address social problems and increase youth participation. There are approximately fifty active Children's Groups in impoverished areas throughout South Africa (though largely in Greater Cape Town where the program began). Collectively, the groups are called 'the Children's Movement.' A youth General Council of two elected delegates per group represents the groups in guiding the overall 'Children's Movement.'

Each group has a name, a youth executive committee, an adult coordinator, and a venue, typically a school or community center. The target age group is 7 to 14 years old, but older youth are also involved through various leadership roles, including 'young adult facilitators' who help the adult coordinator facilitate the participatory process and activities for younger youth. Youth members nominate adult coordinators for each group, and the CRC provides them and young adult facilitators with training and support in implementing and facilitating participatory activities. Participants must agree to abide by rules and expectations established by existing youth members.

The Children's Groups include a wide range of activities associated with different programs, including health awareness campaigns, education, environment, and cultural or arts activities. The CRC's adult staff initiates many of these programs, and youths initiate others. Either way, the activities are intended to provide youth with opportunities to participate in decisions and leadership regarding the planning and implementation of the activities. In many groups, members meet at least two days per week, often in subgroup teams, for around two hours per session. Adults include both dialogue and activity components in the sessions guided by youths' ideas and preferences.

YEP elements

Participation in decision-making – Youths participate in decision-making about types of activities pursued, planning activities, and how they contribute. They establish group rules and help select adult staff. Older youths also facilitate activities with younger youths.

Supportive adults – Each group has an adult coordinator and young adult facilitators intended to support the empowerment process and the youths themselves.

Asset-building activities – Groups implement and participate in a range of developmental activities including arts, advocacy, literacy, and recreation, all of which may offer opportunities for learning and capacity building. Some youths also participate in structured training programs organized by CRC or partner institutions on youth participation or facilitation generally or on specific program areas, such as environmental action.

Prosocial environment – Groups involve collective decision-making and participation, group rules and values, and adults that help inculcate positive group interactions.

Information source: RIATT (2010)

Theory of change

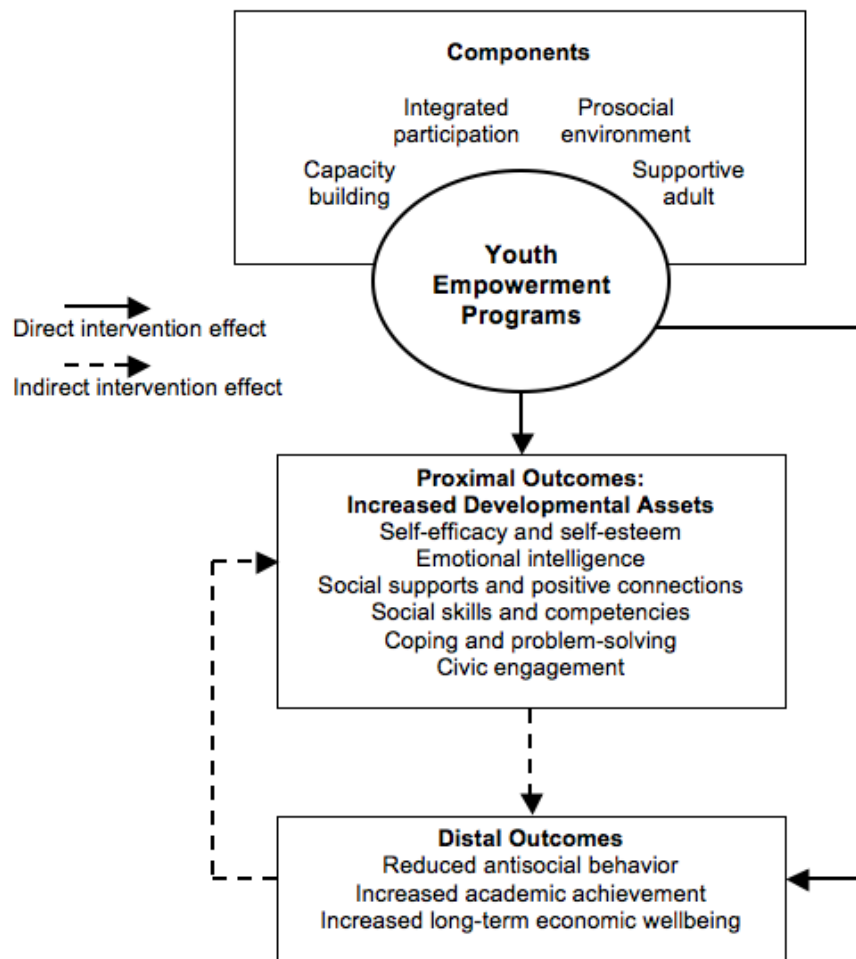
A theory of change articulates how an intervention's processes and components are expected to influence intended outcomes. In other words, it describes how the intervention might work. Evaluation literature expects a theory-driven approach to improve the design and quality of interventions and the evaluations that assess them (Imas & Rist, 2009; Weiss, 1995). As such, this section focuses on the theoretical basis for YEPs' hypothesized change process.

Youth empowerment programs aim to develop psychosocial assets among participating youths through a dynamic process that integrates connections with supportive adults, capacity-building (or 'asset-building') opportunities, prosocial environments, and regular involvement in program decision-making. In turn, youth development literature expects that these assets serve as pathways to distal indicators of success and wellbeing (e.g., academic achievement and economic wellbeing) and as protective factors against problems like antisocial behavior. A basic theory of change diagram for YEPs (created by this author) is provided in figure 1.7.

Positive youth development frameworks for programming goals, such as the *Five C's* (Roth & Brooks-Gunn, 2003a) and the *Forty Developmental Assets* (Search Institute, 2008) have helped set the stage for an attention to strength-based outcomes in YEPs' theory of change. The Search Institute defines *developmental assets* as relationships, opportunities and personal qualities that young people need to avoid risks and to thrive. The Organisation for Economic Co-operation and

Development's (OECD's) establishment of *Key Competencies* reinforced concern for developmental assets (OECD, 2005). While the first OECD competency category is largely technology and knowledge-based, the remaining two categories—'interacting in heterogeneous groups' and 'acting autonomously'—outline critical competencies included in, or closely related to, the developmental outcomes described in this review.

Figure 1.7. Basic theory of change diagram for YEPs



Although researchers have found cognitive skills (e.g., IQ) to be fairly solidified by age 8 to 9, they have found noncognitive skills (e.g., motivational,

emotional, and social skills) to be malleable into adolescence and thus ideal targets for intervention at that stage and likely to yield better return on investment than cognitive remediation strategies (Carneiro et al., 2007; Cunha & Heckman, 2006; Cunha et al., 2010; Heckman et al., 2006). Noncognitive abilities, moreover, have been shown to predict important distal outcomes, such as future educational-level attainment, employment, wages, and adult depression, even after cognitive ability and demographic variables are controlled for (Carneiro et al., 2007; Cunha & Heckman, 2006).

The development of a theory of change for a specific intervention integrates multiple forms of contributing knowledge, including relevant theory, evidence from previous studies, contextual and stakeholder analysis. Later chapters will identify contextual and stakeholder issues that influence the theory of change underlying a particular youth empowerment program (YEP). This chapter considers the theory of change for YEPs as a general intervention, and so it primarily focuses on psychosocial theories and non-evaluative evidence that inform the components and anticipated youth-level outcomes connected to youth empowerment. Chapter two will systematically review evidence from experimental and quasi-experimental evidence that could lend to further development of a theory of change.

Theories behind the theory of change

Notably, the use of the terms ‘psychosocial theories’ and ‘theories of change’ can be confusing because the two are not synonymous. A theory of change is a specification of how a particular intervention—its components and activities—is

expected to achieve specified outcomes (Connell & Kubisch, 1998). Psychosocial theories are not constructed for a particular intervention but rather offer more general explanations for how attitudes and behaviors are affected. As such, a psychosocial theory is not a theory of change in the intervention sense of the term, but it can contribute to formulating one.

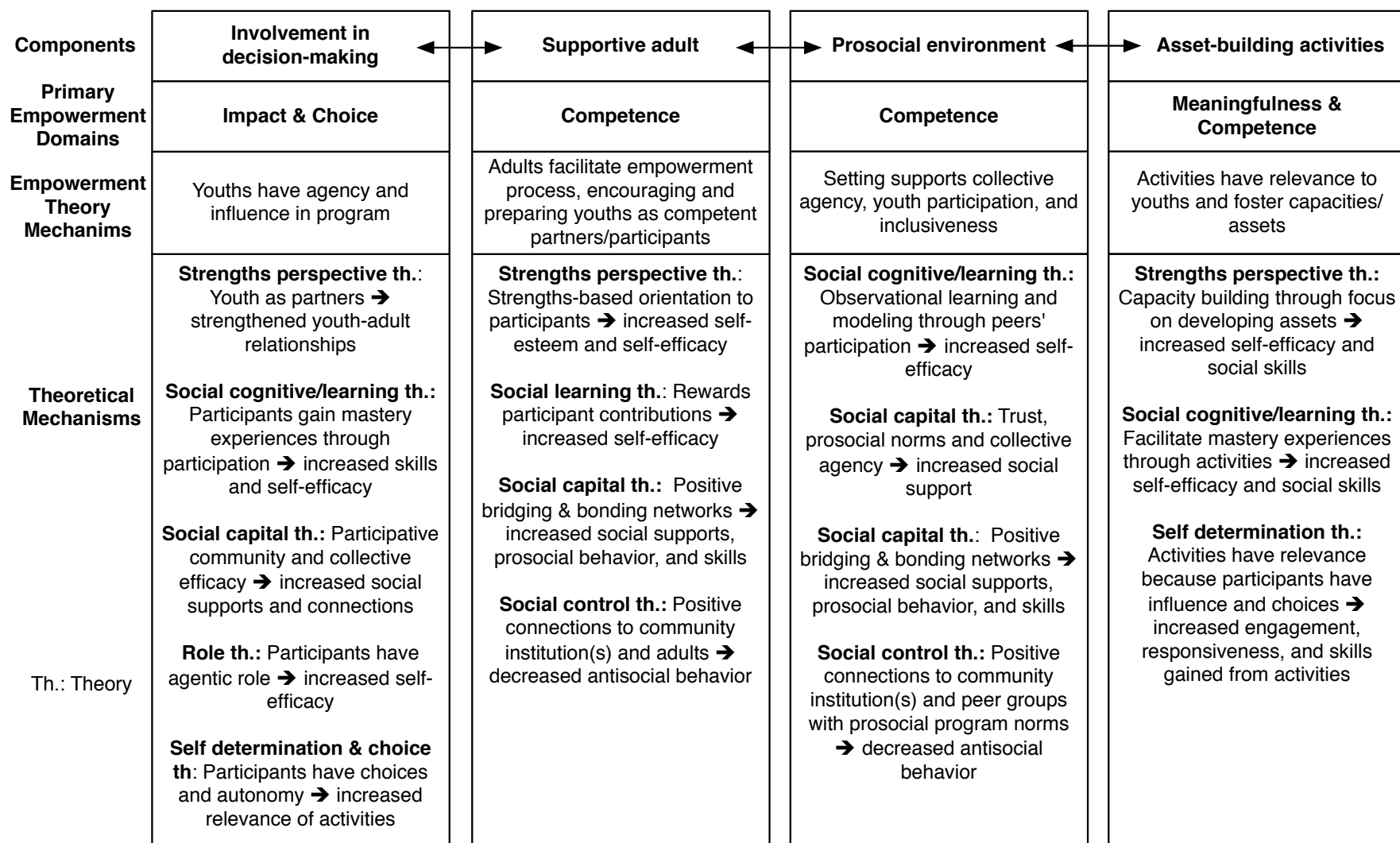
The extant youth empowerment literature draws on a range of psychosocial theories to elucidate the paths by which YEPs are expected to change young people's attitudes and behaviors. Most importantly and explicitly, empowerment theory provides a basis for conceptualizing the processes and activities by which YEPs may affect change in youth outcomes (Holden et al., 2004). It would be inadequate, however, to examine empowerment theory without also considering how other psychosocial theories deepen understanding of how particular functions in the empowerment process relate to developmental outcomes. Hence, the youth empowerment literature draws on more than empowerment theory alone to articulate the expected change process. Key contributions from empowerment theory and related theories are summarized in the following paragraphs and illustrated in figure 1.8.

Empowerment theory. Empowerment theory heavily influences social and economic interventions, like YEPs, that are based on participatory approaches to improving the welfare of marginalized populations (Holden et al., 2004). Yet, despite empowerment theory's importance in community psychology, the theory remains largely enigmatic due to the diverse ways in which empowerment and its

implications are conceptualized (Zimmerman, 2000). This section aims to clarify key notions of empowerment theory, particularly as it relates to youth programming. As Zimmerman explains, “Empowerment is both a value orientation for working in the community and a theoretical model for understanding the process and consequences of efforts to exert control and influence over decisions that affect one’s life, organizational functioning, and the quality of community life” (Zimmerman, 2000, p. 43). This chapter examines empowerment’s contributions as a theoretical model to YEPs’ theory of change.

Empowerment *theory* “connects individual wellbeing with the larger social and political environment, and suggests that people need opportunities to become active in... decision-making in order to improve their lives, organizations, and communities” (Zimmerman, 2000, p. 58). Empowerment theory, historically more centered on marginalized adult populations, has promoted an emphasis on people’s strengths, appreciation for cultural diversity, and shift of language and services to supplant “one up/one down helper-helpee relationships” with collaboration and active participation of disempowered persons in the processes that affect their lives (Chinman & Linnery, 1998; Rappaport, 1981). It is this function of agency through participant engagement in decision-making processes that underpins empowerment theory and largely distinguishes YEPs from more broad-based youth development programs.

Figure 1.8. Theoretical mapping behind youth empowerment programming's theory of change



This diagram provides a suggested framework for how relevant social theory contributes to youth empowerment's theory of change in youth development. Empowerment domains are consistent with those posited by Thomas and Velthouse (1990). Empowerment theory provides the overarching theoretical basis with other social theories (described in more depth in this section) helping to explain how mechanisms within the youth empowerment process might contribute to the change process. Arrows between components indicate that each component in youth empowerment is expected to interact with and reinforce the others.

According to Fymier and colleagues (1996), attention to the philosophy and practice of empowerment was popularized through writings by Block (1987) in the business context of manager-employee interactions. Block promoted an organizational environment in which each employee felt s/he was invited to make meaningful contributions and shape the direction of the company. These ideas were subsequently applied to other organizational contexts, such as schools and classrooms (Glasser, 1990; Shulman & Luechauer, 1993). While definitions of empowerment in organizational contexts evolved and varied, the emphasis on overcoming individuals' sense of powerlessness within systems and increasing intrinsic motivation through fostered feelings of ownership and self-efficacy remained constant (Fymier et al., 1996).

As empowerment-oriented approaches were increasingly applied toward poverty alleviation (e.g., Easterly, 2006; Godinot & Quentin, 2006; Nylen, 2003), Sen (1999) argued that the path to ameliorating the conditions of the world's poor lies in the recognition and development of the poor's *capabilities*. This "freedom-centered" understanding of economics and development, Sen contended, is an "agent-oriented view" moving away from approaching individuals primarily as "passive recipients of the benefits of cunning development programs" (Sen, 1999, p. 11). As such, Sen retains the empowerment theory focus on agency and participation in policies and programs, but he contributes to empowerment theory the notion that the poor are poor because they lack defined capabilities, and that agency and participation can foster development of these capabilities.

Although empowerment theorists often describe how the experience of agency and participation can lead to improved outcomes, they do not often specify what the experience of empowerment comprises. Thomas and Velthouse (1990) defined the experience of empowerment to consist of four dimensions. These dimensions include *meaningfulness* (“the value of a task in relation to one’s own beliefs, ideals and standards”), *competence* (“the person feels qualified and capable to perform the necessary activities to achieve the goal”), *impact* (“the accomplishment of a task is perceived to make a difference”), and *choice* (“the degree to which persons self-determine their task goals or methods for accomplishing them”) (Fymier et al., 1996). These categorizations suggest that participation in decision-making is important, but if, for example, participants do not feel the tasks or subjects of decision-making are not important, they will not feel invested, and if participants do not feel competent to engage in decision-making processes effectively, they will not feel much sense of agency even when the opportunities to contribute to decision-making are present. These dimensions further contextualize the role of supportive adults in YEPs, who can help develop the capacities of young people to participate effectively and facilitate participatory processes in which youths feel they have choice, impact, and meaningful connections.

Empowerment theory has been increasingly applied toward adolescent interventions (Chinman & Linnery, 1998; Lakin & Mahoney, 2006). Mohajer and Earnest (2009) and Wong and colleagues (2010) connect youth empowerment to Freire’s ‘pedagogy’ (1972) by which marginalized populations develop *critical*

consciousness—progressive awareness of one’s environment and one’s ability to affect change within it—through participatory learning and action. Freire’s work has been broadly influential in the empowerment literature. Importantly, Freire helped to articulate empowerment as a process, and one that required help rather than one that simply involved leaving the poor to lead themselves. He believed that it is the job of a ‘facilitator of empowerment’ to help navigate marginalized people through this process using shared dialogue as a key instrument, in which the facilitator sees him/herself as a co-learner with the participant. In the case of YEPs, the facilitator role explains the prominence of supportive adults as a key component in the empowerment process. In fact, several of the YEPs discussed in the next chapter’s systematic review describe their adult staff and volunteers as facilitators.

To summarize, the key elements of empowerment theory are that: (1) the agency of marginalized groups and individuals in decision-making processes that affect them leads to their improved wellbeing (Sen, 1999; Zimmerman, 2000); (2) intermediately, empowerment enhances the capabilities, intrinsic motivation, sense of ownership, and self-efficacy of participants, which mediate empowerment’s impact on wellbeing (Fymier et al., 1996; Sen, 1999); (3) the experience of empowerment in tasks and activities can be categorized multi-dimensionally in terms of participants’ perception of choice, impact, meaningfulness, and competence (Thomas & Velthouse, 1990); and (4) the process of empowerment, particularly with marginalized groups, is enabled by facilitators who help ensure that the dimensions of the empowerment experience are realized (Freire, 1972; Wong et al.,

2010). The following paragraphs examine other psychosocial theories that further develop and explain change processes held by empowerment theory.

Strengths perspective theory. Empowerment theory is closely linked with the strengths perspective theory of social work practice (Healy, 2005, p. 154). Strengths perspective theory particularly mirrors empowerment theory's emphasis on agency and capability-development to improve outcomes. As Holden and colleagues (2005) express, "Empowerment theory replaces the view of youth as community problems, which dominates prevailing intervention theories, with a view of youth as community assets and resources." This reorientation is clearly repeated in strengths perspective theory.

The strengths perspective theory postulates that if social workers "mobilize clients' strengths (talent, knowledge, capacities) in the service of achieving their goals and visions, the clients will have a better quality of life on their terms" (Saleebey, 1996). Similarly to empowerment theory, the strengths perspective emerged as a challenge and contrast to what advocates saw as the prevailing preoccupations in social work practice on individual pathologies and deficits. Foundational exponents of the strengths perspective critiqued latent features of modern social work professionalism that emphasized detachment, individualized treatment, social labeling, and practice language permeated by deficits rather than strengths orientation, all of which they saw as doing more harm than good for clients (Goffman, 1991; Reynolds, 1951).

The strengths perspective ventures to counter these trends by shifting social work practice toward the following principles: (1) adopting an optimistic attitude toward service users; (2) focusing primarily on assets of service users, on which services can build; (3) collaborating with service users, emphasizing service users as *partners* in service processes with workers; (4) working toward long-term empowerment of service users, supporting participants in developing durable capacities through adversities; and (5) creating communities of strengths-focused social support to help service users achieve resilience and enhance quality of life (Healy, 2005). Evidence of the influence of the strengths perspective theory can be found in the codes of ethics for both the British Association of Social Workers and the American National Association of Social Workers, both of which cite partnership, capacity building, and development of potential with service users as leading principles for the practice (BASW, 2011; NASW, 2011).

With respect to youth, the strengths perspective is strongly reflected in the 'positive youth development' movement (which emphasizes focus on assets and strengths of young people over perceived deficits and problems) within the social work community in the USA and increasingly abroad (Park, 2004). The emphasis, however, on collaboration with service users as partners in practice connects the theoretical orientation more specifically to youth empowerment. Again, adults are given an important role in the theoretical framework, which recognizes a partnership between practitioners (adult staff) and service users (youth participants). This partnership is consistent with the role of adults as facilitators

and reflected in Wong and colleagues' (Wong et al., 2010) typology of youth participation and empowerment (TYPE).

Social learning theory, social cognitive theory, and social control theory.

Empowerment involves highly social processes between marginalized individuals and their environment. As such, empowerment theory assumes that certain kinds of social interactions affect positive changes in program participants, but the theoretical mechanics of how empowerment-based social interactions influence attitudes and behaviors are not well-developed by empowerment theory alone. For this, we have to examine empowerment theory's rootedness in several social theories that articulate how attitudes and behavior can be changed by altering the nature of a youth's interactions (Bandura, 1986; Kim et al., 1998).

Social learning theory posits that young people's behavior responds to experiences of rewards and sanctions in their social interactions (Akers et al., 1979). This is particularly important in the development of self-efficacy, because negative social messages could diminish intrinsic motivation attached to one's precepts of ability to achieve tasks, and, conversely, social learning theory suggests that positive reinforcement and experience with achievement can increase self-efficacy beliefs. Building on social learning theory, *social cognitive theory* assumes that a particular attitudinal outcome, self-efficacy beliefs, can be intentionally changed through an individual's social interactions.

The proposed sources of self-efficacy development include mastery experience, vicarious experience, social persuasions, and physiological reactions

(Pajares, 2006). Pajares (2006) suggests that, for most people, mastery experience constitutes the most important influent of self-efficacy—actual experience with achievement. While YEPs may have little role to play with physiological influents of self-efficacy, they do purport to foster processes and environments through which young people have the opportunity to succeed in new challenges (mastery experience), interact with peers, encouraged through the same process to engage in positive and challenging goals and activities (vicarious experience), and receive constant positive feedback that affirms young people's immediate value and lauds their achievements and contributions (social persuasions).

Not only does social cognitive theory posit that certain kinds of social interactions can change self-efficacy beliefs, it also expects that those beliefs go on to affect behavior. As such, social cognitive theory offers theoretical justification for why changing self-efficacy matters. Social cognitive theory postulates that development of self-efficacy translates to stronger performance in cognitive development and functioning (Bandura, 1993). Consistent with social cognitive tenets, emphasis on the importance of self-efficacy assumes that young people's actions follow from a process of self-reflection that interacts with their personal and social experiences. In other words, processing of their beliefs in the level of goals worth pursuing and their ability to achieve them allegedly often precedes young people's actions. As such, if youth empowerment strengthens young people's sense of self-efficacy, this change mediates behaviors attached to that self-efficacy, such as overcoming problems, participating more assertively and effectively, and pursuing task or life goals. Bandura describes the chain of thought to action in this way:

Much human behavior, which is purposive, is regulated by forethought embodying cognized goals. Personal goal setting is influenced by self-appraisal of capabilities. The stronger the perceived self-efficacy, the higher the goal challenges people set for themselves and the firmer is their commitment to them (Bandura, 1993, p. 118).

Like social learning and cognitive theories, *social control theory* emphasizes the role of social interactions in shaping behavior. Social control theory, however, focuses more specifically on social inclusion and exclusion as influents into youth behavior. Young people commit deviant acts, Hirschi (1969, p. 16) suggests, when their bonds to society are weak or broken. Huebner and Betts (2002) oriented this theory toward positive outcomes as well, better aligning it within the strengths-based youth development and empowerment paradigm. The 'bonds' to which Hirschi refers include attachment to parents, schools, and peers; commitment to conventional lines of action; involvement in conventional activities; and belief in a common value system (Huebner & Betts, 2002). Notably, how different types of bonds influence behavioral outcomes may vary by gender. Survey data from 911 students from one American public school, for example, indicates that attachment bonds had more explanatory power toward delinquent acts and academic achievement with females, whereas involvement bonds appeared to have had better explanatory power with males (Huebner & Betts, 2002).

Consistent with social control theory, YEPs prioritize the social inclusion of young people. By engaging adolescents in shared decision-making and leadership processes within positive institutions, YEPs aim to construct a web of attachment and involvement bonds between youths and positive individuals and institutions in their schools and communities. Youth empowerment programs may be especially

well positioned to establish these bonds. Programs that serve young people but do not enable them to influence the activities and plans of the program may fail to foster the level of perceived ownership and association necessary to form sufficiently strong bonds to build adolescents' developmental assets and protect them against competing pressures toward deviant behavior.

Again, the social theories described in this section illustrate important roles for adults in the youth empowerment theory of change process. With or without special interventions, young people are naturally exposed to countless social interactions every day. The change processes articulated by the aforementioned theories assume that specific kinds of social interactions—e.g., new opportunities to succeed in positive and developmentally appropriate challenges, positive reinforcement, and supportive attachments to individuals and institutions—influence positive outcomes. Because the theories assume that these kinds of interactions are not taking place for many struggling young people at an adequate level without intervention, staff or volunteers that are motivated and equipped to facilitate these kinds of social interactions presumably constitute an essential part of the process. Consequently, the literature in these theoretical realms has focused heavily on the roles of adults in young people's immediate ecological settings to change self-efficacy and other developmental outcomes (Pajares & Urdan, 2006).

Social capital theory. An emphasis on 'bondedness' and 'connectedness' in youth empowerment literature reflects scholarship concerning *social capital theory*. Putnam (1995) describes social capital as the "features of social life—networks,

norms, and trust—that enable participants to act together more effectively to pursue shared objectives.” Contrary to traditional conceptions of empowerment as strictly de-contextualized and individualistic (e.g., Riger, 1993), more recent youth empowerment literature is community-oriented, focused on dynamic, prosocial environments (and, more specifically, young people’s participation and influence in such environments) (Jennings, 2006). To this end, social capital theory can help explain the potential value of social networks embedded in youth empowerment’s theory of change.

Social capital theory is multifaceted and even contradictorily defined (Halpern, 2005). Nonetheless, branches of social capital theory clearly tie the concept to key elements of empowerment theory. Namely, Onyx and Bullen’s (2000) hierarchical factor analysis of survey data of 1,200 adults in five Australian communities resulted in three specific factors that best empirically define social capital: community participation, agency, and trust. Likewise, youth empowerment’s most defining component—active involvement of youth in decision-making processes within community institutions—depends on each of these elements. Trust, being perhaps the most signature element of social capital according to other empirical research (Halpern, 2005, p. 33), is essential in the participation of youth in program planning or implementation. This requires a partnership between youths and adults based on trust between the two that is not necessarily present in interventions in which young people receive services without having the power to influence them.

Social capital theorists such as Putnam (2000) and Gittell and Vidal (1998) describe social capital in terms of what they call *bonding* and *bridging*. Bonding social capital emphasizes stronger grouping and relationship building among relatively homogenous people—such as youths of a same class or neighborhood. Bonding social capital draws on psychological literature concerning social support (Halpern, 2005, p. 35). Such bonding is expected to form a durable support network around individuals that can help provide safety and resilience in times of crisis or need. Bridging networks are believed to strengthen individuals' ability to advance in social systems, because connections to heterogeneous groups, particularly those with greater power or resources, can facilitate resources, opportunities, and ideas above and beyond what one homogenous group can provide. Granovetter (1983) popularly described this phenomenon as “the strength of weak ties.” Reflecting both aspects of social capital (as well as elements of social control theory), youth empowerment emphasizes the facilitation of *bonding* ties based on trust that can strengthen young people's resilience and emotional support as well as their *bridging* ties by connecting youths to new groups and institutions through empowerment that can promote youths' socio-economic mobility and healthier communities (Boeck, 2009; Chinman & Linnery, 1998).

Role theory. Though empowerment theory literature does not appear to draw explicitly on the sociological role theory, there are notable parallels and explanatory advantages in the sense that, by definition, empowerment changes the nature of individuals' roles in social settings. Freire's (1972) participatory pedagogy, for example, was unequivocal about redefining the roles of people in power (often adult

service providers in the case of YEPs) to those of facilitators, co-learners, and partners, and the roles of marginalized people (vulnerable youths in the case of YEPs) to that of contributors, teachers, agents, and partners. The prevalence of role and identity formation permeates youth empowerment literature (Chinman & Linnery, 1998; Larson, 2000).

According to role theory, attitudes and behaviors correspond with the expectations (the 'roles') that individuals' social environments implicitly assign to them (Biddle, 1986). As such, youth empowerment involves a process by which a young person's social environment intentionally redefines his or her role as one of value, ability, autonomy, and contribution, which is expected to change the youth's attitudes and behaviors to reflect that role. This expands on arguments by social cognitive and learning theorists that mastery experiences and social persuasions are important for the development of self-efficacy. Role theory contributes the notion that a new role within a social setting acts as a mechanism for change. As such, the point of a YEP is not only to provide opportunities to experience achievements and be rewarded for them, but it is moreover to help youths see these experiences as characteristics of a new role—that of an agent and capable contributor—that becomes a significant part of their evolving identity.

Self-determination theory and choice theory. Youth empowerment is consistent with key elements of choice theory and social determination theory to the extent that empowerment involves young people in decision-making processes that allow them to choose or help choose how they would like a program to run and

the ways in which they participate in the program. A key feature of empowerment theory is individuals' experiences of choice and impact in the processes that affect them. Self-determination and choice theories lend further explanatory aid to potential contribution of choice and impact in the change process.

Choice theory postulates that choice increases relevance of tasks and activities to individuals' needs and preferences and the extent to which individuals feel invested in those tasks and activities (Glasser, 1999). Moreover, choice theory expects that participants are more impacted by activities in which they perceive higher relevance and investment. Similarly, self-determination theory postulates that autonomy, competence, and relatedness are fundamental cornerstones underling people's intrinsic motivation (Dickerson, 1998; Ryan & Deci, 2000). Social determination theory and choice theory are discussed concomitantly here because both emphasize the role of individuals determining the nature of their involvements in fostering greater motivation to actively engage with and benefit from those involvements.

This principle of motivation's link to determination was exhibited in the results of an educational randomized controlled trial (RCT) ($n = 207$) that found that students in classrooms in which they could choose homework assignments showed higher intrinsic motivation, homework completion, and more positive perception of their relationship with the teacher than students in classrooms randomized to not offering choice in homework assignments (Patall et al., 2010). Similarly, youth empowerment advocates contend that programs that give youths agency in

decision-making increase the programs' relevance to youths' interests and objectives, which subsequently increases their engagement in and impact from the program.

In summary, a good theory of change is aided by examination and integration of existing psychosocial theories that lend logic and plausibility to its formulation, and youth empowerment's theory of change has much to draw on to this end. Empowerment theory provides an overall conceptual framework for how YEPs might positively change young people's developmental outcomes, but additional psychosocial theories help to expose in greater detail how mechanisms within empowerment theory contribute to the overall hypothesized change process. Figure 1.8 illustrates how the abovementioned psychosocial theories could conceptually map onto empowerment theory in explaining youth empowerment's general theory of change as it is expressed in this chapter.

Nevertheless, theories need to be tested, particularly in terms of how they materialize (or not) in the context of particular theories of change. Moreover, theories are by themselves generic and abstract; put into practice, ample opportunities for conflicting influences and implementation defects exist in the 'real world,' so intervention research must investigate the extent to which theories are reflected not only in the outcomes of an intervention, but also in its implementation.

Intervention costs

Youth interventions must increasingly demonstrate their competitive edge in contributing positively to communities and countries as they vie for space in the marketplace of social ideas during fiscally ominous times. As this dissertation is written, both the United States and United Kingdom governments, a couple of the biggest funders not only of domestic programs but also of international aid programs, propose deep cuts into social spending in desperate attempts to curtail significant national debts (HM Treasury, 2010; Office of Management and Budget, 2011).

Paradoxically, divergent demographic trends between developed and developing countries may place comparably heavy burdens on paying for youth services. Some argue that the ageing demographic trends among industrialized countries have shifted resources toward the elderly and away from youth and children (Preston, 1984; Weil, 1997). Meanwhile, in developing countries where youth numbers dominate the demographic composition of their societies, such as Egypt and Jordan, less wealthy governments struggle to finance the provisions necessary to help the large proportion of youths transition successfully to adulthood (Assaad & Barsoum, 2009; Kanaan & Hanania, 2009). In other words, for varying reasons, governments and donors are acutely challenged to match limited resources with the scale of need for improving youth outcomes. As such, YEPs confront significant urgency to design, implement, and evaluate their efforts with meaningful and measurable impacts relative to their costs in mind.

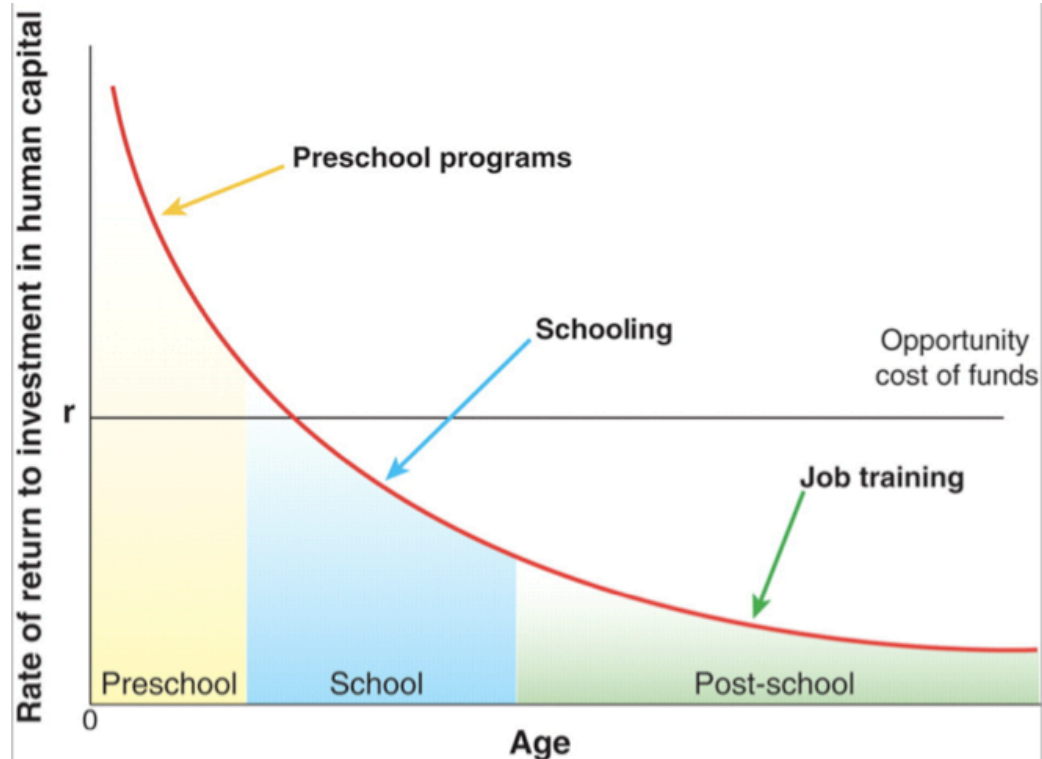
Notably, YEPs generally place a heavy emphasis on human resources, which can drive up the cost of youth interventions. The role of adults, often well trained in youth empowerment practices, tends to form a significant aspect of YEPs' theory of change. Research by The Finance Project that surveyed fourteen YEPs in the USA found that staff salaries and benefits accounted for an average 54% of intervention costs (Gray & Hayes, 2008). Annual costs per youth directly involved in YEPs in the USA can approximate, for example, \$1,270 USD for the Wide Angle Youth Media program and \$1,726 USD for the Hampton Youth Commission (Gray & Hayes, 2008). Costs reflect total reported expenses for 2007 divided by number of yearly participants, not including youth indirectly served through youth-led community outreach or advocacy actions.

On the other hand, if YEPs can help circumvent the long-term consequences of social behavioral problems, the cost-savings to tax payers may be substantial (Cunha et al., 2005; Scott et al., 2001). Yet, the evidence to-date suggests that waiting to intervene in the lives of particularly troubled or disadvantaged adolescents, even assuming positive intervention benefits, requires greater cost for smaller effects compared to the investments that could have been made earlier in the young person's life. In other words, in times of particularly strained economic resources for social investments, those advocating for increased funds for adolescent programming in the broader policy debates confront what Heckman and Masterov (2007) describe as "the productivity argument for investing in young children." Heckman and Masterov (2007) explain, "Because of the dynamic nature of

the skill formation process, remediating the effects of early disadvantages at later ages is often prohibitively costly.”

Heckman and colleagues present the summary of economic modeling of return on investment in human capital at different stages of the life cycle from early childhood to young adulthood in the figure 1.9 graph. This modeling is based on a number of assumptions, limited data, and experimental studies from a limited range of intervention types for young people (e.g., General Education Development (GED) and job training programs). As such, more theory-driven youth development or youth empowerment interventions may produce more substantive effects that could alter such calculations of the value of investment in this age group. To date, however, adequate research into these interventions and their impacts is lacking. Whether YEPs show effective or ineffective results, the economic implications underscore the need for rigorous evaluation to ensure that resources are invested in interventions and practices that produce intended outcomes for intended populations.

Figure 1.9. Rates of return on human capital investment in disadvantaged children



J J Heckman Science 2006;312:1900-1902

“Rates of return to human capital investment in disadvantaged children. The declining figure plots the payout per year per dollar invested in human capital programs at different stages of the life cycle for the marginal participant at current levels of spending. The opportunity cost of funds (r) is the payout per year if the dollar is invested in financial assets (e.g., passbook savings) instead. An optimal investment program from the point of view of economic efficiency equates returns across all stages of the life cycle to the opportunity cost. The figure shows that, at current levels of funding, we overinvest in most schooling and post-schooling programs and underinvest in preschool programs for disadvantaged persons.”

Considering harm and null results

While the literature on youth empowerment overwhelmingly assumes positive benefits, it is possible that YEPs may be ineffective or even harmful. Research by Dishion and colleagues (1999), for instance, has shown that programs that aggregate deviant youth together—even if for the purpose of positive

interactions—can unintentionally reinforce deviant behavior. Many youth programs try to transform the nature of peer relationships through a positive programming environment, but a Campbell Collaboration systematic review of after-school programs showed no evidence that they have been able to improve peer associations (Zief et al., 2006). Participants in one after-school program, in fact, were significantly more likely than control group participants to have deviant peer associations at follow-up (Zief et al., 2006). Results could be different for YEPs, which could argue a more intentional, theory-based process for facilitating positive peer environments. This, however, is yet to be seen in the empirical evidence and thus leaves open the possibility of unintended harmful effects.

The reality of ineffectual youth programming was prominently displayed by the Cambridge-Somerville Youth Study ($n = 650$) that evaluated a multi-year, multi-component intervention for child and adolescent boys providing a range of services, including counseling, academic services, family guidance, and recreation (McCord & McCord, 1959). At 18-year follow-up, the study found the number of participants to have committed crimes in childhood and adulthood and the number of crimes committed to have been approximately equal between intervention and control groups. The investigators deemed the well-resourced intervention a failure.

Other studies have found that higher self-esteem has correlated with higher hostility, and offenders have reported higher emotional intelligence scores than non-offenders (Baumeister et al., 1996; Hemmati et al., 2004). While these studies do not establish causality between self-esteem and hostility or emotional

intelligence and offending, they leave the possibility nonetheless that developing such assets may have unintended consequences for antisocial behavior. The functions of YEPs may curtail any such unfavorable effects, but certainty requires a robust evidence base. This constitutes one justification for including antisocial behavior as a secondary outcome in chapter two's systematic review.

Moreover, YEPs might unsuccessfully strive to improve developmental assets. Programs focused on increasing the role of young people as leaders in program decision-making might in practice downplay or neglect valuable expertise of adult youth workers. By consequence, programmatic decisions made by youths could lead to null or negative intervention effects. Attempts at empowerment might ultimately reinforce traditional power relationships; enter adolescents into challenges that leave them feeling inadequate and disillusioned; or elevate some youths over others, thus yielding 'success stories' with a few participants while circumventing the growth potential of others (consider Cooke and Kothari's (2001) "tyranny of participation").

Youth empowerment is theory-driven, and this basis may augment the likelihood of success in affecting youth outcomes. Nonetheless, even assuming the precepts of theory driving youth empowerment are valid, their translation to practice is, perhaps, highly susceptible to diminishment. Constructs such as social cognitive theory as a foundation for intervention with adolescents assume a considerable concentration of positive social stimuli to overcome competing influences and change outcomes, such as self-efficacy, as hypothesized. The extent to

which YEPs can and do provide the *level* of positive mastery experience and social persuasion required to change otherwise vulnerable adolescents' beliefs and noncognitive skills to a meaningful degree is unknown. As mentioned earlier in the chapter, YEPs operate amidst complex ecological factors, and these can either support or obstruct YEPs' efforts to positively influence adolescent outcomes.

Finally, the broader question remains as to the extent to which programs can achieve desired changes in social and behavioral outcomes for the adolescent age group. As described above, Heckman and colleagues have pointed to results of experimental studies of high-quality early childhood interventions showing significant long-term impacts against the backdrop of less compelling evidence for interventions targeting older age groups, and they have argued that interventions need to be targeted much earlier for optimal results (Heckman, 2000; Heckman & Masterov, 2007). Similarly, Patterson and colleagues (2000) found deviant peer associations to be significant predictors of multiple expressions of antisocial behavior; results of their structural equation modeling indicate that interventions should target children prior to adolescence in order to effectively avert such problematic associations.

These studies give cause for reservation and provide equipoise about the extent to which stakeholders might expect automatic impacts of adolescent programming, and they reinforce the need for high-quality research into the processes and effects of YEPs. This may be particularly true for high-risk adolescents targeted by YEPs who may have lacked the early interventions and

developmental assets to form a foundation for empowerment-based programming. According to Heckman's research conclusions, "skill begets skill" (Heckman, 2000). High-risk adolescents will often not have the foundational non-cognitive skills, as a result of deficits of previous capacity-building, that have the opportunity to multiply in the context of a YEP or other forms of youth development programming.

Contrary to the aforementioned challenges, a small number of youth development programs, such as Big Brothers/Big Sisters and the Quantum Opportunity Program (QOP), have effectively improved adolescent emotional and/or behavioral outcomes as demonstrated by rigorous experimental or quasi-experimental evaluations (Hahn et al., 1994; Tierney et al., 2000). Even still, evidence suggests that effective adolescent programming falls short of the magnitude of effects produced by effective early childhood programming. In the well-known experimental study of the Perry Preschool program for disadvantaged children, for example, long-term follow-ups showed that 20% more intervention participants graduated high school than control participants; this compares to a 7% difference between groups demonstrated by the QOP study (Schirm et al., 2003; Schweinhart & Weikart, 1998). QOP also did not show effects for the same diversity of outcomes (e.g., problem behaviors and academic testing scores) impacted by Perry Preschool. For young people who did not have the advantage of effective early childhood education, even a smaller statistically significant effect can be an important success, but the difference nevertheless highlights the added challenges for adolescent programming. The extent to which YEPs achieve their stated impact

goals is largely unknown and constitutes the purpose of the research presented in this dissertation.

Conclusion

The literature review reported in this chapter aimed to summarize justifications for youth empowerment programs (YEPs) both in terms of the need and the theoretical basis for how youth empowerment might address the need. The background has also examined issues that could undermine these justifications and render YEPs unsuccessful at improving adolescents' developmental outcomes.

Although most young people survive the biological and social trials of adolescence quite resiliently, many groups of youth still face concerning problems thwarting their healthy development and social inclusion. Even youths who avoid the most concerning problems associated with adolescence—e.g., school failure, criminal behavior, and substance abuse—may not have opportunities to maximize their developmental potential. Youth empowerment programs are not always described or justified in a theory-driven manner, and this may contribute to some of the heterogeneity in the field. Nonetheless, much of the youth empowerment literature and the basic tenets of YEPs are connected to substantive theories that articulate how environmentally-focused, participatory, and prosocial intervention could strengthen young people's development assets, which in turn could improve distal social, health, and economic dimensions of wellbeing. Despite the rationale, however, the theory of change has multiple points of vulnerability. These include difficulties with changing adolescent outcomes, susceptibilities to deleterious peer

contagion effects, and the extent to which youth empowerment programming is even an appropriate course of action for many of the adolescent groups to which it is delivered.

Guided by a systematic, evidence-based social intervention approach, this dissertation investigates the impacts and processes of YEPs in order to help better understand what the best available evidence tells us about the applicability of youth empowerment for measurably improving adolescent wellbeing. This process begins with a systematic review of the effects of YEPs on adolescents' self-efficacy and self-esteem reported in the following chapter.

Chapter 2: Secondary Research: The State of the Evidence from a Systematic Review

Introduction

Despite the increasing popularity of involving young people in the processes that affect their lives and communities, little is known about the demonstrated impacts that such participatory programming has on young people (Gray & Hayes, 2008; Zeldin et al., 2000). When an intervention area lacks a “recent, high quality systematic review of the relevant evidence,” the Medical Research Council’s (MRC’s) framework for complex interventions suggests that one should be conducted before starting further primary evaluation (Craig et al., 2008, p. 3). As such, this chapter reports the methods and results of a systematic review of the effects of youth empowerment programs (YEPs) for improving self-efficacy and self-esteem of adolescents (ages 10-19). The review is guided by methodological and reporting standards held by the Campbell Collaboration and the Cochrane Collaboration Handbook (Higgins & Green, 2009).

Background

Description of the intervention

This review defines YEPs as interventions that regularly involve young people as partners and participants in the decision-making processes that determine program design, planning, and/or implementation. With the support of caring

adults, YEPs engage young people in program leadership as a characteristic of their involvement in safe, positive, and structured activities.

Common examples of YEPs are found in *particular* youth councils, teen centers, community-based participatory research programs, social action and advocacy groups, peer education models, and informal and non-formal education programs that regularly integrate youth participation in program decision-making, as stated above. Structurally, this participation within programs often takes the form of advisory councils, committees, youths on boards, workgroups, or staff positions. Sometimes, young people and adults serve together in formal leadership capacities such as committees; other times, membership is reserved exclusively for youths with adults acting in more of a supportive role.

Youth empowerment involves a collective, democratic, and prosocial process of engagement, which implies group interaction (Cargo et al., 2003; Jennings, 2006). Consequently, exclusively one-to-one youth development interventions, such as most mentoring schemes, are not reviewed here.

Like YEPs, many non-empowerment-based out-of-school programs involve structured activities and safe spaces during hours that adolescents need them most. They do not qualify as YEPs, however, if youths are not systematically involved with program decision-making. Some peer education models, for example, may only activate adolescents in content delivery rather than shaping program planning and implementation (Shiner, 1999). Often, youth centers and out-of-school time program schedules, objectives, and activities are adult-driven. Youths may occasionally be asked for their input or sporadically involved in programmatic

decision-making, but if their involvement is not structured so as to ensure opportunity for real influence and regular participation in programmatic decision-making processes, the intervention is not empowerment-based.

While some formal education systems also employ increasingly participatory approaches (Hannam, 2001), this review focuses on youth empowerment initiatives outside of formal schooling. An analysis of effectiveness evidence and unique implementation issues for youth empowerment within formal education would be a valuable undertaking meriting a separate review.

Additionally, many programs and activities involving emerging adults (aged 18 to 25) also share empowerment characteristics. Research over the last decade, however, has demonstrated this age group to be socially and cognitively distinct from adolescence (Arnett, 2000, 2006). Moreover, the conditions under which many such activities take place (e.g., in university settings) can differ in many respects from community-based YEPs serving adolescents; this would introduce substantially more heterogeneity to the review. Given differences between populations and intervention settings, this review only focuses on YEPs that target adolescents.

Outcomes of interest

Youth empowerment programs aspire to develop critical psychosocial assets among participating youths. Supportive relationships, affirming environments, skill-building, and the integrated experiences of ownership, challenge, and mastery associated with meaningful youth participation are hypothesized to impact strength-based competencies and behaviors among youth participants. In turn,

youth development literature expects that these assets serve as pathways to distal indicators of success and wellbeing (e.g., academic achievement) and as protective factors against consequences of social exclusion (e.g., antisocial behavior). The outcomes of interest for this review are discussed below (measures are discussed in the methodology section).

Primary outcomes. Self-efficacy and self-esteem constitute the primary outcomes for this review. As Bandura explained, “perceived efficacy is a judgment of capability; self-esteem is a judgment of self-worth” (Bandura, 2006b). By engaging young people as valued partners, YEPs aim to improve young people’s belief in their own worth as well as in their ability to shape their lives and environments (Kirby & Bryson, 2002, p. 24).

While self-efficacy and self-esteem are distinct concepts (Gilad et al., 2004), Judge and colleagues (2002) have demonstrated that the two traits are highly related and the combination of the two can give a better prediction, for example, of job satisfaction and performance. Moreover, the frequent co-presence of the two constructs as suggested outcomes in literature involving youth empowerment reinforces the sensibility of pairing self-efficacy and self-esteem as primary outcomes from an evaluative perspective (Anderson & Sandmann, 2009; Jennings, 2006; Oliver et al., 2006; Roth & Brooks-Gunn, 2003c; Sinclair, 2000).

High *self-efficacy* has been shown to predict better performance in work, academics, and sports; increased happiness, job satisfaction, and persistence; improved safe sex practices; and successful smoking cessation and prevention (de

Vries et al., 1988; Judge & Bono, 2001; Kalichman & Nachimson, 1999; Martin & Gill, 1991; Multon et al., 1991; Natvig et al., 2003). A meta-analysis conducted by Stajkovic and Luthans (1998) found that self-efficacy accounted for a 28% improvement in work-related performance. Lower self-efficacy, on the other hand, appeared to predict higher levels of depression among young people (Bandura 1999) and was associated with higher alcohol use (Taylor, 2000). Notably, literature on self-efficacy frequently delineated between general and task-specific self-efficacy with arguments both for (Chen et al., 2001; Judge et al., 2002; Luszczynska et al., 2005) and against (Bandura, 1997; Stajkovic & Luthans, 1998) using general or global measures. Both aspects of self-efficacy were included in this review.

Evidence suggests that high *self-esteem* is related to high social support and resilience (Dumont & Provost, 1999; Hoffman et al., 1988) whereas low self-esteem was related to depression, anxiety, and suicidal ideation (Newbegin & Owens, 1996; Overholser et al., 1995; Rosenberg et al., 1995). Boden and colleagues (2008) found self-esteem to be an important “risk marker variable, with low self-esteem being associated with a range of negative outcomes,” and they found high self-esteem at age 15 to be a significant predictor of life satisfaction and peer attachment at ages 18, 21, and 25. Research by Baldwin and Hoffman (2002) indicated that self-esteem changes dramatically during adolescence, emphasizing the special importance of interventions that foster higher and more stable self-regard through this volatile life period.

Secondary outcomes. While self-efficacy and self-esteem were included as primary outcomes for their prominence in the youth empowerment literature and theoretical connections, they were not the only strength-based indicators associated with YEPs. The review's secondary outcomes included several other developmental assets that YEPs are believed to improve in young people.

Developmental assets. Developmental assets can include a range of "internal and external strengths within an individual's social ecology that are predictive of positive outcomes, including health, mental health, and education" (Kia-Keating et al., 2011). Forging *social supports and positive connections* between youths and their peers, communities, teachers, and families constitutes a central pillar of YEP models (Jennings, 2006; Kirby & Bryson, 2002; Roth & Brooks-Gunn, 2003b; Villarruel et al., 2003). Such relationships are frequently discussed as both a key process component as well as an expected outcome of successful YEPs. As young people actively participate in collective decision-making processes, dynamic social environments, and challenging new experiences, it is expected that they acquire and develop transferable *social skills and competencies* (Kirby & Bryson, 2002; Wilson et al., 2006).

Emotional Intelligence (EI) consists of domains related to aspects of processing, understanding, and managing emotions. Gundlach and colleagues (2003) conceptually argued that increasing EI may act as an important pathway to increasing self-efficacy, reinforcing an important role that some believe EI may play in YEPs' theory of change (Barber, 2007). Youth empowerment programs that

involved adolescent populations especially vulnerable to situations of conflict and stressful life events may prioritize youth empowerment as a strategy to strengthen young people's problem-solving and coping skills to navigate difficult situations. Adeptness in problem-solving skills is frequently highlighted as a valued characteristic of organizational and community leaders (Mumford et al., 2000)—roles into which YEPs invite young people.

While this review does not limit itself to YEPs centered on social action activities, many YEPs do heavily incorporate community engagement and social advocacy themes into their programming. Therefore, strengthening *civic engagement* among young people is a key driver for many YEPs (Jennings, 2006). The United Kingdom, for example, has a considerable history of local youth councils through which youth empowerment is often a vehicle for activating young people in neighborhoods and public action (Matthews, 2001). Civic engagement can be expressed in many ways, including volunteering, membership in civil society clubs and organizations, beliefs concerning the importance of civic engagement, expectations of future community involvement, voting, and political participation.

Academic achievement. Several of the aforementioned primary and secondary outcomes (e.g., self-efficacy and emotional intelligence) have been shown to predict academic performance (Pajares, 1996; Parker et al., 2004; Petrides et al., 2004). Through strengthening proximal outcomes concerning young people's psychosocial assets, YEPs may have indirect impacts on academic performance via improvements in noncognitive abilities (e.g., motivation, emotional traits, and social

skills) that mediate academic achievement (Cunha et al., 2010; Heckman et al., 2006). Research by Berndt and Keefe (1995) indicated that youths who reported prosocial peer interactions—a central tenet of youth empowerment—were more likely to participate actively in school and extracurricular activities. YEPs may also directly affect academic performance when educational goals and activities are integrated in particular interventions. Academic performance can refer to standardized test scores, completion, and grades.

Antisocial behavior. Youth empowerment by definition approaches young people from a strengths-based perspective that translates to a primary focus on recognizing and enhancing youths' developmental assets. The focus of YEPs on young people's strengths, however, does not preclude YEPs from having an impact on antisocial behaviors of common concern to communities and policy-makers. As stated by Roth and Brooks-Gunn (2003a), "*The goals of youth development programs promote positive development, even when seeking to prevent problem behaviors.*" Research suggests that enhancing adolescents' assets can be an effective course for reducing problem behaviors (Aspy et al., 2004; Kia-Keating et al., 2011). By enhancing psychosocial protective factors, engaging youth in constructive activities during vulnerable out-of-school hours, and strengthening young people's stake in their environments, proponents contend that YEPs may be at least as effective as those directly and primarily aimed at curtailing antisocial behavior.

Previous related reviews

No systematic review, to the author's knowledge, has been published that specifically addresses the impacts of youth empowerment. This section discusses two reviews synthesizing research for positive youth development and empowerment broadly, which provide relevant insights informing the planned review.

The Catalano and colleagues (2004) review on positive youth development (PYD) was commissioned by the US Department of Health & Human Services, completed in 2002, and conducted by a team of researchers at the University of Washington (Catalano et al., 2004). The review ultimately included twenty-five program evaluations, and findings indicate promising results for strength-based programs serving youth and children.

The study, however, did not report a systematic search strategy, extend beyond the United States, nor prospectively state sought-after outcomes. It did not include evaluations revealing null effects or significant effects that did not favor the intervention. The last characteristic is particularly concerning given that past research has demonstrated unintended harm caused by some well-intended youth interventions, which underscores the importance of understanding the consequences—positive, negative, or neutral—of youth development programs (Arnold & Hughes, 1999).

The review's definition of PYD is very broad, including any intervention that meets at least *one* of fifteen constructs¹; none of which stipulated involvement of young people in program decisions or design. The broad inclusion criteria for PYD programs coupled with a lack of predetermined outcomes largely explain why so many evaluations were included in the review. In contrast to the Catalano and colleagues review, the forthcoming review will focus on youth empowerment, reduce program heterogeneity, and minimize the chance of spurious conclusions due to a lack of predetermined outcomes.

The Wallerstein (2006) review, conducted for the World Health Organization, was the only review identified by the author that directly addressed empowerment strategies. The review explored empowerment broadly for all age groups and with an interest in health outcomes. Wallerstein gives a useful overview of various themes within the broader empowerment movement and gives a framework for empowerment that includes multiple levels of outcomes. The broad, international review of outcomes linked to empowerment offers valuable context for this review. The resulting framework includes a heavy emphasis on self-efficacy, community engagement, and social bonding, which reinforce their importance as outcomes of interest in the present review.

¹ Constructs included the following fifteen objectives: promotes bonding, fosters resilience, promotes social competence, promotes emotional competence, promotes cognitive competence, promotes behavioral competence, promotes moral competence, fosters self-determination, fosters spirituality, fosters self-efficacy, fosters clear and positive identity, fosters belief in the future, provides recognition for positive behavior, provides opportunities for pro-social involvement, and fosters prosocial norms.

A brief section of the review was devoted to youth empowerment. The author, however, did not discuss the quality of the evidence behind listed outcomes linked to youth empowerment, nor is it clear that the studies actually evaluated a YEP as defined by this review, versus less participatory youth development programs. The present review employed a search strategy for empowerment studies specific to youth and controlled impact evaluations.

The Wallerstein review was not conducted according to systematic procedures, no specific outcomes were identified for study inclusion, virtually all types of studies were acceptable for assessing effectiveness with no distinction made between study designs in discussing findings, and again there was no indication that the review made an effort toward including null or harmful effects. The author defended the review's approach with a dismissal of the use of systematic review methods and experimental designs to assess empowerment programs. As the present protocol runs contrary to this perspective, the author briefly addresses assertions stated by Wallerstein:

"Empowerment projects by their nature are complex and do not easily fit into an experimental design...causal relationships may be too complex to uncover within a changing social environment." While many experimental designs in practice are not sufficiently sophisticated to isolate the effects of *different program components* within an intervention, they are still capable of determining the impacts that participation in YEPs has on youths for outcomes of concern. The Zief and colleagues (2006) Campbell Collaboration review, for example, assessed the

evidence for after-school programs, which can be comparably complex and heterogeneous. Conceivably, moreover, a controlled trial could compare youth empowerment with an active comparison in which adolescents participate in a non-participatory though otherwise similar intervention (e.g., youth-driven smoking cessation campaign versus didactic smoking cessation education classes). In other cases, complementary process and qualitative studies can help clarify active ingredients within multi-component programs (Miller et al., 2003; Oakley et al., 2006).

“Expectations for individual health status changes may be unrealistic in short time frames...” For this reason, it is important that program evaluations delineate clear theories of change. While ambitious health changes may indeed take much time, proximal psychosocial outcomes mentioned by Wallerstein and earlier in this protocol may arise in shorter time periods. Such short-term impacts have been documented by other evaluations of positive youth development programs broadly (Catalano et al., 2004; Tierney et al., 2000) and can lead to more distal health and behavioral outcomes of concern (Kia-Keating et al., 2011).

“Evaluation should...use a mix of methodologies and designs – rather than randomized control trials...” The author appears to treat mixed methods and RCTs as mutually exclusive. To the contrary, experimental designs of social interventions are often used within a mixed methods approach and are most useful when complemented by other forms of research addressing process and implementation (Mayo-Wilson, 2007; MRC, 2008; Oakley et al., 2006). Experimental designs of social

interventions are increasingly an area of interest for participatory research approaches as well (Jones et al., 2008). Experimental and quasi-experimental trials are best suited to answer a *specific* type of research question concerning causality and effectiveness. This is not the only type of question of concern for youth empowerment, but it is the primary focus of the present review.

Objectives of the review

This review aims to systematically investigate and summarize the state of the evidence on the impacts of YEPs on adolescents' self-efficacy and self-esteem. Specifically, this review endeavors to address the following questions in order to contribute to the body of evidence available to stakeholders and researchers so as to improve services and supports for young people:

1. Impacts: Do YEPs affect adolescents' sense of self-efficacy and self-esteem? Additionally, does the intervention affect hypothesized secondary outcomes, including social supports, emotional intelligence, social skills, academic performance, and antisocial behavior? If so, is there sufficient evidence to indicate that the secondary outcomes correlate with this review's primary outcomes as suspected?

2. Heterogeneity: Do YEPs affect various sub-groups differently? Do variations in program design or implementation—with special consideration to levels of participation—also reveal trends by which outcomes differ? Does heterogeneity in evaluation quality and design correlate with certain outcome patterns?

3. Future research: What are the knowledge gaps revealed by this review, and how can they inform future research on youth empowerment – especially future impact evaluations?

The expectation of policy-makers, funders, and community organizations to actively involve young people in program design and implementation is becoming increasingly popular. This review works toward a better understanding of the measured merits behind that option. If evidence allows, the review further aims to advance knowledge with respect to the decisions within youth empowerment (e.g. how youth empowerment works best and for whom).

Methodology

Criteria for considering studies for this review

Types of studies. The evaluation must have involved either an experimental or quasi-experimental design with a prospectively assigned control group. Quasi-experiments needed to have taken steps to establish a reasonably credible counterfactual. As such, only quasi-experiments that used matching or statistical methods (e.g., propensity scores) to ensure that the control group was similar to the intervention group at baseline were included.

This review accepted trials involving control groups with no service provided and/or trials with comparison groups that involved alternative services. Alternative services should not have facilitated youth involvement in program decision-making or active leadership roles. Basic recreational or educational

activities, such as instructional sessions or presentations, games, and informal athletic activities, for example, could constitute a comparison to youth empowerment programming.

Types of participants. The target population is adolescents. The age definition of adolescence is not consistent across institutions and cultures. For inclusivity, this review follows the definition of adolescence as ages 10-19 as classified by the World Health Organization, United Nations, and World Bank (UNFPA, 2008; World Bank, 2003; World Health Organisation, 2009). According to recommendations from Campbell Collaboration Social Welfare Group peer-reviewers, it was decided that at least 75% of the study sample must have met this age criteria.

Types of interventions. This review investigated the impacts of YEPs that regularly involve adolescents in determining program design, activities, and/or implementation. Structurally, this participation often takes the form of democratic decision-making processes involving, for example, youth councils, committees, youth on boards, workgroups, staff positions or other youth groups with regular opportunities for program decision-making. Sometimes, youths and adults serve together in formal leadership capacities such as committees; other times, membership is reserved exclusively for youths with adults acting in more of a supportive role. Programs must have involved regular access to a supportive adult or older youth leader, though this need not have involved one-to-one mentoring.

Delivery could have taken place in community-based or school-based settings so long as the intervention occurred regularly and outside of formal education. Interventions primarily within formal education, juvenile justice, residential programs, therapeutic interventions, conferences, or workshops were not included. Includable programs must have convened regularly (i.e., not a one-off event).

Types of outcomes. Studies must have measured at least one of the review's primary or secondary outcomes. Outcomes could have been measured by way of self-reports, third party or researcher observations, interviews, or official records. The review accepted measures that were and were not well validated.

This review's primary outcomes were self-efficacy and self-esteem. Self-efficacy included both general self-efficacy measures and task-specific self-efficacy (e.g., drug avoidance self-efficacy or sexual behavior self-efficacy). General and task-specific measures, however, were not meta-analyzed together given differences between the constructs. Self-esteem is most commonly assessed by the ten-item Rosenberg Self-Esteem Scale (Rosenberg, 1989). Some research has delineated between global self-esteem (e.g., measured by the Rosenberg scale) and specific self-esteem (e.g., measured by the area-specific Hare Self-Esteem Scale), but no specific self-esteem outcomes were measured by this review's included studies.

Secondary outcomes include several areas of other developmental assets—social supports and connections, social skills, emotional intelligence, coping and problem-solving skills, and civic engagement—as well as academic performance and

antisocial behavior. The review's protocol listed prominent example measures for each outcome.

Search methods for identification of studies

Electronic searches. The investigator searched twelve major electronic databases for this review: Applied Social Science Index and Abstracts, Australian Educational Index, British Educational Index, CINAHL, Cochrane Library (CENTRAL), Dissertation and Theses Abstracts, EMBASE, ERIC, Medline, PsycInfo, Social Service Abstracts, and Sociological Abstracts. One review author (MM) conducted the literature search.

Additional relevant institutional web-based publication databases searched included Chapin Hall (University of Chicago), Out-of-School Time Program Research & Evaluation Database (Harvard Family Research Project), Innovation Center, National Clearinghouse on Families & Youth (US Administration of Children & Families), Public/Private Ventures, Search Institute, the UNICEF Evaluation and Research Database (ERD), the Australian Clearinghouse for Youth Studies (ACYS), National Council for Voluntary Youth Services (NCVYS) Publications, the UK DCSF Inclusion Development Programme (IDP) Publication Catalogue, and the World Bank Poverty Impact Evaluations Database. All databases and websites searched, along with search dates, are included in appendix 2.1.

Search terms. The following search terms were used for each of the
aforementioned databases²:

Population:

(young OR youth OR child* OR teen* OR adolescen* OR minors OR school ADJ student*
OR boy* OR girl* OR NEETs OR NEET OR 14-19).ab,ti.*

AND

Intervention:

(pyd OR cyd OR empowerment OR youth ADJ engag OR volunteerism OR volunteering
OR youth ADJ advocacy OR youth ADJ activism OR youth ADJ development OR youth
ADJ leader* OR youth ADJ inclusion OR community ADJ service OR after ADJ school OR
afterschool OR youth ADJ1 decision-making OR youth ADJ driven OR youth ADJ run OR
youth ADJ adult ADJ partnership* OR youth/adult ADJ partnership* OR youth-adult
ADJ partnership* OR youth ADJ action OR youth ADJ1 involvement OR youth ADJ
participation OR young ADJ people* ADJ participation OR youth ADJ led OR peer ADJ
education OR peer ADJ led OR peer ADJ participation OR youth ADJ voice OR service
ADJ learning OR youth ADJ council* OR teen ADJ council* OR non-formal ADJ education
OR nonformal ADJ education OR informal ADJ education OR teen ADJ cent* OR youth
ADJ cent* OR participatory ADJ research).ab,ti.*

AND

² Variations of Boolean operators, wildcard symbols, and field indexes were used depending on the nature of the specific database.

Methods:

(control OR random* OR trial* OR effectiveness OR efficacy OR compar* OR clinical* OR experiment* OR impact ADJ evaluation OR impact ADJ study OR impact ADJ assessment OR outcome ADJ evaluation OR outcome ADJ study OR outcome ADJ assessment).af.*

Searching other resources. In order to explore potentially eligible studies among unpublished as well as published literature, institutions and individuals regarded as professional leaders in the area of youth development and research were contacted individually and asked for any leads on specific studies, or databases likely to include studies, that might have met the review's inclusion criteria. Contacts were made to seventy professionals representing over fifty institutions (e.g., foundations, non-governmental organizations, international organizations, government agencies, and research institutes). A typical email to relevant professionals and institutions is given in appendix 2.2. Most institutions contacted had country-specific focuses in seven countries, but eleven institutions had international scope (e.g., The World Bank, United Nations agencies, and global foundations). While responses provided insightful information and resources related to youth empowerment, the professional outreach did not yield any additional eligible studies for this review. Respondents generally felt that the field lacked examples of rigorous impact evaluation for youth empowerment programs.

Data collection and analysis

Selection of studies. This review accepted both published and unpublished studies for inclusion, and there were no exclusion criteria based on where the study was conducted or the reporting language. A second reviewer (Dr. Paul Montgomery, Centre for Evidence Based Intervention, University of Oxford) contributed to the screening and selection of studies, as well as data extraction. The Cochrane Handbook states that it is essential that reviews be undertaken by more than one person so as to improve the methodological integrity of the process (section 2.3.4.1; Higgins & Green, 2009).

Both reviewers reviewed all citations and discussed and resolved issues concerning study inclusion and exclusion. A screening guide was used to determine inclusion or exclusion and is provided in appendix 2.3. An abstract was automatically excluded if it was rejected by both reviewers according to any of the six screening criteria. Full reports of studies were retrieved (by this author), reviewed and discussed (by both reviewers), and coded (by both reviewers) using the screening guide for any study that was not excluded based on its abstract.

Data extraction and management. Studies selected for inclusion or as relevant excluded studies were further coded by this author using the forms in appendix 2.4, primarily for intervention characteristics, and appendix 2.5, to guide discussion of study quality. Both reviewers reviewed the studies and the coding, and any disagreements were discussed and resolved between the two reviewers.

Relevant data on intervention and study characteristics (summarized in table 2.1) were extracted from the coding forms for analysis and discussion.

Assessment of risk of bias in included studies. A systematic approach to assessing study quality on the basis of predetermined criteria was used by the reviewers, based on a previously published systematic review (Zief et al., 2006). Forty-one characteristics of study design and reporting were used to appraise study quality, and four standards in particular were considered priorities for judging study quality. These standards included evidence of (a) no significant control group contamination, (b) no significant overall study attrition nor differential attrition that would bias the results, (c) appropriate statistical methods used, and (d) primary outcomes having been measured at follow-up for all available sample members, thereby meeting the qualification of ‘intention-to treat’, not ‘treatment-on-treated’, analysis. The checklist used to appraise study quality is included in appendix 2.5. The checklist was intended as a discussion guide to facilitate meaningful analysis of the quality of included studies; it was not part of the inclusion criteria nor were studies ranked according to a particular grade or score.

Measures of treatment effect. For the included studies with comparable continuous outcomes, Hedges g was calculated using means and standard deviations (SDs). Unlike Cohen's d, Hedges g corrects for small sample size and may give a more conservative estimate of variance.

In future review updates, the following analysis plans apply. Continuous data that must have values greater than 0 will be considered skewed if the mean is less

than the sum of two standard deviations (Altman et al., 2001; Higgins & Green, 2009). Primary authors will be contacted for more information, log transformed data, or raw data if skewed data is suspected. The reviewers will calculate and compare standardized mean differences across studies if the same outcomes are measured in different ways; weighted mean differences will be calculated for outcomes measured in the same way. Log odds ratios with 95% intervals will be calculated for dichotomous outcomes data (Higgins & Green, 2009).

When means and SDs are unavailable, the author will calculate Hedges g using other available statistics, for example an F-test and p -value or t-test and p -value. When data are presented in several forms that could be used to calculate an effect size, we will select the least form that is closest to the raw data. That is, when mean changes are reported in addition to ANOVAs, we will select mean change scores. When means and SDs are not available and Hedges g is calculated using other statistics, we will note this in the text.

A random-effects model was used to calculate combined weighted mean effect sizes. The author assumed that differences between studies' interventions and populations were likely to have caused variations in effect sizes, rendering a fixed-effects model inappropriate for this review (Borenstein et al., 2007), especially given that included studies came from different regions of the world.

Unit of analysis issues. The included studies in this review all treated individuals as the unit of analysis. If future updates identify includable studies in which groups (e.g., programs, neighborhoods, classrooms, or schools) are the unit of

analysis, the reviewers will have to determine if results can be meta-analyzed without a confounding interaction between the intervention effect and unit of analysis. In order to combine individual-level and cluster-level trials, studies will need to have at least reported adequate information to adjust for possible design effects, including data to calculate 'effect sample sizes' (Donner & Klar, 2002). Sensitivity analysis may be used to assess effects of varying levels of randomization.

Dealing with missing and incomplete data. Only one study (Berg et al., 2009) did not provide means and standard deviations for the review's primary outcomes in the write-up, but the authors provided this data upon request. Two studies (Berg et al., 2009; Olson-Merichko, 2006) did not impute or include data for those lost to follow-up, nor, in Berg and colleagues' case, for those not analyzed due to lack of compliance (per-protocol analysis). Missing data for Olson-Merichko's study was unlikely to be consequential with only one participant lost to follow-up, and Berg and colleagues' study was not meta-analyzed with the other two included studies.

Assessment of heterogeneity. Heterogeneity according to study quality, population characteristics, and intervention characteristics was assessed and summarized in both table and narrative format. With respect to intervention heterogeneity, differences in the levels of youth empowerment and the nature of program activities were descriptively assessed from reported intervention descriptions according to extent of youth participation in decision-making, types of

skill-building activities, and nature of youth-adult relationships, assisted by prominent frameworks for youth empowerment described in chapter one.

Assessment of publication bias. Both published and unpublished studies were included. Authors of included studies were contacted and asked to provide statistics for any of the review's primary or secondary outcomes that were measured and not reported. Any potential biases from selective or incomplete publication and/or reporting are discussed in section, 'risk of bias in included studies.'

Treatment of qualitative research. Qualitative studies can contribute usefully to a more holistic understanding of youth empowerment processes and experiences. This review, however, centered on a research question concerning effectiveness as assessed by impact evaluation designs capable of establishing a credible counterfactual. As such, the author maintained a focus on controlled trials for the purposes of this review, though he discusses qualitative process research associated with included studies (as outline in section 20.2.3 of the Cochrane Handbook; Higgins & Green, 2009).

A qualitative systematic review exploring process, mechanisms, and perceptions underlying youth empowerment would indeed be a valuable enterprise, but it would be a distinct and considerable undertaking meriting unique criteria, methods, and a separate review altogether (Dixon-Woods et al., 2006; Higgins & Green, 2009; Jones, 2004).

Process and implementation studies play a vital role in making sense of the results of an impact evaluation and therefore should ideally accompany any trial measuring the effectiveness of a social intervention (Mayo-Wilson, 2007; MRC, 2008; Oakley et al., 2006). This is particularly important for YEPs, which rely heavily on process and can vary significantly in implementation. Process studies can involve a combination of qualitative and quantitative methods assessing programming aspects such as program quality; levels of youth engagement, participation, and satisfaction; fidelity to intervention manuals, curricula, or plans; program environment; nature of youth-adult and peer dynamics; and program 'dosage' and consistency. While accompanying process studies were not required for study inclusion, they were discussed in relation to study quality, better understanding of impact study results, and implications for YEPs.

Data synthesis

Subgroup analysis, moderator analysis, and investigation of heterogeneity. There were an inadequate number of includable studies with comparable outcomes to conduct subgroup analyses. If future review updates yield adequate data, the following plans apply. The review will explore potential differential interactions using baseline data on subgroup and program characteristics. In order to model heterogeneity among target populations, potential moderators will be gender, age range (within adolescence), race/ethnicity, and household income. Previous research has suggested relationships between these demographic characteristics and outcomes for adolescents in youth development settings (Altman, 1998; Eccles et al., 1997; Harris et al., 2001).

Program covariates will include duration and frequency of intervention, presence of a civic engagement component (i.e., volunteering, service-learning, advocacy, or public awareness), existence of a training or preparation component prior to or in tandem with leadership opportunities, and levels at which young people are involved in program decision-making. These program qualities all have important implications for the theoretical literature regarding which aspects of the various approaches to youth empowerment serve as active ingredients for achieving positive outcomes (Billig et al., 2005; Catalano et al., 2004; Chinman & Linnery, 1998; Jennings, 2006; WHO/UN, 1999).

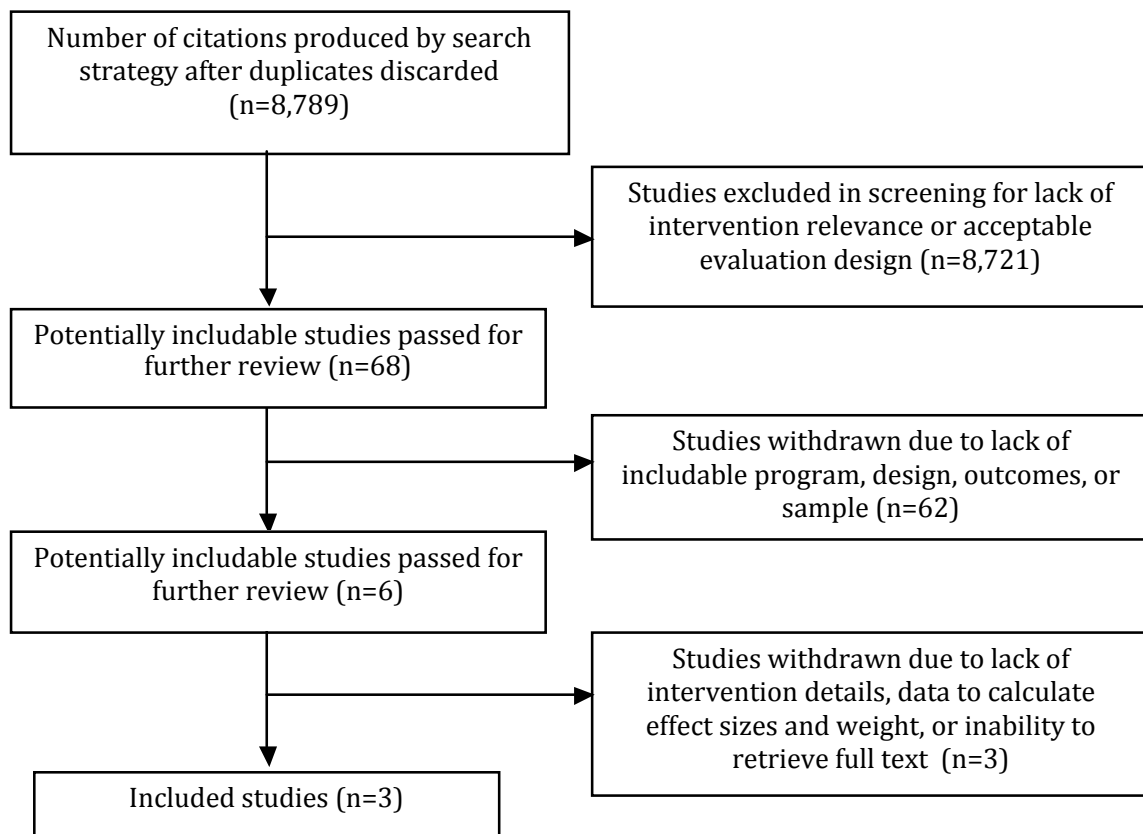
Sensitivity analysis. There were an inadequate number of includable studies with comparable outcomes to conduct sensitivity analyses. If future review updates yield adequate data, the following plans apply. Sensitivity analysis of included studies will be conducted to assess trends between study qualities and synthesized outcomes. Quality indicators will include allocation concealment, intention-to-treat, evidence of contamination, and, in the case of quasi-experimental trials, methods used to establish a credible counterfactual. Sensitivity analysis will examine whether weighted mean effect size differs between randomized trials and quasi-experimental studies, and between groups of trials with varying units of randomization. If there are no differences, then studies using either of these designs will be combined.

Results

Results of the search

From the electronic databases, a total of 7,985 citations were retrieved. An additional 804 citations were identified from relevant institutional web-based publication databases. The total number of citations retrieved was 8,789. Sixty-eight studies were passed for closer inspection by the reviewers. Of these, 62 studies were excluded for not meeting inclusion criteria and 3 studies were excluded due to insufficient data or intervention details; 3 studies ultimately met all of the review's inclusion criteria. Reasons for exclusion and citations for the 65 excluded studies that were reviewed in-depth are provided in appendix 2.6. Figure 2.1 provides a flow diagram of the review process.

Figure 2.1. Review flow diagram



Description of the studies

Included studies. The review identified three studies that matched all of the review's eligibility criteria. The studies included the Youth Action Research Project (YARP) in the USA evaluated by Berg and colleagues (2009), the Youth Leadership Program (YLP) in the USA evaluated by Olson-Merichko (2006), and the Questscope Non-Formal Education (QS NFE) program in Jordan evaluated by Morton and Montgomery (2011).

Two studies involved randomized controlled trials but had small sample sizes ($n = 127$ and $n = 40$), were both described as pilot studies, and were unpublished at the time of this review (Morton & Montgomery, 2011; Olson-Merichko, 2006). One study was a peer-reviewed quasi-experimental evaluation with a matched comparison design ($n = 316$) (Berg et al., 2009). The Morton & Montgomery (2011) study is in submission for peer-review. Review of the Olson-Merichko (2006) study is limited to details and data provided in the dissertation, as the author is deceased. Further details and data were required to adequately review the Berg and colleagues (2009) study, which were supplied by the primary author.

None of the three studies used a truly active-comparator design to assess youth empowerment against a comparison group with an intervention of comparable exposure without empowerment-based methodology. Berg and colleagues (2009) compared the intervention group to a group of young people participating in other summer employment programs, but the nature and dosage of these programs were not assessed. In Morton and Montgomery (2011), youths

randomly assigned to the wait-list control condition were offered a basic biweekly recreational activity that did not use empowerment-based methodology, but this was conducted at a lesser dosage than the YEP and primarily intended to maintain contact with participants in the control. Olson-Merichko's (2006) study involved a no-treatment control.

The mean ages of the study samples were similar (ranging from 15.2 to 16.0); the samples were otherwise notably heterogeneous. There were differences between study samples in terms of urban versus rural settings, cultural contexts, ethnicity, gender make-up, and life circumstances (e.g., in-school versus out-of-school). Posttests for Morton and Montgomery (2011) and Olson-Merichko (2006) were conducted at 4 months. While Berg and colleagues (2009) conducted a data collection at 3 months, this only captured a summer training institute component; the 12-month posttest captured the more empowerment-based components of the intervention (youth-led projects). The 12-month posttest for Berg and colleagues is reported in this review unless otherwise indicated.

Two of the evaluated programs, YARP and YLP were based on participatory action research models in the United States (Berg et al., 2009; Olson-Merichko, 2006). A substantial amount of youth empowerment literature in recent years has focused on participatory research as a means for engaging young people in programs, schools, and communities (Kirby, 2004; Ozer et al., 2008; Suleiman et al., 2006b; Worrall, 2000). One evaluated program, QS NFE, involved an empowerment-based non-formal education model for out-of-school youth in Jordan (Morton &

Montgomery, 2011). Non-formal education is often associated with empowerment approaches to working with marginalized populations through participatory learning (Castelloe & Watson, 1999; Moulton, 1997). This is a central aspect of the QS NFE theory of change with youth participants.

The programs have different characteristics in structure, context, and content; basic characteristics are outlined in table 2.1. Questscope non-formal education is a 24-month intervention, but only the first 4 months of the program are captured by the RCT. YARP is a 10.5-month intervention, but three data collections are conducted (3, 6, and 12 months). The YLP is a 4-month intervention. All three interventions involved weekly programming with minimum intended weekly exposure ranging from 2 to 4 hours for most portions of the interventions.

Given the intensive and manualized 7-week capacity-building institute, constant adult facilitator presence, emphasis on youth-adult partnerships not only in short-term decisions but also in ongoing youth-led projects, emphasis on collective action, and the 10-month intervention exposure, the YEP captured by Berg and colleagues' evaluation appears to reflect the most thorough application of the youth empowerment process theory of change among the three studies.

While youth in YLP and QS NFE did not receive financial compensation for their participation, YARP participants were employed and financially compensated as youth researchers. Youth participation in YLP and YARP was in both cases centered on explicit social action projects involving school or community advocacy. Direct social action or civic engagement was not a formal aspect of the QS NFE

intervention model. The YARP and YLP were both interventions designed and implemented by the respective researchers. Questscope non-formal education is an ongoing program jointly led by a non-governmental organization, Questscope, and the Jordanian Ministry of Education; the study authors were unaffiliated with the program design and implementation.

Youth participation in decision-making. The YLP and YARP interventions were developed specifically for the purpose of testing the impacts of youth empowerment. As such, these programs facilitated not only a high level of youth involvement in program decision-making, but also in initiating and planning activities and projects. Youth participants were primarily responsible for designing and implementing research projects, activities, and meetings, with adult facilitators acting as supporters and guides throughout the implementation process. In terms of Hart's Ladder of Children's Participation, empowerment generally ranged from 'adult-initiated, shared decisions with children' with respect to program initiation and training implementation, to 'child-initiated, shared decisions with adults' with respect to the research and community action activities.

The QS NFE program is based on a participatory methodology that involves youths in determining learning topics and social or recreational activities. This methodology stipulates regular involvement of young people in daily program decision-making and a co-learning relationship between youth and adults. While some QS NFE sites have included longer-term youth leadership roles and youth-led

planning, the intervention methodology does not formally necessitate this level of participation prompted by YARP and YLP. In terms of Hart's Ladder, the methodology is best described as involving programming that is adult-initiated with shared decisions with children. The associated process study suggests that in sites with lower implementation fidelity, programming slips to an approach more reflective of the 'consulted and informed' level of participation.

While all three studies included qualitative research into the implementation of the interventions, only Morton and Montgomery (2011) used a quantitative instrument to measure the extent to which participants felt empowered by the program process.

Adult involvement. All three studies described adults involved in the programs as 'facilitators.' To this end, each program articulated the adults' roles as facilitating youth participation and contribution rather than controlling or directing the nature of youth involvement. The YLP's adult involvement primarily included the lead researcher in training youth leaders and supporting youth-led projects. Similarly, YARP adults, consisting of researchers and trained project staff, led the youth research training, supported youth-led projects, and facilitated reflective discussions. QS NFE identifies and retrain teachers from the formal education system to serve as facilitators for the program outside of school hours.

The task of implementing quality youth empowerment processes can require special skills and competencies. This is especially true for those responsible for implementing empowerment programs with more marginalized youth with fewer

previous empowering experiences. Youth empowerment literature has emphasized the importance of adult training to help staff or volunteers develop facilitation and youth development skills for empowerment processes and overcome any conflicting deficit-based inclinations toward youth (Jennings, 2006). To this end, facilitators in YARP and QS NFE received structured training specifically in empowerment methodologies prior to intervention implementation. The YLP did not include an adult training component in empowerment methodology, though youth team leaders that led project teams did participate in 25 hours of training in leadership skills.

The YLP gave a particularly high level of autonomy to participants, often with youth groups meeting without adults to discuss and implement projects. Youth ‘team leaders’ were selected and trained to lead each group. QS NFE involved a higher level of adult control in program planning and implementation, partly due to needs in meeting certain educational curriculum standards established by adult officials. Among the three programs, YARP appeared to most closely approximate Wong and colleagues’ (2010) ‘pluralistic’ standard of youth-adult shared control, though all three programs stressed shared control to varying degrees.

Skill-building. YARP provided a distinct training component that engaged all program youth in a 7-week, 20-hour per week summer institute. The YARP training curriculum covered topics particularly focused on developing youths’ skills in participatory research. YLP involved 25 hours of leadership training curriculum for youth team leaders and QS NFE delivered 24 hours of leadership and research

training curriculum for the program's youth advisory council, but these subgroups of program participants were not included in the RCT samples.

All three intervention descriptions stressed the development of young people's social and leadership skills through a diversity of social activities and empowerment-based methodologies. Only YARP included a defined curriculum for the summer institute component of its skills-building activities. Otherwise, YARP and YLP skill-building activities largely constituted ongoing team research project activities, and QS NFE included a range of educational games as well as cultural, vocational, and recreational activities depending on the program site and youth preferences. Although QS NFE included a diversity of activities with potential to foster youth development, it did not engage youth in ongoing skill-building opportunities through team-based projects as the other programs did.

Literature on YEPs often suggests an important role for structured, ongoing training components specially designed to help prepare young people for meaningful participation in empowering program processes, specific program tasks (e.g., participatory research, photography, etc.), and broader civic engagement (Jennings, 2006; Wilson et al., 2006). The absence of such components for most youth in programs like QS NFE and YLP may have limited the programs' ability to develop young people's skills and fully engage them in participatory opportunities.

Table 2.1. Characteristics of included studies

	Berg et al., 2009	Olson-Merichko, 2006	Morton & Montgomery, 2011
Study Characteristics			
Design	Quasi-experimental with matched control group	RCT	RCT
Comparison	Alternative summer employment programs	No treatment control	Wait-list with basic biweekly recreational activities
Review's primary outcomes	Self-esteem, drug avoidance self-efficacy, sexual behavior self-efficacy	General self-efficacy	General self-efficacy
Review's secondary outcomes	Social assertiveness skills, social connectedness, school bonding, drug use, sexual behavior	Proactive coping, proactive attitude, team skills, delinquency, drug use	Social skills, social supports, prosocial attitude, conduct problems, emotional symptoms, local adult connectedness
Follow-up (months)	3, 6, 12	4	4
Sample size	316 (114 Intervention, 202 Control)	40 (20 Intervention, 20 Control)	127 (67 Intervention, 60 Control)
Analysis method	Completer analysis	Completer analysis	Intention-to-treat
Attrition (%)	26.3 ³	2.5	6.3
Sample age: mean (range)	15.2 (14-17)	16 (14-18)	15.9 (13-21)
Female (%)	51	72	15
Process study	Observations of facilitator performance	Qualitative youth interviews	Quantitative empowerment survey; qualitative youth and facilitator interviews
Program Characteristics			
Intervention	Participatory action research and community advocacy with training	Participatory research with training	Participatory non-formal education
Location	USA, urban	USA, rural	Jordan, urban
Dosage	20 hours per week (summer), 4 hours per week (school year)	2-3 hours per week	4-10 per week
Duration	11 months (7-week summer institute, 8-month projects)	4 months	24 months (3 8-month cycles)

³ Although the published article reports 17.4% attrition, later author contact updates attrition to 26.3%.

Structured youth training component	Yes (7-week leadership & research skills institute)	No (not for study participants)	No (not for study participants)
Service/advocacy component	Yes	Yes	No
Primary setting	Community; single site	School (meetings); single site	Schools (special program center venues); multi-site
Level of Hart's Ladder of Participation according to intervention description	Youth-initiated, shared decisions with adults (for projects; overall program adult-initiated)	Youth-initiated and directed (for projects; overall program adult-initiated)	Adult-initiated, shared decisions with youth

Excluded studies. Sixty-five studies were reviewed in-depth at the final level of screening and ultimately excluded. Reasons for exclusion are given in appendix 2.6. The majority of studies were excluded due to inadequate study design or the intervention description not matching the review's inclusion criteria. In some cases, studies were excluded because program descriptions in write-ups or author communications did not qualify the interventions as empowerment-based according the criterion of regular youth involvement in program decision-making. For some youth development programs that were experimentally or quasi-experimentally evaluated, author communications indicated isolated or ad hoc instances of youth participation in program decision-making but that this was not a regular element of the intervention design.

Three program evaluations matched all of the review's inclusion criteria except for having been primarily based in formal education. These included Allen and colleagues (1997), Lakin and Mahoney (2006), and Winkleby and colleagues (2004). All three studies evaluated school-based programs that implemented youth

empowerment through service or advocacy-oriented group projects. Because the study designs match the review's eligibility standards and the interventions contribute to the limited evidence-base for youth empowerment, the studies are briefly described here.

Allen and colleagues (1997) ($n = 695$) conducted a multisite experimental study of the Teen Outreach Program (TOP), a national volunteer service program that combined community service activities with classroom-based discussions on service experiences and broader adolescent development issues. The program was designed to incorporate regular involvement of youths in decision-making processes concerning discussion topics and service projects. Students participated in TOP for nine months (an academic school year). The intervention targeted changes in problem behaviors (pregnancy, school suspension, and academic failure) and did not measure any strength-based outcomes, such as self-efficacy or self-esteem.

The TOP research design mixed randomization between classes or students as units of assignment depending on agreements with particular schools. Twenty-five schools and 695 students were included in the study sample. Problematically, the study did not conduct a multilevel analysis to account for design effects from mixing individual and cluster-level units of assignment. As such, even if the study were included, the results could not be included in a meta-analysis unless primary data were available to redo analysis and calculate design effects (Donner & Klar,

2002). The study was conducted before these issues of design effect became prominent in the statistical literature in the late 1990s and early 2000s.

Lakin and Mahoney (2006) ($n = 43$) conducted a small, pilot experimental study that randomly assigned three classes to either a class-based participatory research and community-service intervention (two classes) or control (one class). Forty-five students participated in the evaluation (29 intervention and 14 control), and the intervention lasted ten weeks with two sessions per week. The small student sample size and small number of cluster units randomized rendered the study particularly susceptible to underpowered results and confounding influences. Unlike TOP, this intervention was not an ongoing program and was developed specifically for the purpose of the study. Also in contrast to TOP, Lakin and Mahoney were primarily interested in improving strength-based outcomes—particularly, self-efficacy, empathy, and civic engagement indicators.

Winkleby and colleagues (2004) ($n = 813$) conducted a cluster-RCT involving ten continuation high schools and 11th and 12th grade students (5 intervention schools, $n = 375$; 5 control schools, $n = 438$). Schools were randomized to either an intervention that engaged youth in participatory research and community advocacy projects to prevent tobacco use or to a standard drug prevention curriculum (standard treatment control). Students participated in a day-long advocacy institute and a semester (18 weeks) of youth-led participatory research and advocacy activities. The study was primarily interested in changes in tobacco use, but

strength-based outcomes, such as advocacy self-efficacy, were also incorporated in the theory of change and impact assessment.

Findings: Effects of the three YEP studies that met all exclusion criteria except having been based in formal education are briefly described here. With respect to this review's primary outcomes, Lakin and Mahoney (2006) detected an intervention effect at trend-level statistical significance for general self-efficacy ($p = .09$; standard mean difference (d) = .57), and Winkleby and colleagues (2004) found a highly significant intervention effect on tobacco-related advocacy-specific self-efficacy ($p < .01$; $d = 2$).

Allen and colleagues (1997) did not measure either of the primary outcomes for this review, but the study did find statistically significant intervention effects favoring the intervention group for all three of the study's problem behavior outcomes (secondary outcomes for this review), including academic failure ($p < .001$; odds ratio (OR) = .42), school suspension ($p < .001$; $OR = .39$), and female pregnancy rates ($p < .05$; $OR = .41$). For secondary outcomes, Lakin and Mahoney (2006) found significant intervention effects favoring the intervention group on intent to be involved in future community action ($p = .046$, $d = .67$) and empathy ($p < .01$, $d = 1.03$), but not for sense of civic responsibility ($p = .95$, $d = 0$). Winkleby and colleagues (2004) found favorable intervention effects on community advocacy ($p < .001$, $d = 5.55$) and smoking status for baseline regular smokers ($p < .001$, $d = 3.22$),

but not on smoking status for baseline light smokers ($p = .13$, $d = 1.10$) or non-smokers ($p = .93$, $d = .05$).

As is common with school-based intervention studies, all three of the described excluded studies used designs that randomized either group units (schools or classrooms) or a mix of groups and individuals. Even if the studies met inclusion criteria, including them in meta-analysis would not be possible due to lack of information on design effects in order to appropriately combine individual-level trial data with cluster-level trial data. While the individual studies reported generally positive results, varying degrees of potential methodological problems and the small number of studies warrant a level of caution before interpreting the results too ambitiously. Notably, adequate information on design effects was unavailable from any of the study reports to account for cluster designs and therefore the effect sizes provided above probably represent exaggerated estimates.

Risk of bias in included studies

Allocation. The Morton and Montgomery (2011) and Olson-Merichko (2006) randomized studies reported comparable groups at baseline on demographic and dependent variables, indicating successful randomization. Demographic variables included gender, age, and working status in both studies as well as additional variables (e.g., race, parent's education, and household income) measured by Olson-Merichko (2006), and dependent variables included a wide range of outcome measures. Berg and colleagues (2009), however, conducted a quasi-experimental design, and, even with matching techniques to create two similar groups on

demographic variables, there were significant baseline differences according to some of the intended outcome measure variables. As such, although the quasi-experimental design involved the largest sample size of the three included studies, it also invited the most susceptibility to selection bias. Berg and colleagues attempted to adjust for baseline differences in the analysis by treating the variables as covariates. Despite common practice, however, statistical literature increasingly denounces this practice in non-randomized designs since the variance of an observed dependent variable caused by the covariate is not likely to be independent of the variance caused by the group (Miller & Chapman, 2001).

Blinding. No blinding to trial arm membership was reported for randomization or assessment for any of the included studies for allocation or outcomes assessment. Morton and Montgomery (2011) did utilize computer-based randomization and data collection programs in order to minimize opportunities for bias in the absence of blinding.

Attrition and missing data. Morton and Montgomery (2011) and Olson-Merichko (2006) had relatively low attrition rates (6.3% and 2.5%, respectively). Even though the latter did not incorporate intention-to-treat analysis, the attrition was small enough that this decision was not likely consequential. Berg and colleagues (2006), however, reported relatively high attrition at 26% (author contact). Berg and colleagues did not find any significant differences on demographic variables between completers and non-completers. Nevertheless, high attrition could have had an impact on results, as non-completers could have been

different from completers in terms of response to intervention or other unobserved characteristics. Berg and colleagues' study was not included with the other two studies in this review's meta-analysis.

Selective reporting. There was no evidence of selective reporting of outcomes in the sense that particular measured outcomes were simply not reported. Berg and colleagues (2009) only reported statistics for outcomes with statistically significant intervention effects in the published paper, but the authors did make full data for all outcomes readily available. Olson-Merichko (2006) did not report statistics for antisocial behavior outcomes, which did not show statistically significant intervention effects. Only Morton and Montgomery published a protocol prior to trial commencement prospectively stating outcomes to be measured.

Other potential sources of bias. Only Morton and Montgomery (2011) accounted for 'contamination.' Given the circumstances of the QS NFE target population (out-of-school youths spread across several communities) and that siblings were block-randomized together, notable contamination was unlikely. Documentation of the activities of control group participants confirmed little evidence of contamination effects. Potential for contamination effects was not addressed by Berg and colleagues (2009) nor Olson-Merichko (2006); contamination potential may been a particular concern with the Olson-Merichko study, given that intervention and control participants were all members of the same school during the course of the intervention. Berg and colleagues reported that control participants took part in alternative summer employment programs,

which also *may* have integrated empowerment-based programming components, thus inviting potential contamination effects.

Effects of the interventions

Only Berg and colleagues (2009) measured self-esteem, but, given significant baseline differences, the measure was treated as a covariate in the study's analysis rather than an outcome. Olson-Merichko (2006) and Morton and Montgomery (2011) used the same ten-item measure for general self-efficacy (Schwarzer & Jerusalem, 1995). Berg and colleagues (2009) only included task-specific self-efficacy measures for drug avoidance and sexual behavior. None of the included studies independently showed statistically significant impacts on any of the primary outcome measures for self-esteem or self-efficacy.

General self-efficacy was the only primary outcome measured by more than one included evaluation and therefore the only outcome that was meta-analyzed. Meta-analysis of data from Morton and Montgomery (2011) and Olson-Merichko (2006) did not show a combined intervention effect on self-efficacy (95% CI -0.12 to 0.49; $z = 1.21$). Meta-analysis data and a forest plot using a random-effects model are given in figure 2.2. Given the limited sample sizes and includable studies, the results of this meta-analysis should not be interpreted as an authoritative statement on the effects of YEPs. The results simply reflect a small number of impact studies' aggregate effects on self-efficacy. While some secondary outcome areas were assessed by more than one included study (social supports, social skills, and

problem areas), these were not meta-analyzed because of the level of heterogeneity among constructs captured by the different measures.

Independent results for all of the review's primary and secondary outcomes measured by the included studies are displayed in tables 2.2 and 2.3 (adaptations of table templates used by the Zief et al (2006) review). Berg and colleagues and Olson-Merichko conducted one-tailed analyses to test for intervention effects. Because this review accommodates the possibility of unintended adverse outcomes, significance levels were recalculated using two-tailed tests. Although the tables present relatively little data given the small number of includable studies, future updates of the review could follow the template with additional material. The same outcomes tables including results from the three excluded school-based YEP studies are provided in appendix 2.7.

Figure 2.2. Meta-analysis for general self-efficacy

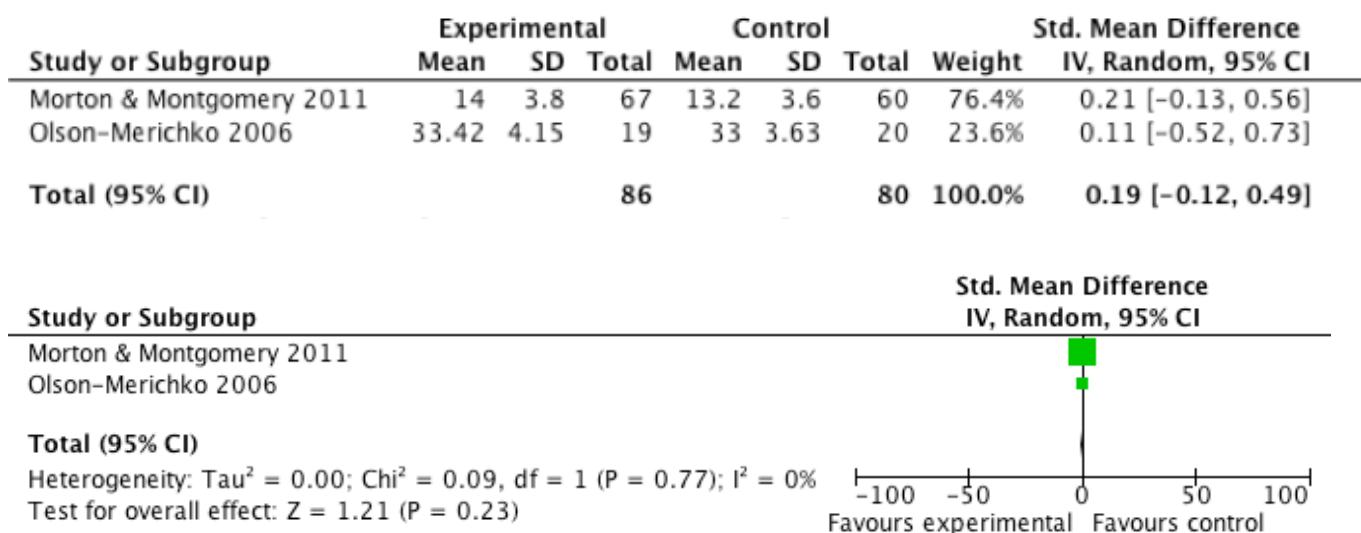


Table 2.2. Primary outcomes from included studies

Study	Outcome	Results						
Outcome	# Studies (Combined N)	# Measured Outcomes	Outcomes favoring intervention		Outcomes favoring comparison		Null effects ($p > .10$)	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
General self-efficacy ^a	2 (n=167)	2	0	0	0	0	2	100
Specific self-efficacy ^b	1 (n=316)	2	0	0	0	0	2	100
Self-esteem ^c	1 (n=316)	1	0	0	0	0	1	100

^aOlson-Merichko (2006) and Morton & Montgomery (2011)
^bBerg et al (2009): drug prevention and sexual behavior
^cBerg et al (2009) (had baseline differences, reported here but otherwise not treated as an outcome in the primary study or by this review)
^{*} $p < .10$, ^{**} $p < .05$, ^{***} $p < .01$ intervention effect

Table 2.3. Secondary outcomes from included studies

Study	Outcome	Results						
Outcome	# Studies (Combined N)	# Measured Outcomes	Outcomes favoring intervention		Outcomes favoring comparison		Null effects ($p > .10$)	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
Social supports and connections ^a	2 (n=443)	5	0	0	0	0	4	80
Social skills ^b	3 (n=483)	5	1	0	0	0	2	66
Emotional intelligence	0	0	0	0	0	0	0	n/a
Coping and problem-solving skills ^c	1 (n=40)	1	1	0	0	0	0	0
Civic engagement	0	0	0	0	0	0	0	n/a
Academic performance	0	0	0	0	0	0	0	n/a
Problem behavior ^d	3 (n=483)	8	3	1	0	0	4	50

^aBerg et al (2009): social connectedness and school bonding (latter had baseline differences, not treated as outcome); Morton & Montgomery (2011): social supports of friends, social supports of family, and adult connectedness

^bBerg et al (2009): social assertiveness skills and social skills; Olson-Merichko (2006): team skills; Morton & Montgomery (2011): social skills and prosocial attitude

^cOlson-Merichko (2006): proactive coping**

^dBerg et al (2009): alcohol use*, marijuana use**, had sex, and number of sex partners**; Olson-Merichko (2006): delinquency, drug use, and alcohol use; Morton & Montgomery (2011): conduct problems**

* $p < .10$, ** $p < .05$, *** $p < .01$ intervention effect

Discussion

Summary of the main results

Due to a small number of includable studies with combinable data, only a very modest meta-analysis was possible for one outcome: general self-efficacy. The meta-analysis did not demonstrate intervention effects for self-efficacy. Despite the considerable amount of literature and institutions promoting the believed impacts of YEPs on positive attitudes and behaviors, this review concludes that there is thus far insufficient empirical evidence to adequately support the claim. There is currently insufficient evidence for reviewers to make conclusions concerning the effects of YEPs.

None of the three included studies independently demonstrated significant intervention effects on this review's primary outcomes, self-efficacy and self-esteem. Many outcomes had null effects, and, for all three included studies, these outnumbered outcomes that did show significant intervention effects. All three included studies assessed self-efficacy, but none showed significant intervention effects compared to control groups at posttest. Differences were insignificant for both general (Morton & Montgomery, 2011; Olson-Merichko, 2006) and task-specific (Berg and colleagues 2009) measures of self-efficacy.

On the other hand, no study showed evidence of harm, nor were there any studies that failed to produce at least one statistically significant, positive intervention effect for the review's secondary outcomes—including social skills (team skills), coping skills (proactive coping), and problem behaviors (conduct

problems, marijuana use, and number of sex partners). Given the large number of outcomes measured by each study, however, a small proportion of outcomes could have shown intervention effects by chance alone (Feise, 2002).

Overall completeness and applicability of evidence

In general, the review demonstrates a paucity of evidence from high-quality impact studies of YEPs outside of formal education contexts. The lack of effects reflected by the three includable studies on the primary outcomes reported by included studies could have been attributable to low attendance rates. In the Questscope non-formal education program in Jordan, for example, 52% of study participants attended less than the minimum amount expected (two days per week). Given that the duration of intervention exposure evaluated by the included studies ranged from 4 to 10 months, longer durations of program exposure could be required to achieve intervention effects with adolescents, particularly those of higher risk. To this end, YEPs that take place in formal education may be advantaged over those outside of formal education by higher overall exposure.

Alternatively, it is possible that YEPs, at least in the forms represented by the included studies, could be an ineffective approach to changing social and emotional youth outcomes hypothesized in theory of change descriptions of youth empowerment.

Quality of the evidence

Only three studies met all of the review's inclusion criteria; these consisted of mixed levels of methodological quality and relatively small sample sizes. The two

randomized controlled trials (Morton & Montgomery, 2011; Olson-Merichko, 2006) both had fairly short posttest periods at 4-months follow-up, and neither had undergone peer-review at the time of this review. The more intensive youth empowerment model that involved longer participatory experiences and the most structured leadership training (Berg et al., 2009) collected data for up to 12 months, but the non-randomized design was more susceptible to bias. Only one study (Morton & Montgomery, 2011) had published a protocol prior to recruitment. Some study reports did not provide data for outcomes that did not show positive, statistically significant intervention effects (most evaluators were willing and able to provide unreported data after contact). Such reporting bias can skew the public's understanding of the full effects of an intervention (Smyth et al., 2011).

Potential biases in the review process

Youth development programs frequently aim to affect changes across multiple behavioral, social, and attitudinal outcomes. As such, the limited emphasis on self-efficacy and self-esteem, though commonplace in the youth empowerment literature, could bias conclusions around a narrow set of measures. Including and reporting a broader set of secondary outcomes mitigated such potential. Additionally, no study was excluded based on its outcomes measures, and the search strategy did not include terms related to outcomes. Consequently, although the review focused its analysis and discussion based on particular outcomes often associated with youth empowerment, the author's decisions on primary and secondary outcomes did not constrain the number of studies that were included in this review. The main reason for the small number of includable studies was a lack

of studies that met both the review's intervention criteria (youth programs outside of formal education that regularly involved participants in program decision-making processes) and study methods criteria (experimental or controlled quasi-experimental designs).

The search for grey literature through professional outreach was limited to the contacts in the author's networks or those that the author could identify. It is possible that other groups could have conducted includable unpublished studies.

Additionally, the inclusion of 'youth empowerment programs' is susceptible to interpretation as to what constitutes youth empowerment. The author tried to establish objective and inclusive intervention inclusion criteria based on regular participation of youths in program decision-making. Moreover, two reviewers independently screened the studies, and there were no disagreements.

Agreements and disagreements with other studies or reviews

This is the first known review of the effects of YEPs. The review findings agree with previously stated expectations in the literature that very little evidence has been generated on youth empowerment interventions through high-quality impact evaluation (Crowley & Skeels, 2010; Gray & Hayes, 2008; Zeldin et al., 2000). This review adds to those statements by identifying the few studies that do exist and outlining directions for future research.

Conclusions

Implications for practice

The review demonstrates an insufficient evidence-base for YEPs' impact on self-efficacy and self-esteem. As such, the review is unable to offer definitive conclusions about the impacts of YEPs. While the few includable studies do not show positive intervention effects on these primary outcomes, there is also no evidence of harm from reported outcomes data. There is limited evidence for intervention effects on the review's secondary outcomes (e.g., social skills and antisocial behavior) suggesting a potentially important role for youth empowerment in changing these outcomes, but further research is needed.

As expected, there was heterogeneity between YEPs evaluated by the included studies with respect to program activities, the extent of youth participation in program decision-making, the nature of adults' roles in the program, the amount of program exposure, and the characteristics of the study sample. As required, all three programs shared a commitment to regular involvement of young people in programming decision-making as an aspect of their intervention experience, samples of similar age groups, a supportive adult presence, and asset-building activities intended to build on young people's strengths.

The reviewers could not identify sufficient data to assess the extent to which differences or similarities between YEPs help to explain outcomes. It remains to be seen, for example, whether different degrees of youth participation in program leadership and decision-making explain different levels of program effects on youth

development outcomes, and for different populations of youth. Youth empowerment has varying intervention implications across different cultural contexts; these should be explored in future studies.

Although this review concentrated on YEPs outside of formal education, it also identified three studies of YEPs that met all of the review's criteria except for having been based in formal education. In general, research interest in schools as contexts for youth development and engagement has gained increasing traction (Eccles & Roeser, 2011; Shinn & Yoshikawa, 2008), and these studies reflect that interest. Notably, despite methodological concerns, results from the few randomized studies of YEPs based in formal education present a generally more positive picture of program effects than do the included studies of YEPs outside of formal education. The evidence is too little and problematic to be generalizable, but it does suggest promise for schools as settings to improve developmental and problem-behavior outcomes through youth empowerment—particularly those with service or advocacy-based curricula.

It may be that more favorable intervention effects in the formal education YEPs were due to methodological differences or more sophisticated program designs. Alternatively, programs based in schools may have advantages of increased overall attendance and the ability to capitalize on setting-level process factors specific to schools, such as improved connections with school staff and positive program “contamination” among peers within the school environment. Out-of-school YEPs included in this review also targeted more disadvantaged populations,

including those not attending school, presenting additional challenges for producing high attendance and intervention effects.

Implications for research

Impact study. Though still very limited, the concentration of studies over the last six years is a promising sign for impact evaluation of YEPs. Five out of six of the included studies and relevant excluded studies were published since 2004; one was published in 1997. This relatively recent production of experimental and quasi-experimental evaluations of YEPs suggests some momentum for interest in youth empowerment as a modality of intervention as well as understanding its effects. As such, future updates of this review may yield more substantive syntheses of the evidence for YEPs if the primary research trend continues. For example, a large, multisite RCT (anticipated sample of about 3,400 participants) of the YouthBuild program in the US, which involves empowerment-based methodologies, is currently underway and may add significantly to the evidence-base on YEPs (MDRC, 2011).

Nevertheless, the small number of includable studies highlights an ongoing tension between demands for high-quality evidence from scientifically rigorous impact evaluation designs and the difficulties of applying experimental designs to participatory programs. The use of RCTs, for example, is unrealistic when participants *initiate* empowerment programs, in which case the same participants could not be randomly assigned into or out of an intervention they created. In situations where experimental or controlled quasi-experimental designs are not feasible or appropriate, programs may be left to rely on less scientifically robust

methods for impact evaluation, such as cohort studies, life histories, and ethnographies, to discern evidence about program effects, but higher susceptibility to bias is a major concern for confiding in their findings on intervention effects.

The prevalence of self-efficacy measures among the few includable studies corroborates the attention to the outcome in the broader youth empowerment literature. Self-esteem was only measured in one study, but it was not included as an outcome in the study's final analysis due to significant group differences at baseline (Berg et al., 2009). The greater attention to self-efficacy compared to self-esteem could reflect clearer congruence between self-efficacy as a motivational construct and the personal agency emphasized by empowerment theory (Chen et al., 2004).

Increased impact evaluation of YEPs could be supported by more research investment in measures that are consistent with YEPs' general theory of change and have been sufficiently tested for the range of cultural contexts to which researchers plan to apply the measures. For example, despite frequent reference to the role of improved trust-based relationships between youth and community adults in the youth empowerment process, both Morton and Montgomery (2011) and Berg and colleagues (2009) were unable to identify adequately tested and sensitive measures to assess young people's sense of connectedness to community adults.

The baseline differences on several outcome measure variables in Berg and colleagues' (2009) quasi-experimental study reinforce the difficulties of establishing two comparable groups without random assignment, even when matching procedures based on basic demographic variables are used. When practicable and

appropriate, future research into the efficacy and effectiveness of YEPs should strive to implement randomized designs, which are best equipped to avoid selection biases that can produce misleading results (Craig et al., 2008; Glazerman et al., 2003; Jadad & Enkin, 2007).

Where sample size calculation was conducted (Morton & Montgomery, 2011), the study was not powered for subgroup analysis. As such, the studies could not authoritatively assess the extent to which demographic characteristics (e.g., age, gender, family income) or process characteristics (e.g., attendance, support, or empowerment) moderated or mediated intervention effects. Though statistical power was limited, Morton and Montgomery (2011) found that the program center in which youth rated the highest empowerment outperformed the center with the lowest youth scoring on empowerment on nine out of eleven measured outcomes. Potential interactions between young people's experiences of empowerment in program processes and program effects should be investigated further with better process measures and larger sample sizes in future research.

Finally, follow-up data collection for included studies ranged from 4 to 12 months following beginning of program participation. Future studies should follow participants for longer time periods (over a year), as some outcomes may require longer program exposure to reveal measureable change, or, alternatively, changes that do occur in outcomes in the first few months may not be sustained.

Process study. Some researchers have challenged the applicability of experimental studies to empowerment-based interventions given the complex

processes underlying their potential impacts (Wallerstein, 2006). While this review assumes the importance of using experimental designs to increase the evidence-base for the efficacy and effectiveness of complex social interventions, including YEPs, complementary methods are indeed important to help isolate the particular components, interactions, and processes working inside of such interventions to facilitate or stifle outcomes of interest (Craig et al., 2008; Oakley et al., 2006; Raudebush et al., 2008).

All of the included studies integrated some level of process study in tandem with experimental and quasi-experimental designs. Each of these involved a qualitative dimension, such as qualitative activity observations or interviews with youth participants and/or adult staff. In each case, such qualitative investigation enabled researchers to tease out more nuanced insights into program implementation that might not have been revealed otherwise. Observations in Berg and colleagues (2009), for example, indicated that reflective activities helped youth engage with their action research projects more meaningfully when reflections facilitated opportunities for youth to personalize the experiences; in qualitative interviews associated with Morton and Montgomery's (2011) study, youth described the most important attributes of effective adult facilitators (e.g., making sure that *all* youths' voices get heard in decision-making processes) (Morton, 2011); and youth feedback in Olson-Merichko (2006) suggested that future programming could be improved by extending the duration of projects and shifting meetings to school hours when students found it easier to meet.

Of the three included studies, only Morton and Montgomery (2011) used a quantitative process study component—a modified version of the Learner Empowerment Survey, which was also used in the discussed excluded study by Lakin and Mahoney (2006). The Learner Empowerment Survey (Fymier et al., 1996) was initially created for college communication classes. Theory-driven instruments developed more specifically for YEPs for adolescents could provide more useful information. Efforts have been made in recent years toward better general youth program quality instruments, such as the High/Scope Youth Program Quality Assessment (Blazevski & Smith, 2007) and the READY Tool (Sabaratna & Klein, 2006). Given the continued development and evaluation of YEPs, specific attention should be given to how such instruments validly and reliably assess empowerment constructs concerning the contribution, participation, preparation, and support of young people in community programs.

Summary conclusion

This review reveals an insufficient evidence-base from experimental or quasi-experimental studies to substantiate the hypothesis that YEPs have an impact on developmental assets such as self-efficacy and self-esteem. More research into YEPs using rigorous impact study designs is needed. Researchers should further develop methods and measures to enable high-quality, mixed-methods process studies to complement impact studies of YEPs so as to provide more useful evidence for practitioners and policy-makers. Given the relative nascency of impact evaluation in the YEP field, this review's findings should be interpreted as a stimulus

for further research investment and action rather than a basis for generalizable conclusions about the effects of youth empowerment.

References for included studies

- Berg, M., Coman, E., & Schensul, J. (2009). Youth Action Research for Prevention: A Multi-level intervention designed to increase efficacy and empowerment among urban youth. *American Journal of Community Psychology*, 43(3), 345-359.
- Morton, M., & Montgomery, P. (2011). Empowerment-based non-formal education for out-of-school youth in Jordan: A pilot randomized controlled trial. *In submission*.
- Olson-Merichko, J. A. (2006). *Youth leadership program evaluation: Implications for implementation*. Unpublished Dissertation, Indiana University of Pennsylvania, Indiana, PA.

Chapter 3: Context for Primary Research

Introduction

The preceding chapter highlighted the general paucity of high-quality impact evaluation of youth empowerment programs (YEPs), especially outside of the United States. Accordingly, the chapters to come will contribute to the evidence-base concerning youth empowerment in a non-Western cultural context by reporting impact and implementation research on a particular YEP, Questscope non-formal education (QS NFE), for out-of-school youth in Jordan. This chapter reviews a wide range of literature to provide context for such programming in the Hashemite Kingdom of Jordan as well as the Middle East and North Africa (MENA). Descriptive research and reliable data systems are notoriously lacking in Jordan as with much of the developing world. Nevertheless, even the limited available information helps set the backdrop for why effective youth interventions are needed for out-of-school Arab youth, as well as for potential political, economic, and cultural challenges that could confront the application of youth empowerment programming in the region.

This chapter begins by providing a general introduction to Jordan's demographic, political, economic, and civil society circumstances. Youth programming does not operate in isolation from these dimensions of context. Demographic trends, for example, can inform the social-cultural undertones that influence applicability of and responsiveness to an intervention. Economic circumstances can help demonstrate need for effective youth interventions, but they can also contextualize the difficulties of funding them. Political and civil society

trends offer insights into whether, and how, important institutions might support or hinder participatory community-based efforts like YEPs.

The chapter then goes on to focus on context more specific to YEPs. This includes the challenges and opportunities for intervention and policy with Arab youth, and the educational and vocational trends that have led to the emergence of non-formal education as a vehicle for youth empowerment with vulnerable young people in Jordan. The chapter considers youth empowerment in light of cultural and historical nuances within Arab society, which could welcome empowerment-based programming in some respects but could also present significant challenges for diffusing a methodology that has been largely developed and researched in Western societies. Finally, the chapter examines the limited data available about the characteristics of out-of-school youth, the population targeted by the YEP under study in this dissertation.

This background investigation can contribute to better theoretical development and cultural feasibility assessment for the randomized controlled trial described in the next two chapters, as encouraged by the MRC Framework for Complex Interventions (MRC, 2008). Imas and Rist (2009, p. 142) suggest beginning with “front-end analysis” in the early stages of evaluation, which comprises “an investigation of an issue or problem to determine what is known about it and how to proceed in developing an evaluative approach to it.” Several aspects of a front-end analysis (e.g., identifying relevant theory and evidence-base) were largely addressed in previous chapters, but this chapter focuses particularly on understanding issues and problems related to youth (and especially out-of-school

youth) in the Jordanian context.

According to Imas and Rist (2009, p. 151), a review of the literature can help contribute to better understanding of the policy and environmental context that in turn enables identification of critical assumptions. Similarly, Weiss (1995) reasoned that a good theory of change identifies the assumptions embedded in the hypotheses about how an intervention will affect specific outcomes. Better understanding of the multifaceted context of a problem and intervention helps to ascertain such assumptions that could influence a program or program evaluation's design, implementation, or results.

Country profile

Demographics

The Hashemite Kingdom of Jordan (Jordan, for short) is both strategically and precariously situated as a geopolitical buffer between the West and other Arab nations (see figure 3.1 map). At 34,445 square miles, Jordan borders Syria, Iraq, Saudi Arabia, and shares the longest border with Israel. The Jordan population estimate for 2011 is 6.26 million (IMF, 2011), which primarily ethnically consists of Arabs (98%) with the remaining mostly including Circassians (1%) and Armenians (1%) (Sharp, 2010a). The country's religious composition includes 92% Sunni Muslim, 6% Christian, and 2% small Muslim sects (Sharp, 2010a). As reflected by the population chart in figure 3.2, Jordan has experienced tremendous population growth.

With the establishment of the state of Israel in 1948 and subsequent events such as the 1967 war in which Israel conquered the West Bank and Gaza Strip, large groups of Palestinian refugees fled to Jordan, resulting in a substantial Palestinian basis to the country's demographic (UNHCR, 2004). The USA's Congressional Research Service estimates that Jordanians of Palestinian origin comprise 55% to 70% of the population (Sharp, 2010a). Just in the last twenty years (since 1990), the country's population has nearly doubled. Growth rates are attributable to increased fertility and life expectancy as well as repatriation and immigration flows driven by discrete geopolitical events. Growth rate spikes responsive to events can be seen in the chart presented in figure 3.3, which shows especially high growth rates in the

1960s (leading up to the war of 1967 with vast numbers of Palestinians coming in from Israel) and early 1990s around the Gulf War of 1990-1991. The ramifications of the first Gulf War included the arrival of over 300,000 repatriated Jordanian nationals and even larger numbers of refugees fleeing from Kuwait. More recently, population trends have been compounded by an estimated 300,000 Iraqi refugees fleeing violence to Jordan following the US-led invasion that began 2003 (USCRI, 2004).

These growth spikes have constituted significant demographic shocks making rapid contributions to urbanized poverty. The inflows of Palestinian and Iraqi refugees and displaced persons have presented particularly challenging demographic burdens for the country to address. Among other problems, these groups experience considerable legal and social obstacles to accessing adequate education and employment. Moreover, pressures from rapid demographic growth have been compounded by Jordan's increasing water crisis. Jordan is one of the world's ten most water deprived countries, and recent droughts have worsened the situation (Sharp, 2010a). Severe water shortages stand to threaten Jordan's already-fragile social and economic stability.

Figure 3.1. Jordan map

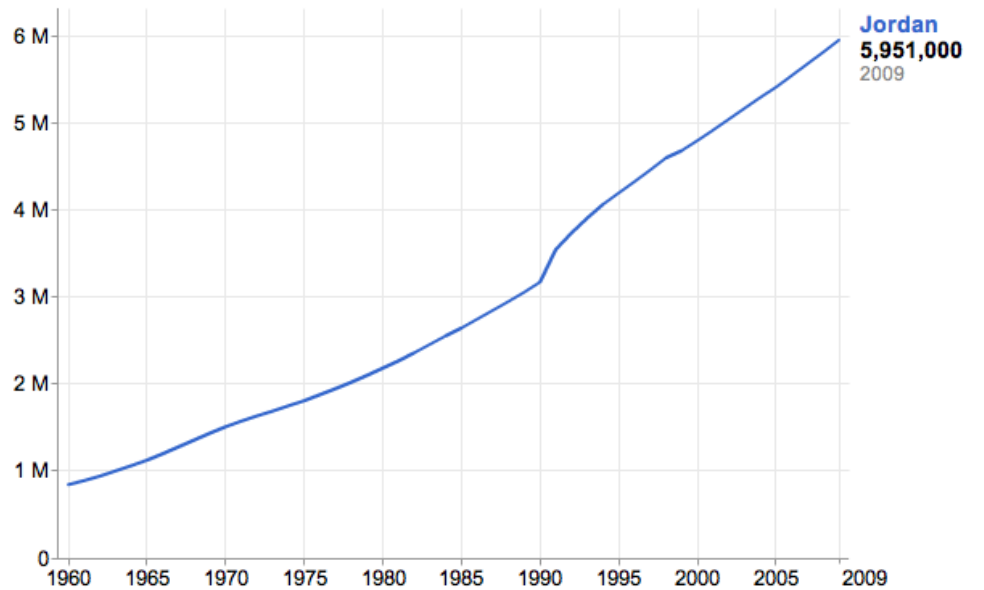
Figure 3.2. Jordan population trends

Figure 3.3. Jordan growth rate trends



Population

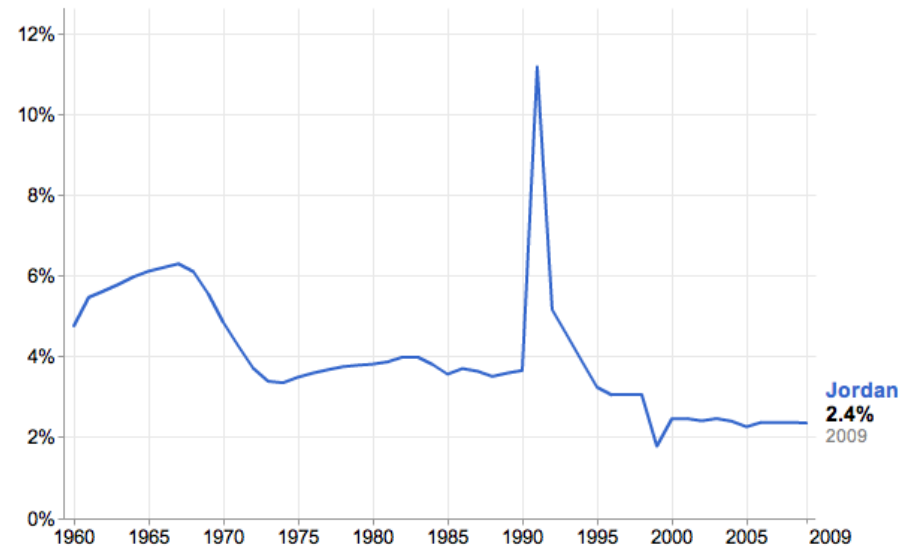
Midyear estimates of the resident population.



Data source: [World Bank, World Development Indicators](#) - Last updated Apr 26, 2011

Population growth rate

Percentage change of resident population compared to previous year.



Data source: [World Bank, World Development Indicators](#) - Last updated Apr 26, 2011

Urbanization

Jordan's population growth has been largely urbanized, especially with growth concentrated in the greater Amman area (Kanaan & Hanania, 2009). Approximately 70% of the population lives in urban areas (USDOS, 2011); approximately 82% of adolescents (aged 10 to 19) reside in urban areas (Population Council, 2009). Forty-two percent (over 2.5 million in 2009) of the population lives in the capital city of Amman (USDOS, 2011), which comprised "little more than a collection of dwellings and some 2,000 to 3,000 inhabitants" as early as the 1920s (Potter et al., 2009).

Urbanization patterns have further entrenched social disparities often encapsulated by Jordanians' reference to "two Ammans"—Western Amman, with neighborhoods of wealth and upscale commercial areas, and Eastern Amman, comprised of the urbanized poor, including low-income, generally more religiously conservative Jordanian communities and Palestinian refugee camps (Potter et al., 2009). Although these area discrepancies have been mapped by Potter and colleagues and are commonly acknowledged in Amman, to date there is a lack of publically-available data to confirm and deepen understanding of the disparities between "East Amman" and "West Amman." Nevertheless, the research presented in chapters four through six of this dissertation focuses on programming set within disadvantaged areas of Amman as shown by the limited available data.

Politics

Jordan's political system is defined as a constitutional monarchy (USDOS, 2011). The King is the head of state and has sweeping political powers, including

appointment of the Prime Minister who serves as head of government. The national bicameral legislature is determined by elections—the last of which took place in November 2010 after the King dissolved the government. The regime has recurrently dissolved the government for the stated purposes of advancing political reform (Sharp, 2010a). As George's (2005) qualitative investigation into contemporary Jordanian politics illustrates, Jordan's situation is characterized by dueling forces promoting progressive political reforms and freedoms on one hand and corruption and constraints on press and political openness on the other.

Despite the government's intensive efforts to advance reforms, difficulties remain. Corruption is seen as a particularly acute problem in Jordan, with 40.7% of businesses surveyed in the World Bank Investment Climate Survey describing corruption as a major constraint to doing business in Jordan (World Bank, 2006a), and the strong cultural tendency for job placements, among other things, to be based on connections (*wasta*) rather than merit thwarts social and economic mobility and fosters adverse incentive structures (Sharp, 2010a). El-Said and Harrigan (2009) argue that, in good times, *wasta*, an enduring cultural artifact of Jordan's Bedouin tribal roots, can provide a positive force for 'bridging' social capital expanding social networks in Jordan, but, in bad times (e.g., wars and civil wars), *wasta* is used as a harmful tool that encloses 'bonding' social networks and reinforces biased and corrupt public policy.

Moreover, notwithstanding efforts toward political liberalization, a 2009 national opinion poll conducted by the University of Jordan's Center for Strategic Studies found that 74% of people fear criticizing the government openly (Open

Source Center, 2009) and *The Economist* Intelligence Unit's Democracy Index ranks Jordan 117 out of 167 countries (The Economist, 2011). Protest activity in Jordan, especially among young people, against corruption and for political reforms has recently increased following revolutionary successes in Tunisia and Egypt (BBC, 2011; Leigh, 2011). The regime has taken additional steps to advance political reforms and hold high-profile public officials accountable for acts of corruption (Sharp, 2010a), but the state of affairs remains very much in flux. Even with internal and external unrest, however, Jordan is considered to be relatively stable and potentially less susceptible to the scale of regime change observed as-of-late with other Arab countries due largely to political agility and stature maintained by the monarchy (Hamid, 2011b).

Economy

The size of Jordan's nominal GDP was estimated at \$27.5 billion in 2010 (IMF, 2011). Marking a deepening financial predicament, Jordan routinely runs budget deficits and public debt was estimated at \$13.6 billion in 2009, nearly 64% of GDP (Sharp, 2010a). With the exception of large phosphates and potash reserves, Jordan lacks the natural resources—principally oil—that drive economic growth and independence for other Arab countries. Instead, Jordan's economy has largely relied on tourism and external sources of growth, such as remittances and foreign aid (Sharp, 2010a). The foremost economic challenge for Jordan has been described as its need to reduce these dependencies (Kanaan & Hanania, 2009).

Jordanians working abroad (an estimated 10% of Jordan's population), particularly in Arab Gulf countries rich in oil resources, have constituted a major

source of the Jordanian economy. An estimated 20% of Jordan's GDP comes from remittances (Sharp, 2010a). According to the USA's Congressional Research Service, total American aid to Jordan from 1951 through 2010 totaled approximately \$11.4 billion (Sharp, 2010b). Under a joint 2008 Memorandum of Understanding (MOU) between the American and Jordanian governments, the USA committed to \$660 million in annual foreign assistance (\$360 million in Economic Support Funds (ESF) and \$300 million in Foreign Military Financing (FMF)) for a five-year period. Jordan was the fourth largest recipient of American foreign aid in 2008 (after Israel, Afghanistan, and Egypt; Tarnoff & Lawson, 2009). Jordan's economy also depends on tourism, the health of which precariously oscillates with geopolitical events such as war and terrorism in the broader region (Sharp, 2010a).

In order to move past external dependencies, Jordan is working to attract private investment and develop its human capital as a driver for economic growth. This helps explain the substantial investments the Jordanian government makes in education and social policy, as described later in this chapter. Nevertheless, these efforts have considerable challenges to overcome. The Jordanian government is the country's largest employer by far, with estimates ranging from one-third to two-thirds of all workers, and the country still experiences significant "brain drain" with many of its most skilled and educated workers pursuing better opportunities abroad (Sharp, 2010a).

The official poverty rate in Jordan was 14% in 2004 (based on a JD 392, or \$554 USD, per capita per year income poverty threshold) a decline from 21% in 1997 (UNICEF, 2007; World Bank, 2004). Chronically poor make up 29% of those

below the poverty threshold (UNICEF, 2007). Despite the impressive poverty reduction, however, a World Bank (2004) report questions its sustainability without stronger economic growth, better social and economic policy targeting the poor, and dealing with persistent unemployment.

Civil society

In many respects, Jordan is well-positioned for piloting innovative social interventions and high-quality evaluation in the Arab region. While a developing context, Jordan has consistently performed above the Arab states and world averages for the Human Development Index, which includes indicators related to health, education, and income (UNDP, 2010). Jordan ranked 89 out of 169 countries included in the index in 2010, placing Jordan at the low end of what the index considers “high human development.” The country’s relative stability and faithfulness to the rule of law foster a more conducive environment for social development initiatives than do many Arab countries characterized by greater political and institutional volatility. Consequently, the number of non-governmental organizations (NGOs) in Jordan has nearly doubled since political liberalization reforms in 1989 (Wiktorowicz, 2002). A 2009-2010 mapping study by the European Union estimated that there are approximately 3,000 civil society organizations in Jordan (including grassroots community-based organizations (CBOs), non-governmental organizations (NGOs), umbrella organizations, and thematic alliances and coalitions) (Williamson & Hakki, 2010).

Despite the conditions that have encouraged the growth of NGOs, however, Wiktorowicz (2002) argues that specific government strategies have also stifled the

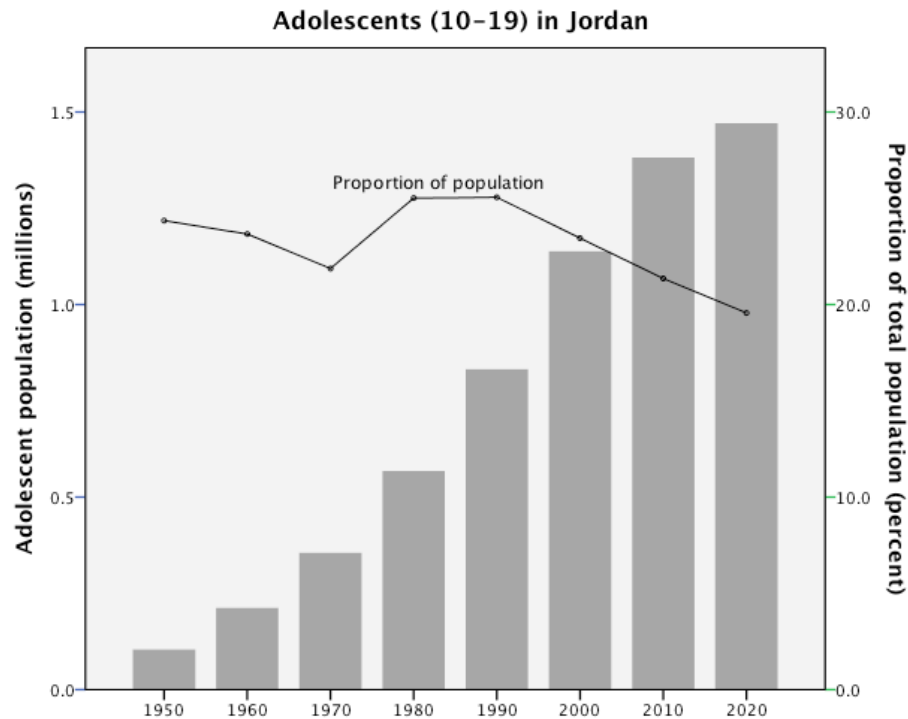
ability of NGOs in Jordan to genuinely advance democratization and civil society development. Such strategies to control the NGO community, according to Wiktorwicz (2002), include: “(a) administrative repression and oversight; (b) civil society ‘infiltration’ through royal nongovernmental organizations and other government NGOs; and (c) centralization through the General Union of Voluntary Societies.” The trends present tensions between Jordan’s advances in civil society development that could facilitate various levels of youth empowerment in political and social life, which is frequently promoted by the government and royal family (HCY, 2004), and nervous maneuvers applied by the regime that could constrain the spread of empowerment initiatives.

Youth challenges and opportunities in Jordan

Youth population growth and social exclusion

Promoting stronger evidence for youth intervention has special urgency in MENA. The demographic phenomenon of a highly dense youth population (commonly referred to as the ‘youth bulge’) in MENA is commonly considered one of the most pressing challenges facing the region (Fuller, 2003; Salehi-Isfahani & Dhillon, 2008; The Economist, 2009). This demographic challenge is visibly pertinent in Jordan. Young people aged 15 to 29 currently make up the largest demographic group in Jordan with more than 29% of the population (Kanaan & Hanania, 2009). Population trends for Jordanian adolescents are included in figure 3.4 (data were drawn from the United Nations Population Division’s World Population Database (UNPD, 2009)).

Figure 3.4. Adolescent share of Jordan population over time



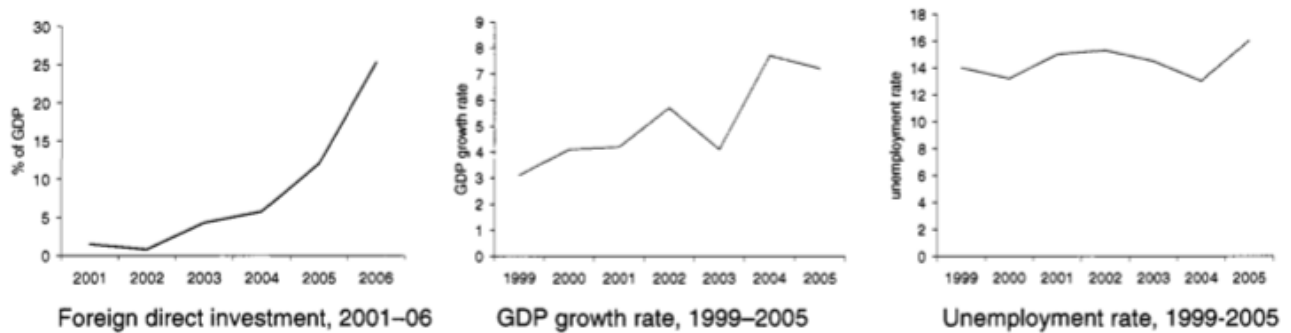
The coupling of severe demographic growth trends with economic volatility and dependence has presented looming challenges for Jordan’s young people. While the official unemployment rate stands at approximately 14%, youth unemployment remains at 30%. According to Kanaan and Hanania (2009, p. 145), as huge swaths of young people enter the labor market, “the government estimates that 50,000 new jobs must be created annually just to maintain the current employment rate.”

Considered in this light, the Middle East’s youth demographic stretches public resources and reinforces the social and economic fragility of the region.

Paradoxically, even with increased foreign investment and GDP growth in Jordan, unemployment has risen (see figure 3.5) (World Bank, 2008). Economists largely attribute this contrast to the influx of migrant workers taking service jobs that

Jordanians are less willing to pursue and to failure of Jordanian educational and vocational systems to prepare Jordanians with the skills and competencies demanded by the labor market (World Bank, 2008).

Figure 3.5. Paradox of increased unemployment with investment and growth



Source: World Bank (2008)

At the same time, however, others have described the disproportionately young society as a 'demographic gift' (Dhillon, 2008; Fuller, 2003, p. 6). This optimistic outlook holds that a larger youth population aged 15 and older translates to a lower proportion of dependency-age people (younger than 15 or older than 64) who can potentially contribute to the country's labor force and economic growth (Chaaban, 2009). If met with adequate investments in the right kinds of education and diverse private sector jobs, Brookings Institute papers suggest that the Middle East youth concentration could be a driving force behind economic prosperity and positive social change (Dhillon, 2008; Fuller, 2003).

Failure to capitalize on this opportunity with effective investments in youth will continue to render considerable youth exclusion that bears costly implications. In fact, Chaaban (2008) developed economic models to calculate the costs of several

aspects of youth exclusion—youth joblessness, school dropouts, adolescent pregnancy, and youth migration—and estimated that the total economic costs of youth exclusion amounted to approximately \$1.5 billion USD (7.3% of GDP) for Jordan in 2006, and up to \$53 billion USD (17.4% of GDP) for larger countries in the region such as Egypt. Such models involve considerable assumptions, but the estimates nevertheless provide perspective into the magnitude of the fiscal dimension of the problem. Of particular relevance for the present research are the estimates for school dropouts, including an estimated economic cost of 1.5% of Jordan's GDP (Chaaban, 2008). Discounted lifetime earnings for primary school leavers in Jordan amount to \$18,217 USD according to Chaaban. That is \$4,365 USD less than those with secondary education and \$22,369 USD less than those with higher education. As such, youth who drop out from school are at significant disadvantage for reaching quality living standards.

Education trends

In many respects, Jordan has made noteworthy educational progress. Literacy and education in Jordan have steadily risen over the last several decades. In 1999, 98.2% of 15 to 19-year-olds were literate compared to 54.7% in 1961 (Al-Khalidi, 2006). Illiteracy rates for the whole population dropped from 33.5% in 1979 to 10.3% in 2005. During the same time period, the percentage of total population with formal education rose from 52.1% to 85.3% (Al-Khalidi, 2006). Particular gains were made over time with school enrolment of Jordanian youth; 81.3% of 6 to 24-year-olds were enrolled in school in 1998 compared to 34% in 1961 (Al-Khalidi, 2006). While the status of gender equality remains problematic (especially with

discrimination and widening gender gaps in the labor market; Kanaan & Hanania, 2009; Miles, 2002), Al-Khaldi also shows trends of narrowing grade school education gaps in Jordan between males and females over time, and Jansen (2006) highlights increased participation in higher-education among Jordanian females. In fact, Jordan has one of the region's highest female literacy rates at 82% (Eyango & AlShaer, 2003).

Despite these gains, however, Al-Khaldi (2006) juxtaposes illiteracy rates in Jordan around 10% in 2004 (still good relative to other Arab countries) to near universal primary school *enrollment*; given high demand for skilled labor indicated in his analysis, Al-Khadi's suggests that reduction of the dropout ratio in Jordanians' education, along with improved rural educational access, should be the country's chief priority for improving its labor force. Indeed, although positive progress has been made on school enrollment, school attendance is still a problem. Survey data collected in 2007 by the Population Council (2009) showed that 6.8% of girls and 8.8% of boys between the ages of 6 and 17 (school age according to UNESCO) were not in school. Looking at a longitudinal rather than snapshot perspective, a series of cohort studies tracking 722,883 children indicated that 24% of Jordanian youths that entered first grade did not graduate from twelfth grade (high school) (NCFA, 2004).

This out-of-school youth population is particularly concerning for Jordan given the country's high youth unemployment rates. Data analysis by Al-Khaldi (2006) shows that educational attainment is one of the most important factors affecting labor force participation in Jordan over recent decades, especially for

females. To this end, prospects for developing economies to integrate productively into the global economy can be strongly diminished by inadequate engagement and preparation of youth through education. As a Brookings Institute report (Fuller, 2003, p. 15) argues: “Lower levels of education and skills immediately translate into a less productive workforce that impacts negatively upon the economy and offers little attraction to potential foreign investors.”

Reasons for early school leaving

More research is needed into specific reasons for why young people drop out of school in Jordan, but available data gives useful insights. For many females, family pressures—especially early marriage—can pose particularly important challenges for school attendance. Among 15 to 19-year-old females in Jordan who have ever been married, only 9.3% are enrolled in school; this compares to 86.7% for those who have never been married (Population Council, 2009). This problem is underscored by the fact that more than 26% of females between the ages of 20 and 24 report having been married by age 18 (3% report having been married by age 15) (Population Council, 2009). However, males are still more likely than females to stop attending school in Jordan (Nimeh & Bauchmuller, 2010).

Males and females are more likely to leave school prematurely when household economic conditions drive them toward labor. One in eight youth reported feelings that the cost of education would prevent them from continuing their studies at higher levels (UNICEF, 2002). Indeed, analysis of national household survey data in Jordan has shown a high incidence of child labor in the household to be negatively correlated with school attendance (Nimeh & Bauchmuller, 2010).

Pressures to work can be particularly strong for young males. A 2001 national UNICEF survey of 7,421 youth, ages 10 to 24, found 35% of young males work compared to 4% of young females (UNICEF, 2002). A separate survey found that 40% of working children are below age 15 and 60% are aged 15 to 18 (UNICEF, 2007).

Finally, inadequate quality of education may be the most prominent contributor to early school leaving. In a 2003 countrywide household survey ($n = 15,376$), out-of-school youth in Jordan, ages 6 to 15, ranked *not liking school* as the foremost reason for not attending, even above work, poverty, and family reasons (Nimeh & Bauchmuller, 2010). This was true for males and females alike. Quality of education is frequently described as a principal concern for Jordan's educational progress. A country educational assessment by USAID explained that "[r]elatively high enrollment rates mask some of the prevailing problems such as the rote memorization methods that entail no critical thinking" (Eyango & AlShaer, 2003, p. 8). Similar concerns have been echoed by multiple other institutions, such as UNICEF and the Brookings Institute, assessing the state of education in Jordan (Kanaan & Hanania, 2009; UNICEF, 2007). Likewise, assessments in nearby Arab countries such as Egypt have found poor educational quality to constitute the most important factor behind early school leaving (Azer, 1993).

The incentive structures of education systems in the Middle East are known for placing a huge emphasis on standardized test preparation and rote memorization at the expense of broad skill sets and competencies that actually prepare young people for competing in the economic marketplace and contributing

to society (Salehi-Isfahani & Dhillon, 2008, p. 14-15). Not only does this approach fail to prepare young people for positive transitions to adulthood, but in many cases it also fails to captivate students. In Jordan, performance on the national qualifying standardized examination for higher education, the *tawjihi*, essentially constitutes the singular determinant of young people's eligibility to attend public university or community college as well as the profession for which they can receive higher educational training (Kanaan et al., 2009). As such, the formal education system in Jordan notoriously teaches students almost exclusively toward this test at the expense of engaging young people, fostering an enthusiasm for learning, or helping youths exercise critical thought and intellectual agility. This challenge provides important context for the intervention described in subsequent chapters. By fostering relevance, ownership, and greater support, youth empowerment may constitute a critical strategy for re-engaging marginalized youths and preparing them for productive transitions.

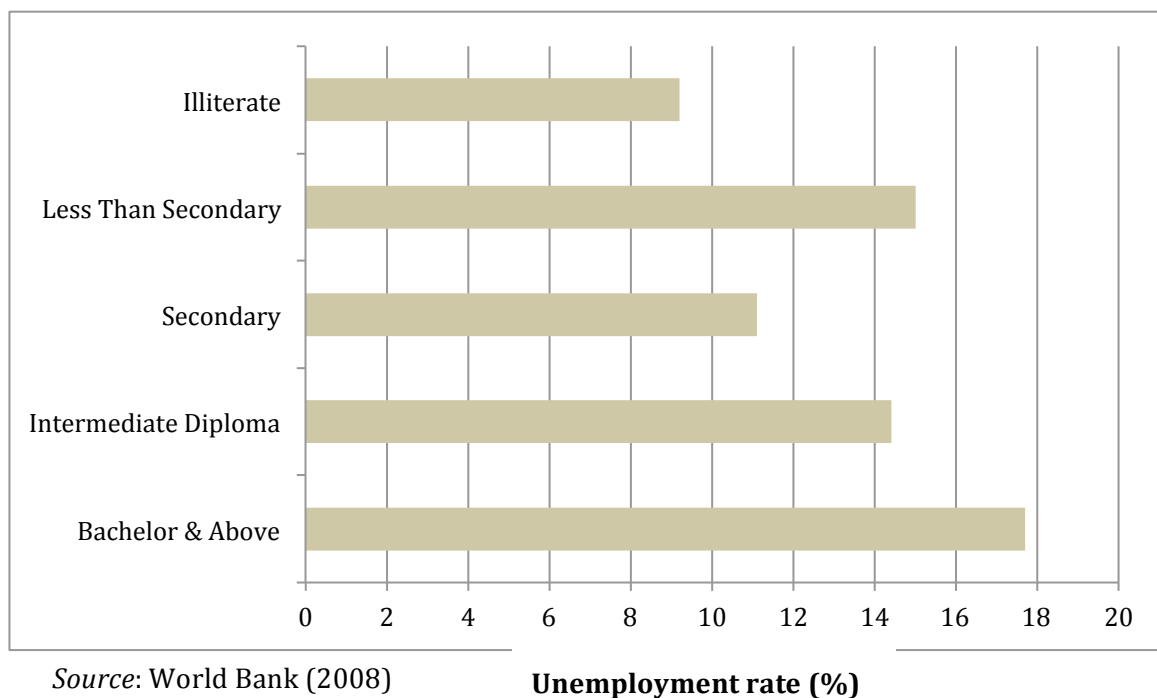
Reasons for youth unemployment

In addition to retention challenges, another significant problem related to educational quality involves Jordan's frequently cited shortcomings in preparing young people for workforce needs (Eyango & AlShaer, 2003). The attainment of education by itself is not a guarantee for success in the labor market. In fact, 2005 census data exhibited in figure 3.6 showed that the most educated segment of Jordanian society has the highest unemployment rate. The hypothesized reasons for this situation are many and complex, but some relevant themes have emerged from the available evidence. Many of the new jobs created in Jordan have been taken by

foreign workers (World Bank, 2008). The level of education of foreign workers is no higher than that of unemployed Jordanians, reinforcing that education level alone cannot explain the problem (World Bank, 2008).

From a survey of unemployed Jordanian workers, 54% of respondents were considered “voluntarily unemployed” (World Bank, 2008). In other words, for this segment of the population, the unemployed desire to work but currently available occupations at prevailing wages from a list provided in the survey were perceived as not worth taking. For this group, the labor market fails to supply jobs that meet job-seeker demand. For the other 46%, however, unemployed Jordanian workers were willing to accept at least one occupation from a list of available jobs, and this “involuntarily unemployed” group reflects a situation in which the Jordanian unemployed fail to supply the skills that the meet labor market demands.

Figure 3.6. Unemployment by education level (aged 15-29), 2005



According to a 2006 survey conducted by the University of Jordan's Center for Strategic Studies, "61% of employers believed that Jordanian workers were less valuable than foreign workers because they put in less effort, took less responsibility, worked less, and were absent more than foreign workers" (reported in World Bank, 2008, p. 16). A World Bank report suggests that while Jordanians are not disadvantaged in terms of quantity of education compared to foreign workers, they can be in terms of quality of education, which is seen as devoid of teaching initiative, critical thinking, and specialized skills valued in the labor market (World Bank, 2008).

This dilemma reinforces arguments by Heckman (2000) that non-cognitive skills (e.g., motivation and social skills) can be even more important than cognitive skills (e.g., academic IQ or standardized test ability) in predicting young people's long-term employment and wellbeing. Indeed, remedial education interventions such as the General Education Development (GED) program in the USA are costly and have failed to show intervention effects on income or employment outcomes (Cunha & Heckman, 2006). As such, unlike such remedial education programs, effective programs aimed at out-of-school youth may need to prioritize methodologies that augment young people's non-cognitive, developmental assets, rather than focusing merely on a general degree and basic cognitive skills.

As such, an educational system that fails to prepare young people in developmental assets and vocational competencies may forfeit the opportunity to foster the proximal characteristics of young Jordanians that will enable them to compete effectively in the labor market irrespective of educational attainment.

Moreover, relatively lower unemployment rates among the least educated Jordanians should not mask the reasons described above for improving quality of education to better engage these groups. Jobs available to the least educated young people can be of low quality and ill-suited for quality living standards (Chaaban, 2008).

This background highlights at least two latent needs that should be reflected in the theory of change of an intervention for out-of-school youths in Jordan. Firstly, the intervention should be adequately engaging and relevant to attract and retain participants, especially given that many out-of-school youths report having left school chiefly because their educational experience failed in this regard. Secondly, although strictly academic competencies such as math and literacy are important, the program needs to facilitate opportunities for young people to develop the non-academic competencies, such as self-efficacy, social capital, and social skills, that will increase their access and competitiveness in the labor market. By fostering relevance, ownership, and greater support, youth empowerment may constitute a critical strategy for reengaging marginalized youths and preparing them for productive transitions.

Educational reform, vocational training, and non-formal education

Despite the aforementioned challenges for educational provision, Jordan has also been recognized as a regional leader in education reform. At 13% of public expenditure for basic and secondary education, Jordan commits one of the highest proportions of public expenditure on education in the region (Eyango & AlShaer,

2003; Kanaan & Hanania, 2009). Jordan's first major educational reform came about in the 1970s to increase vocational training, expand teacher training and certification, and develop schools' infrastructure (Eyango & AlShaer, 2003). The most recent phase of major educational reform, Education Reform for Knowledge Economy (ERfKE), took place from 2003 to 2007. The main components of ERfKE included reorienting the educational system toward the needs of the knowledge economy; improving administration, data management, and information to drive performance; supporting provision of quality physical learning environments; and promoting learning readiness through early childhood education, including a legal requirement for the Ministry of Education to provide kindergarten education targeting the 4 to 6 age group (Eyango & AlShaer, 2003).

Complementary national initiatives have bolstered ERfKE's objectives, such as the Jordan Education Initiative (JEI), supported by the World Economic Forum, and the Madrasati ("my school") Initiative, launched by Her Majesty Queen Rania Al-Abdullah. Both JEI and Madrasati orchestrate public-private partnerships in support of national educational goals. JEI focuses on improving information and community technology (ICT) in Jordanian schools for grades 1 through 12 (Light et al., 2008). Madrasati's aims are to upgrade and improve the physical facilities of 500 schools designated as the country's neediest and to promote continuous activities that facilitate connections between schools and their communities (Masri et al., 2010). Because of the sizeable economic burden placed on Jordan's government as it tries to adequately prepare and educate its disproportionately large youth population, reforms have required the help of such externally sourced initiatives to better fulfill

their goals. Considered in the context of these initiatives, an innovative non-formal education program to reintegrate socially and economically marginalized young people fits well in Jordan's broader pattern of leadership toward more and better investment in the human capital of Arab youth.

In order to respond to the challenge of dropout and ill-prepared youth, several groups including policy-makers, funders, and researchers have called for increased development of non-formal education and vocational training in Jordan (Aboudan & Murray, 2009; Kanaan & Hanania, 2009; NCFA, 2004). Ahmad and Coombs (1975) defined non-formal education rather broadly as "any organized educational activity outside the formal education system...that is intended to serve identifiable learning clientele and learning objectives." The works of Freire, a Brazilian educational theorist, often influence the methodology of empowerment-based non-formal education models and youth empowerment literature in general (Mohajer & Earnest, 2009; Wallerstein et al., 2005; Wilson et al., 2006; Wong et al., 2010). Freire described a 'pedagogy' through which 'facilitators of empowerment' help marginalized people experience stages of 'critical consciousness'—a progressive transformative process involving the development of awareness of one's relationship to, and ability to affect, change in his or her life and environment (Freire, 1972, 2005). This participatory basis for non-formal education has generated high interest in developing contexts like Jordan.

Official Jordanian policy under the Ministry of Education's Annual Directives (No: 68\1\20680, 12 April 2010) does not allow dropouts to return to formal education after three years of leaving. The premise for such policy is to avoid, for

example, adolescents participating in classes with younger children, but this arrangement nevertheless poses significant barriers for social inclusion of adolescent school dropouts. Consequently, school leaving can be very limiting, giving particular importance to alternative pathways such as non-formal education. To this end, the Jordan Ministry of Education outlined its commitment to non-formal education (officially labeled by MOE as an “education reinforcement program for the drop-out” because “non-formal education” has historically been used in Jordan to describe adult literacy programs under the MOE) as a social inclusion strategy for out-of-school youth in the kingdom in its national report on the development of education (MOE, 2008). As the report describes, “the program aims to disseminate education for all through integrating marginalized children and young people who are vulnerable to risk in a compensatory educational program that meets their needs” (MOE, 2008, p. 58).

The Ministry of Education signed a directive making Questscope non-formal education (QS NFE) the official alternative tenth grade certification path in 2007. In 2009, a memorandum of understanding was signed between Questscope and the Ministry of Education designating the two entities as the responsible parties for all expansion and quality control/assurance of the non-formal education programs. Completion of a 10th grade equivalency test makes Questscope graduates eligible to study a trade of their choice in centers of the national Vocational Training Corporation (VTC). The vocational training curriculum includes an entrepreneurship course for QS NFE graduates, which was designed by Questscope. Graduates who complete the VTC curriculum have the option to secure business

start-up loans from national lenders. Currently, the Ministry of Education is reviewing a policy for allowing QS NFE students to continue in secondary education, leading to college and university.

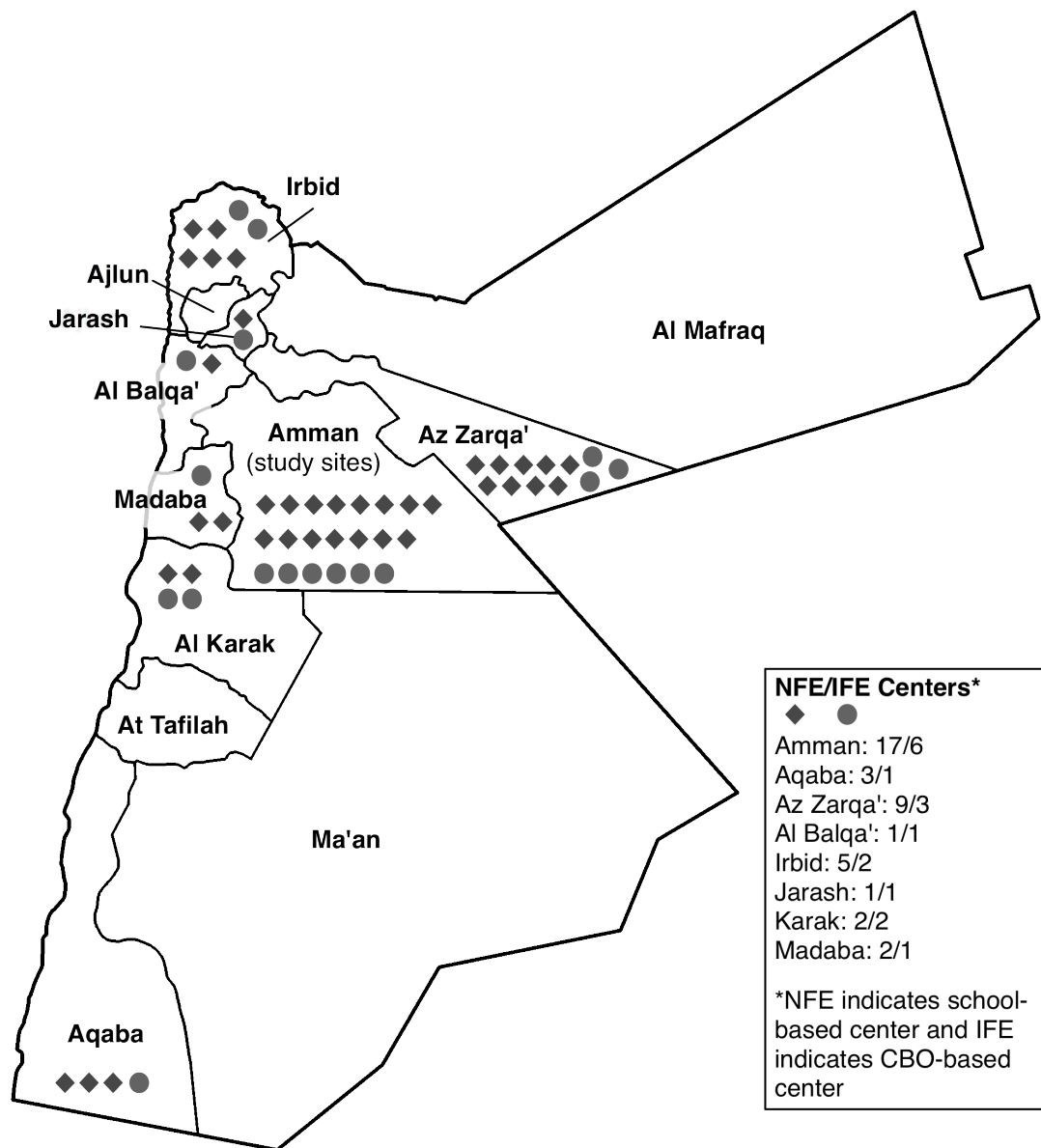
The intervention is further detailed in chapter four, but, in terms of its reach, there are currently forty QS NFE centers operating out of public schools—though they are separate from the formal education system and given their own physical space within school facilities—in eight out of twelve governorates in Jordan and seventeen based out of community-based organizations, where the first educational cycle (called ‘informal education’) can take place. Participants include 13 to 18-year-old males and 13 to 21-year-old females who are out of school. A map of Jordan showing the national distribution of QS NFE centers is provided in figure 3.7. Since 2004, over 7,000 participants have been enrolled in QS NFE.

The evidence-base to support the effectiveness of such interventions, however, is lacking. Substantial pressure on public expenditure in the face of increasing social and economic challenges raises the stakes for evidence-based intervention in the region. In addition to international emphasis on rigorous evaluation in the evidence-based social intervention framework, the Jordanian National Council on Family Affairs, chaired by Her Majesty the Queen, has also made explicit calls for impact evaluation and intervention research into what works for changing out-of-school youth outcomes (NCFA, 2004).

The intervention studied in this research provides social and educational programming to out-of-school youths in Jordan. In doing so, QS NFE aims to engage

youth participants in an empowerment-based programming process, which strives to prepare socially excluded young people with marketable skills and competencies that will enable them to make positive transitions to adulthood and productive contributions to their families, communities, and country.

Figure 3.7. Distribution of QS NFE/IFE centers in governorates in Jordan



*Center symbols on this map do not represent precise locations within governorates.

Youth empowerment in Jordan

Apart from piloting rigorous impact evaluation methodology with a youth intervention in the Middle East, the basic objective of this research is to assess the psychosocial effects of youth empowerment in an Arab context. Much of the existing literature and policy concerning youth empowerment has roots in Western society (Hart, 2008). The few impact evaluations of youth empowerment programs (YEPs)—both based in formal education and not—identified by the international systematic review reported in chapter two were all conducted in the USA. Cultural orientations or aversions toward youth empowerment in the Arab context could have important implications for both implementation fidelity and youth outcomes.

Even Hart, the creator of the iconic ladder of children's participation so often used to illustrate standards for youth empowerment, cautioned that the discourse regarding youth participation in the West that tends to emphasize the assumed programmatic benefit of young people's "individual agency in participating with others to make decisions" may be inappropriate for more collectivist cultural contexts (Hart, 2008, p. 26), as Arab society has been typified (Dwairy, 1998). As such, this section considers environmental conditions that could both promote and obstruct the effectiveness of YEPs with marginalized youth in Jordan.

Amenable conditions for youth empowerment

Despite potential challenges for integrating YEPs in an Arab cultural context, several compelling indicators point toward growth of local conditions that could be considered amenable to youth participation. For example, in the 2010 ASDA'A

Burson-Marsteller Arab Youth Survey, the largest multinational youth opinions survey in the region ($n = 2,000$; aged 18 to 24), respondents reported that living in a democracy was the most important issue for them, even above finding a job, living in a safe neighborhood, or being able to get married (ASSDA'A Burson-Marsteller, 2011). Ninety-two percent of overall respondents and 85% of Jordanian respondents rated living in a democratic country as very important. Arab young people still report that traditional values mean a lot to them and should be preserved (82% overall; 71% in Jordan), but the survey demonstrated that young people believe that traditional values should accompany broader participation, not obstruct it.

These sentiments have been eminently reflected as of late in largely youth-led movements throughout the Arab region advancing freedom and democracy. Recent revolutionary successes in Tunisia and Egypt are believed to have emboldened democratic and participatory sentiments among young people throughout the region (Hamid, 2011a). Though the hypothesis generally lacks data, the broader credence in democratic and participatory youth involvement at the national level could very well translate to desires for similar processes at meso- and micro-level settings, meaning communities, schools, and programs.

Although a European Union mapping study indicated that most 'youth' groups in Jordan are adult-led and linked to governmental institutions, the study also found an emerging category of community-based organizations (CBOs) led by youths and for youths, with youth majorities on governing structures and

responsible for planning and implementing activities (Williamson & Hakki, 2010, p. 21). Because CBOs operate locally and tend to be membership- and volunteer-based in Jordan, this development offers another indication of emerging organic demand for youth empowerment. The study recommended further development of such youth-driven community efforts, particularly in neighborhoods with high concentrations of disadvantaged unemployed young males.

Several broader initiatives within Jordan since 1999 have also exhibited domestically-driven momentum for youth participation. Examples of national initiatives promoting youth involvement in decision-making processes include the development of a children's parliament and national youth conferences providing platforms for youth voice in national issues (Hammand, 1999). In 2006, Jordan launched the initiative "We Are All Jordan" (*Kulna Urdan*) to engage youth nationally in decision-making processes through local youth commissions (Jordan Times, 2008). Though some projects have materialized through this effort, the commission responsible for facilitating its activities has reportedly been largely inactive due to lack of funding (Aboudan & Murray, 2009, p. 68). The government's Higher Council for Youth presented youth participation as its first priority in the 2005-2009 National Youth Strategy for Jordan (HCY, 2004). The very development of the national strategy integrated active youth involvement in the process with approximately 91,000 young people reportedly having participated in the preparation of the strategy (Kanaan & Hanania, 2009; UNICEF, 2007). The document (section 7.1) set out wide-ranging strategic objectives to strengthen youth participation in political, economic, community, and cultural life. Operational

objectives such as “providing greater support and direction to youth centers to become more youth friendly, especially to young people with special needs,” “promoting participatory structures in educational settings,” and “activating participatory methodologies and structures in cultural and recreational activities” speak directly to the aims encapsulated by the Questscope non-formal education intervention examined in chapters four through six of this dissertation.

Jordan claims to be the first country among Arab nations to have produced such a national strategy outlining a “comprehensive and long term vision and plan for the development of its young women and men” (HCY, 2004, p. 3). Impetus for the national strategy came in part directly from the country’s head of state. His Majesty King Abdullah II articulated youth empowerment as a national urgency in an official address to parliament in 2003: “The change to which we aspire can only be achieved through mobilizing young people and listening to their views as they are the pillars of tomorrow and the substance of change” (HCY, 2004, p. 3). Both the King and Queen of Jordan have reiterated the priority of engaging young people in political and social processes in many addresses and official documents since.

Challenging conditions for youth empowerment

The tone set by national initiatives and young people’s own expressed priorities does provide some evidence of amenability within the Jordanian context to involving young people in the decision-making processes that affect their lives and communities. In other words, youth empowerment may in fact have appeal that reverberates more broadly than through Western society alone. Nevertheless, social

contexts for interventions like YEPs can be convoluted with multifaceted and contradictory trends and attitudes at work. UNICEF's Jordan Situation Analysis, for example, contended that school leadership often disregards progressive guidelines for youth participation through student councils set out by the Ministry of Education (UNICEF, 2007). Adolescent females in Jordan have especially reported feelings of limitation in choosing academic disciplines, careers, and participation in extra-curricular activities due to over-protective family control (UNICEF, 2007). Even with emerging signs of proclivity toward youth participation, implementation can still falter if local circumstances and traditional authority paradigms in the home, organization, or community negate empowerment-based programming methodology.

Cultural propensity toward youth empowerment may be further constrained when a program's target population is stigmatized or considered less worthy or capable of meaningful participation. A survey of 2,000 young people in Jordan, for example, found generally negative attitudes toward vocational training (Al-Sa'd, 2007). Ninety-six percent of youths agreed that their parents want them to study at an academic or college preparatory school. Several researchers in the Middle East have alluded to a "culture of shame" that relegates out-of-school and vocational youth to lesser status in a society that largely equates high social prestige with high education (Al-Sa'd, 2007; Sharp, 2010a). Given the lower social regard for vocational education, out-of-school youth could conceivably be held in even lower regard.

This challenge extends throughout the broader Arab region. In fact, only since the 1990s has Egyptian law distinguished youth who work on the streets or remain there during the day from juvenile delinquents (Bibars, 1998). Forty-eight percent of interviewed street youth in Alexandria and Cairo reported community disapproval as a problem they faced (UNODC, 2002). Surveyed children and adolescents working in the streets of Yemen have also reported disappointment that their work activities were often viewed with contempt by society (IRIN, 2008). Indeed, their views were triangulated by interviews with Yemeni teachers who associated working children with late attendance, poor cleanliness, poor concentration, and unacceptable work (Dyer, 2007). Without adequate staff selection and training, this context may translate to a deficit-based disposition toward out-of-school youth that could obstruct quality of delivery with YEPs like Questscope non-formal education (QS NFE). Previous research has indicated an effect of teachers' expectation bias for students on long-term student performance (de Boer et al., 2010); expectation bias based on stigmatization could similarly circumvent program staff members' propensity to facilitate empowerment with youth for whom they hold lower expectations.

The diffusion of youth empowerment through program implementation could also be considered in the context surrounding organizational behavior. Understanding whether Arab society is amenable to empowerment-based approaches in general—and what kinds of empowerment—is complicated by historical, cultural, and religious nuances. As Abu-Saad and Hendrix wrote, "In Islam, there remains a tension between participatory/consultative approaches and

authoritarian approaches to management and organization, both of which are supported in the Quran itself, and in Islamic tradition” (Abu-Saad & Hendrix, 1995, p. 142).

Arab organizational culture was moved closer to authoritarian orientation with the influence of colonial bureaucratic and Ottoman autocratic systems resulting in what Al-Kubaisy (1985) described as “sheikocracy” in Arab organizational behavior. Sheikocracy’s influence to Arab organizational behavior is characterized by: “Hierarchical authority, rules and regulations contingent on the personality and power of individuals who make them, an ‘open door’ policy, subordination of efficiency to human relations and personal connections, indecisiveness in decision making, informality among lower-level managers, and a generally patriarchal approach” (Ali, 1990, p. 16). Abu-Saad and Hendrix (1995) also noted how Bedouin tribal roots can promote kinship, but also “we-they” approaches to dealing with non-kin and paternalistic leadership styles. The sum of these influences may augment the supportive, relational aspects of youth empowerment environments described by youth empowerment literature, but they may also present subversive circumstances for involving youth as equal partners and valued contributors in program decision-making processes.

The extent to which youth empowerment is a culturally valid modality for improving the wellbeing of Arab young people has thus far lacked empirical substantiation. One humbling development to this end has come from investigations into the relationships between parenting styles and mental health of adolescents

across cultures. In Western cultural contexts, research has linked authoritative parenting—characterized by attentive but “non-coercive and democratic” approaches—to positive youth outcomes, whereas authoritarian parenting—characterized by “harsh, restrictive, and psychological methods of control that do not foster psychological autonomy”—has predicted negative mental health outcomes (Dwairy et al., 2006, pp. 262-263). These parenting typologies come from Baumrind (1991) who classified parenting styles as *authoritative*, *authoritarian*, *permissive*, and *rejecting-neglecting*. Baumrind (1991) found in a study of American adolescents and families that adolescents raised in authoritative and democratic families fared better than peers who were exposed to other types of parenting styles on a range of indicators, including competence, maturity, motivation, academic achievement, social responsibility, self-esteem, internal locus of control, problem behaviors, and perceptions of their parents as loving and influential (Baumrind, 1991). The evidence provided a compelling basis in Western society for empowerment-based orientations to adolescents in the household.

Research in other cultural contexts, however, has at times sharply contrasted with results detected in Western Caucasian youth populations. Authoritarian parenting styles, for example, have been linked to higher assertiveness and independence with African American girls (Baumrind, 1972), higher competence for youths in high-risk environments (Baldwin et al., 1990), better adjustment and academic performance among Asian American adolescents (Steinberg et al., 1994), and better achievement outcomes for youths in China and Hong Kong (Leung et al., 1998). In a series of recent studies, Dwairy and colleagues have tested the

consistency of such findings in the Arab world (Dwairy, 2004; Dwairy et al., 2006a; Dwairy et al., 2006b). While authoritative parenting was still linked to the best mental health outcomes, unlike in Western societies, there was no, or only weak, evidence for negative outcomes associated with authoritarian parenting styles in Arab societies (Dwairy, 2004; Dwairy et al., 2006a). The findings led Dwairy to suggest that “the implications of different parenting styles are contextual rather than universal and depend on culture and gender” (Dwairy, 2004, p. 249). Plausibly, this could hold true for teacher and youth workers’ styles as well.

Further to the point, Dwairy and colleagues (2006) noted research that suggests a relationship between ‘cultural double standards’ (e.g., tensions between traditional Bedouin tribal society and modern, urban society) and higher risk for psychological distress in the Arab world. In principle, this cultural double standard perspective could apply to tensions between Arab young people’s experience with empowerment-based programs like QS NFE and their experiences in society outside of the program. In other words, at least in the short-term, the tension between an empowering program culture and a potentially disempowering household or societal culture could cause more strain than benefit on Arab adolescents’ mental health.

Ultimately, the seeming contradictions in various leadership styles and cultural attitudes within the Arab context may better reflect a need for ‘Arab versions’ of youth empowerment rather than a premise for accepting or rejecting ‘Western’ youth empowerment. To this end, it is probably advantageous that the QS

NFE program was developed and implemented within Jordan rather than simply transported from another cultural context. In assessing the development of school counseling in Jordan, for example, Farah (1992) described the incorporation of Islamic values and family involvement as culturally important aspects of professionals' counseling strategies. Such features of youth-adult interactions are conceivably important for staff responsible for facilitating youth empowerment processes with out-of-school youths as well.

Furthermore, Hart (2008) suggested that, in non-Western contexts, youth participation may involve a high level of youth engagement in community youth projects while still maintaining a level of respectful adult authority and adult initiation in project design. The process factors specific to implementing youth empowerment in community-based programs are explored in chapter six. Although regular involvement of youth in decision-making processes is central to the definition of youth empowerment, the youth-adult relationship and functions through which young people are engaged may be necessarily nuanced by local context.

Out-of-school Arab youth

Although several initiatives in Jordan, such as those mentioned in this chapter, have promoted further youth participation, they have largely concentrated on involving university students and otherwise socially and economically advantaged groups of youth. Less investment and research have been devoted to the process and impacts of engaging marginalized subpopulations of youths in program

or policy decision-making. This gap in services and attention is not unique to Jordan. Researchers in Yemen contended that of the 6,000 civil society organizations across the country, only 3 to 5 work specifically with street children (IRIN, 2008). Although youth empowerment literature and international organizations strongly encourage youth participation among socially excluded groups like out-of-school youths, special circumstances may require additional supports and strategies to engage them effectively. The UNESCO (1994) Salamanca Statement on Special Needs Education specifically referenced working and street children as having special educational needs, and, for the same reasons, these special needs conceivably remain important in the broader framework of social policy and intervention. The Salamanca Statement broadly includes various groups of disadvantaged or marginalized children as those with special needs, which includes but is not limited to disabled children.

Various terms such as marginalized, disadvantaged, socially excluded, and dropout youth are used to describe the target population for non-formal education in Jordan. This research generally refers to the specific non-formal education target population as “out-of-school” youths. Burns and colleagues (2004, p. 3) defined “out-of-school youth” as those who have dropped out of school, those who have never attended school, or those who participate in non-formal school programs. Sometimes, the youths targeted by non-formal education in Jordan are also described as “street children” (NCFA, 2004). Street children have been defined as “minors who earn their living by working on the streets or as children who reside on the streets full or part time” (Lusk et al., 1989). The vast majority of street

children live with their families but work on the streets to earn money for their families (Aptekar, 1994). Based on these definitions, not all youths who are out-of-school classify as street children, and not all street children are out-of-school. Nonetheless, there is substantial overlap between the two categorizations for the population of concern in this dissertation's study.

Specific and high-quality research evidence on the needs and conditions of out-of-school youths in Jordan is lacking, but relevant (albeit limited) research from the region can provide some contextual background. The first major study on street children by the Yemini government surveyed 4,760 (of the country's estimated 30,000) street children, aged 6 to 17, in 8 out of the country's 21 governorates (IRIN, 2008). Highlighting household financial hardships driving youths to the streets, 82.6% of respondents reported that their earnings went to helping their families. Common jobs included street vendors, porters, car washers, and beggars. Twenty-five percent of respondents reported experience with various forms of violence, including sexual abuse, robbery, beatings, and harassment by municipality workers. Most street children (62.2%) came from urban areas, and many suffered from a range of health problems.

Findings from research in Egypt have reiterated many of the aforementioned themes. Interviews and focus groups with street children ($n = 185$) conducted by a United Nations Office on Drugs and Crime (UNODC) study in Cairo and Alexandria provided a range of data on the population useful for informing related interventions and theories of change (UNODC, 2002). For example, nearly two-

thirds of interviewees reported habitual use of harmful substances—especially cigarettes (70%) and glue sniffing (60%). As for reasons for substance use, 70% of respondents cited relief from the pressures of the street, and 60% cited peer pressure. Learning a profession (30%) and having a permanent job (28%) were identified as participants' greatest needs and hopes for the future. Many youths (26%), however, could not identify their needs or hopes for the future, suggesting a lack of vision and aspiration, perhaps by consequence of life experiences that failed to foster hopeful sentiments.

Similarly, a study that surveyed 159 urban street children (aged 8 to 13) in Gaziantep, Turkey found a significant relationship between several variables that often characterize street youths' experiences (e.g., physical and sexual abuse, running away, and gang membership) and negative scores on the Hopelessness Scale for Children (Duyan, 2005). The study also found a high prevalence of poor family relations as well as both physical and verbal abuse among the sample. Experience with violence (by peers, community members, or employers) was also the problem most frequently cited (86%) by research participants in Cairo and Alexandria (UNODC, 2002). The study team determined that factors contributing to the prevalence of street children include: low income and family educational level, family break-down, dropout out of education, large family size, unplanned rural-urban migration, and the declining role of the extended family.

A study by Bibars (1998) based on qualitative interviews with street youths in Cairo also illustrated commonplaceness of physical and sexual abuse both in the

household and the community, and damaging stigmatization that street youths experienced. Based on qualitative insights, Bibars recommended increased activities that are “built around the child and her/his needs, and not the negative and traditional perspective of adults,” as well as involving the child’s participation in determining programs’ actions and activities along with “facilitators” with whom youths can build mutual trust. These recommendations echo central elements of the theory of change behind the empowerment-based non-formal education program described in subsequent chapters.

Quality and accessible research on the conditions of out-of-school or street youth in Jordan is lacking, but a country study of disadvantaged children by Jordan’s National Council for Family Affairs suggested similar conditions for this subpopulation of youths (NCFA, 2004). The report contended that street youths are at higher risk of exposure to early and abusive sexual experiences, physical violence, substance abuse, child labor, experience with juvenile authorities, and community rejection. These claims, however, are apparently driven by the report committee’s field expertise rather than specific, local scientific data. The report called for future studies into risk factors and relationships between them. Nonetheless, the sum of the available evidence from even limited research into the attitudes and conditions of out-of school and street youths in the Arab region presents a bleak situation for youth development. Community-based interventions that can effectively engage out-of-school youths and facilitate the positive attachments, motivating environments, and skill-building experiences they need to make positive transitions to adulthood are urgently needed.

Conclusion

Context plays a vital role in planning and interpreting intervention research. Important background is comprised of various levels of interacting historical, economic, political, and social factors that influence an intervention's implementation, uptake, and impact. One of the most salient realities that emerge from the background research presented in this chapter is Jordan's precarious social and economic position as a country strained by unplanned growth, surrounded by geopolitical strife, and largely void of natural resources. In this context, Jordan has no choice but to transform its greatest challenge – its oversized and underprepared youth population – into its greatest asset with results-driven interventions to develop human capital. This need is appreciated by the Jordanian government, which has led some of the region's most ambitious investments and reforms in youth education and workforce preparation. To this end, youth empowerment, implemented through non-formal education, has enjoyed early support from the Jordanian government as a plausible solution for fostering human capital among some of the country's most socially and economically marginalized youth.

Nevertheless, the background research also highlights limits to political freedoms, enduring issues with corruption, and a heavy government apparatus, all of which pose confusing messages and challenges for civil society, including YEPs. Challenges for youth empowerment are exacerbated by deep-rooted hierarchical and patriarchal tendencies in organizational systems and society that could jeopardize the applicability of participatory methodology to Arab society. On the

other hand, the context also reveals growing support for youth participation from top Jordanian government leaders, and Arab youths continue to signal desire for greater opportunities to participate in the processes that affect their lives and communities.

This chapter began by describing the value of context for identifying the assumptions behind a theory of change. Imas and Rist (2009, p. 157) posited that key assumptions include major contextual or environmental factors or events that will likely influence the intervention. The following chapter will describe the specific theory of change for QS NFE, but here the author briefly discusses key assumptions informed by the secondary background research presented above that underlie an evaluation of empowerment-based non-formal education for out-of-school youths in Jordan. Assumptions informed by the background research include the following:

- *The policy context is amenable to QS NFE's quality implementation.*

Jordan's political and economic context has created both opportunities and challenges for civil society organizations to thrive. QS NFE's ability to realize intended impacts, if the intervention model is in fact effective, depends on the extent of its implementation fidelity. Implementation fidelity, in turn, could be influenced by environmental factors influencing whether Questscope and implementing centers will have the freedom, stability, and capacity to adequately implement youth empowerment programming

- *Youth empowerment is culturally viable among the program's key stakeholder groups.*

Although several advances have been made for youth participation within the Jordanian context, aversions to horizontal, empowering relationships, especially with marginalized groups, still remain in aspects of the societal and organizational paradigm. Staff members' propensity to implement high-quality youth empowerment processes and young people's responsiveness to them could be obstructed by rival cultural dispositions and stigmatization associated with the target population.

- *Empowerment-based programming will overcome the reasons target youth left education in the first place to attract and retain participants.*

The QS NFE program has the additional challenge of engaging youth already disconnected from formal education, the principal social system beyond the family for their age group. Many such youths left school because of family pressures, poverty, child labor, and, most importantly, poor experiences with the quality of educational settings. As such, the program's ability to affect youth outcomes is contingent upon its ability to provide the flexibility, supports, and quality programming environment needed to attract and retain the very young people that schools lost.

- *The program is able to change psychosocial outcomes with a particularly vulnerable target population.*

Although data on the conditions of out-of-school youths in Jordan is deficient, limited qualitative and quantitative research on out-of-school and street youths in the Arab region illustrates an austere picture of social exclusion across social, emotional, physical, and economic dimensions. QS NFE's theory of change

assumes that the intervention comprised of its multiple elements is adequate to measurably improve developmental outcomes in a target population potentially characterized by high distrust for adults, peers, and institutions as well as multiple forms of psychosocial problems.

Community-based social interventions never operate in isolation from broader organizational, community, and societal environments fraught with variables that could help or hinder a program's theory of change. Some of the key assumptions for the program evaluation described in the following chapters have emerged from the literature review presented here. The extent to which these assumptions materialized in the QS NFE program evaluation and their potential ramifications for the implementation and impacts of youth empowerment were explored in the implementation and process study reported in chapter six.

Chapter 4: Randomized Controlled Trial: Intervention and Methods

Introduction

The systematic review conducted in this dissertation set out to assess the state of the evidence on effects of youth empowerment programs (YEPs) on psychosocial outcomes. Excluding work conducted in the context of this dissertation, only two includable studies were identified from 8,789 international citations netted by the initial search strategy. The two included studies demonstrated some promise for YEPs' role in improving young people's developmental outcomes. The effects across outcomes, however, were mixed and modest, and the studies were confined to the USA. The review highlighted a need for further use of rigorous impact evaluation designs to assess the effectiveness of YEPs, especially in different cultural contexts where youth empowerment may have varying implications. The present study aims to respond to that need.

The preceding chapters have reported efforts undertaken under the auspices of the 'development' phase of the MRC's complex interventions framework through theoretical development, systematic review, and intervention contextualization. The contents of this chapter are primarily situated in the 'feasibility and piloting' phase, outlining the methodological considerations and procedures supporting impact evaluation of a YEP reported in chapter five. More specifically, chapter four describes the methods of a randomized controlled trial (RCT) in the Hashemite

Kingdom of Jordan and the youth empowerment intervention on which the trial was conducted.

This chapter and the next strive to meet standards outlined by the CONSORT statement for reporting randomized trials (Boutron et al., 2008; Schulz et al., 2010). The intervention studied, Questscope non-formal education (QS NFE) is a multisite program for out-of-school youths in Jordan. The primary hypothesis tested by this study and embedded in youth empowerment's theory of change is that participatory community-based programming that engages young people in decision-making and positive group processes facilitate mastery and social experiences that improve participants' sense of self-efficacy, social skills, positive connections, and prosocial attitudes and behaviors.

The program aims to empower young people by way of a participatory education model that regularly involves young people in program decision-making concerning class learning topics, group expectations, and educational and recreational activities. According to the program model, formal education teachers are retrained and supported by Questscope to foster an empowerment-based methodology, positive program environment, asset-building activities, and meaningful youth-adult relationships for marginalized young people. The extent to which actual implementation of the empowerment process is realized in these centers is examined in chapter six.

Intervention

Implementing organization

Questscope is a non-profit, non-governmental organization working in low-income communities in the Middle East. The organization, a registered charity in the United Kingdom and Jordan, was established in 1988. With an annual budget of \$5.3 million (USD) for fiscal year 2008-2009 (Questscope, 2010), the organization's staff consists of over forty Arab professionals. By collaborating with community organizations and national leaders, Questscope works to design and implement sustainable programs that provide pathways for educational, occupational, social, and personal growth. Questscope programs include mentoring/volunteerism, non-formal education, juvenile justice reform, women-led initiatives, spiritual formation, and Iraqi children initiatives.

Questscope non-formal education

A large diversity of programs has been implemented worldwide under the banners of 'non-formal education' and 'informal education.' While some follow a participatory approach for youth programming, others do not. More still are not even geared toward young people at all, but rather function as literacy programs, for example, with adults (Ambasht, 1992). Ahmad and Coombs (1975) defined non-formal education rather broadly as:

Any organized educational activity outside the formal education system—whether operating separately or as an important feature of some broader activity—that is intended to serve identifiable learning clientele and learning objectives.

So as to avoid confusion with other varieties of non-formal education, this paper uses the term 'Questscope non-formal education' (QS NFE) to indicate the YEP of interest in this research.

Questscope non-formal education began in 1998 as a largely unstructured effort involving a small number of community-based organizations trained by Questscope staff in supportive relationships with local young people, using learning as a vehicle to help out-of-school youths gain positive connections to their community and a sense of self-worth. At that time, the program was called 'street education.' While an important beginning for Questscope in developing its training, programming and educational methodologies, the loosely organized street education program left more to be desired for sustaining the initiative and enabling the youths to make positive transitions to adulthood.

Official policy in Jordan under the Ministry of Education's Annual Directives (No: 68\1\20680, April 12, 2010) does not allow dropouts to return to formal education after three years of being out, resulting in a considerable barrier for young people desiring second chance pathways out of exploitation, child labor, poverty, and other forms of social exclusion. As demand for the program appeared to grow, and as feedback from the youths continued to indicate a strong desire for a second chance at an officially recognized education that would open formal pathways to vocational training and perhaps further schooling, Questscope approached the Ministry of Education in 2004 to further institutionalize the

program. The result was (at the time) a unique governmental-non-governmental partnership between the Ministry of Education and Questscope.

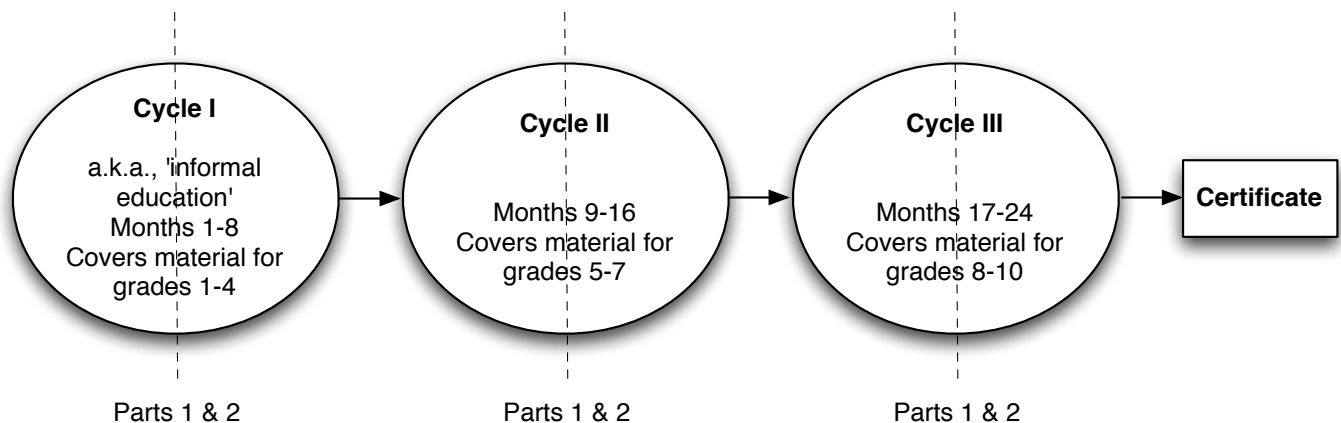
Jordan's Ministry of Education signed a directive making QS NFE the official alternative tenth grade certification path in 2007. In 2009, a memorandum of understanding was signed between Questscope and the Ministry of Education designating the two entities as the responsible parties for all expansion and quality control/assurance of the non-formal education programs. Completion of a 10th grade equivalency test makes QS NFE participants eligible to study a trade of their choice in centers of the national Vocational Training Corporation (VTC). Vocational training curricula include an entrepreneurship course for QS NFE graduates, which was designed by Questscope. Students who complete the VTC curriculum have the option to secure business start-up loans from national lenders. Currently, the Ministry of Education is reviewing a policy for allowing QS NFE students to continue in secondary education, leading to college and university.

Program methodology

The QS NFE program targets out-of-school youths and particularly those with few or no safe and positive opportunities available to them. Males can enter the program from 13 to 18 years of age, and females can enter from 13 to 21 years of age. Class size generally ranges from fifteen to thirty youths with four 'facilitators' (program staff) per center. The intervention is a two-year program involving three eight-month 'cycles.' Educationally, each cycle is designed to correspond with the basic academic goals of specific grades in the formal education system. The program

structure is illustrated in figure 4.1. Cycle one covers material for grades one through four, cycle two for grades five through seven, and cycle three for grades eight through ten. At the end of each cycle, participants take a test specifically designed for QS NFE by the Ministry of Education in partnership with Questscope staff, which determines participants' ability to progress to the next cycle or, in the case of completing the final cycle, to obtain the alternative tenth grade certificate. Youths are recruited to QS NFE through a variety of channels, including facilitators, community-based organizations, mosques, Questscope staff, and peers who have themselves participated in the program.

Figure 4.1. QS NFE program cycles



The program occurs five days per week (though participants are only expected to participate two to four days per week) after traditional school hours. Female centers tend to operate earlier in the day, starting around 3:00pm, whereas male centers tend to provide programming in the evening, beginning around 6:00pm. This timing makes the program accessible for out-of-school youths who are often involved in some form of labor during the day, and it provides an exclusive

environment for QS NFE to develop its own program culture while utilizing the school facilities.

Although QS NFE is outside of the formal education system, the partnership with the Ministry of Education includes the designation of Ministry of Education teachers to staff the program and spaces in public schools to function specially as QS NFE centers. The physical space of QS NFE centers is intended to create a comfortable, youth-friendly environment integrating youth involvement in the layout and appearance. Centers include computers and roundtables more conducive to dialogue and community-building than the simple rowed desks format of traditional classrooms.

While QS NFE clearly involves academic goals, Questscope considers non-formal education principally a social activity. The primary objective, especially in the first cycle, is to develop out-of-school youths' sense of self-efficacy, social skills, positive connections to adults and peers, and an enthusiasm for learning.

Location within the youth empowerment framework

The systematic review described in chapter two defined youth empowerment programs (YEPs) minimally by four components: (1) regular youth participation in program planning or implementation, (2) supportive adults or older youth mentors (these need not be one-to-one mentoring), (3) asset-building activities, and (4) a prosocial environment. The QS NFE program model meets these criteria in the following ways:

1. *Participation in decision-making* – The QS NFE methodology is based on a participatory learning scheme in which youths are actively engaged in program decision-making via regular involvement in processes such as determining program guidelines (the ‘educational contract’), deciding learning topics and activities, and planning cultural and community events.

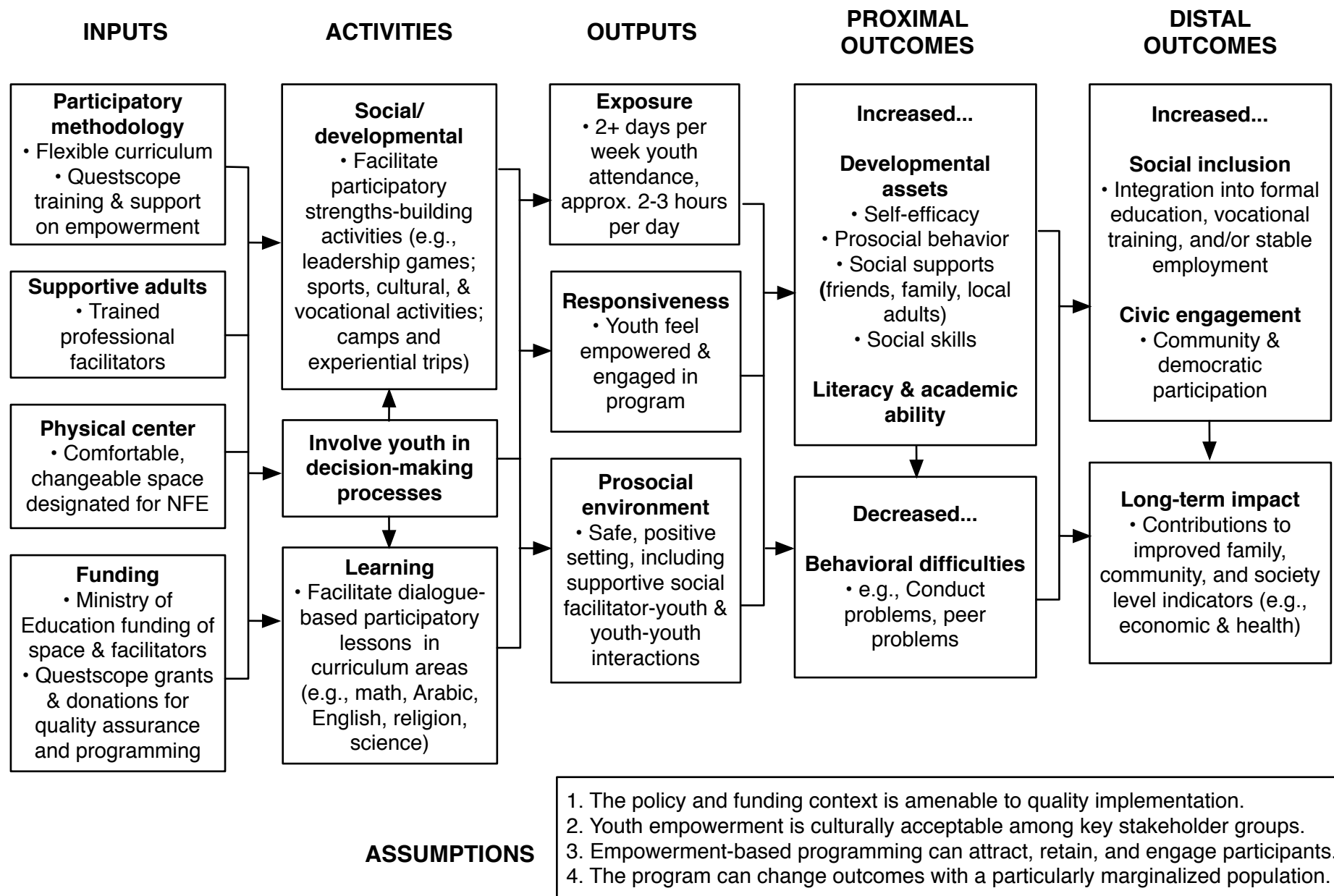
2. *Supportive adults* - Teachers are retrained by Questscope as facilitators specifically for the purpose of cultivating positive, supportive relationships with at-risk youths. Youth participants are expected to come into contact with these adult supporters for at least four to six hours per week.

3. *Asset-building activities* - Youths participate in regular skill-building and asset-building activities through literacy development, vocational training, arts and culture, debates, decision-making, and leadership practice, among other experiences facilitated by QS NFE.

4. *Prosocial environment* – Facilitators are trained to foster a positive social setting conducive to youth development. Young people establish group expectations to encourage more positive peer and facilitator interactions with youth. Collective group decision-making is intended to inculcate trust and positive group norms.

The extent to which these intended components are fully implemented in practice is examined in chapter six.

Figure 4.2. Questscope non-formal education theory of change



Following the MRC's complex intervention guidance in developing appropriate theory (MRC, 2008), a diagram of the intervention's theory of change incorporating these components is included in figure 4.2. A theory of change specifies how activities are expected to lead to interim and long-term outcomes and identifies the contextual conditions that could affect the implementation or outcomes (Connell & Kubisch, 1998). Important assumptions emerged from reviewing contextual conditions in chapter three, and those are included in the diagram. The extent to which QS NFE succeeds in implementing the abovementioned youth empowerment framework is explored in the process and implementation study discussed in chapter six.

Notably, while these intervention characteristics situate QS NFE within the youth empowerment framework, it is not necessarily as comprehensively participatory as some youth empowerment models, particularly those developed in the West. As pointed out by Wong and colleagues (2010), the extent of youth control may necessarily vary according to developmental stages, and youths of low-income households and otherwise marginalized backgrounds may not have the time or energy to exercise youth-driven programming at all levels and in all respects due to additional life responsibilities. In the context of QS NFE, youths are not currently involved at a partner level with adults in *all* levels of decision-making and program design in the organizational structure (e.g., regularly facilitating classroom learning experiences for other youths).

Over the last year, Questscope has moved further toward involving youth voice in higher levels of the organizational structure. In January 2010, Questscope initiated a pilot Youth Advisory Committee (YAC), which consists of three youth leaders from several QS NFE centers to engage in structured leadership development, participatory research activities, and making recommendations to Questscope and the Ministry of Education for the development of the program and policy concerning out-of-school youth. Most program youths, and those included in this impact study, however, will not have had this degree of direct participation in higher levels of the organization's structure. Additionally, many youth empowerment models emphasize regular and direct commitment to social action activities (Mohajer & Earnest, 2009; Zeldin et al., 2008). While sociopolitical issues are often addressed in critical reflection with youth participants, if social action takes place among QS NFE youths, it is typically haphazard and not a direct feature of the intervention design.

Modifying the teacher paradigm and approach is an essential step to creating an environment that is different from the one that dropout youths left in the first place with the formal education system. Zimmerman (2000) argued that an "empowerment approach to intervention design, implementation, and evaluation redefines the professional's role relationship with the target population. The professional's role becomes one of collaborator and facilitator rather than expert and counselor." There is general consensus in the youth empowerment and positive youth development literature that youth need access to supportive, caring adults who work to leverage their strengths rather than focus on their perceived

weaknesses (Cargo et al., 2003; Zeldin et al., 2008). Substantial research has found quality, supportive connections between youths and non-parental adults to be both a significant protective factor (Blum, 2003; Garmezy, 1985; Rutter, 1987; Werner & Smith, 1992) and a mechanism for effective interventions (Grossman & Tierney, 1998).

In the educational context, the quality of relationships between students and their teachers is considered a main contributor to classroom climate and student outcomes (Jones et al 2008; Pianta & Allen 2008; Steinberg & Rollins 1995; Mantzicopoulos 2005). A review of mentoring programs found that programs with mentors from helping professions, such as teaching, reported larger effect sizes than those with mentors not from helping professions (DuBois et al., 2002); this may reinforce the valuable role that retrained teachers can play in the lives of QS NFE youths. Studies suggest that supportive interpersonal relationships between students and teachers correlate with more positive academic attitudes and values, more satisfaction with school, and increased academic engagement (Battistich et al 1995; Shouse 1996; Solomon et al 2000; Felner et al 1997; Skinner & Belmont 1993; Klem & Connell 2004). To this end, QS NFE not only facilitates connections between youths and adults in schools and community-based organizations, but its training also aims to help adults to better understand disadvantaged youths in such a way that enables them to identify, affirm, and develop their personal assets and goals through caring, long-term relationships.

All teachers selected to serve in the program as ‘facilitators’ must participate in a special training curriculum designed and implemented by Questscope. The capacity of QS NFE facilitators to foster an empowering learning environment is critical to the program’s execution. The full training involves approximately 85 hours of sessions and exercises divided among three levels. Training covers a range of topics intended to orient facilitators’ perspective and skills to effectively encourage youth development and participation through learning with marginalized youth. The first level orients facilitators to participatory and dialogue-based learning, adolescent development, and skills to prepare for program sessions. The second level of the training puts special emphasis on ‘participatory rapid appraisal,’ which helps facilitators connect to the community, understand the context in which dropout youths are living and working, and build local channels for recruiting youths into the program. Participatory rapid appraisal is a methodology that evolved from ‘rapid rural appraisal’ and has strong linkages with Robert Chambers at the University of Sussex, a figure whose works have had an important influence on QS NFE’s approach (e.g., Chambers, 1992). The third level goes deeper into topics of educational tools and social change and emphasizes skill-building for group facilitation and participation.

The first level of training is generally conducted in an outside environment with facilitators from multiple centers (e.g., in a hotel or conference center), and subsequent levels of training are typically conducted in the program environment either at the center or in the local community. Moreover, Questscope assigns an organization staff person (called a ‘field coordinator’) to each center. In addition to

coordinating administrative documents and payments, the field coordinator is intended to provide ongoing coaching and support as well as an accessible channel for communication between the centers and the organization.

The Questscope educational model involves learning targets so as to meet certification standards held by the Ministry of Education, but the topics are generated from the students' interests and choices. Youths determine as a group what issues or questions they will study, and facilitators integrate lessons from different subject areas to help the participants grapple with the learning topics they choose. As such, the curriculum offers a skeletal structure consisting of critical knowledge that the youths must acquire, but, through a participatory process, the young people provide the content of their own ideas, topics, and activities. This empowerment-based learning process is intended to foster relevance and engagement among program participants.

The participatory methodology of QS NFE is influenced by the works of Paulo Freire, a Brazilian educational theorist, on 'popular education' and 'critical consciousness' and of Forrest Tyler, from the faculty of psychology at the University of Maryland, on 'prosocial communities' for youths. Freire's book, *Pedagogy of the Oppressed* (1972), challenged authority in the education process and advocated that marginalized populations and other members of society become co-learners. Freire described a 'pedagogy' through which participants experience stages of 'critical consciousness'—a transformative process involving the development of awareness of one's relationship to the world around him or her and ability to affect change in

one's life and surroundings. Freire believed that it is the job of a 'facilitator of empowerment' to help navigate marginalized people through this process using shared dialogue as a key instrument. A sizeable amount of literature concerning youth empowerment beyond Questscope has also integrated Freirian concepts into theoretical underpinnings for studies and evaluations (Mohajer & Earnest, 2009; Wallerstein et al., 2005; Wong et al., 2010). In a similar vein to Freire, Tyler (2004) emphasized the development of confidence in a youth's ability to affect his or her environment, and he described 'prosocial communities' as a conceptual framework for the connection of community members and institutions to a young person so as to strengthen the youth's role and value in society.

In summary, with the influence of established theory in youth development and participation, the QS NFE intervention has been developed on the theoretical underpinnings of advancing the program's relevance and ability to change psychosocial outcomes with marginalized young people by establishing a participatory, positive group setting made possible through retrained teachers acting as facilitators, youth-friendly physical space, and engaging activities.

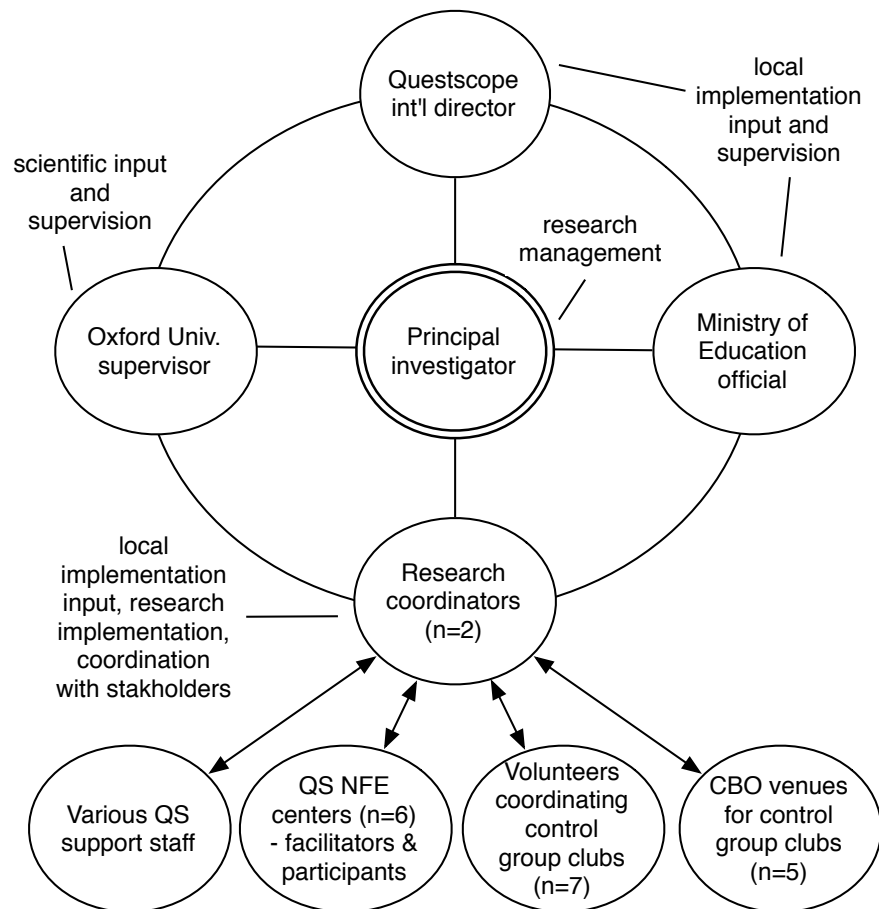
Methods

Research coordinating team

A research team was assembled to advise the planning and implementation of the project. In order to balance academic and practitioner considerations in the research process, the team included representation from Oxford University's Centre for Evidence-Based Intervention (the principal investigator and his supervisor, co-

director of the centre), the executive director and NFE project staff from Questscope, and the official responsible for the QS NFE program in the Ministry of Education. A diagram of the research coordinating team and associated stakeholder groups is included in figure 4.3. The research team gave input and approval into the major decisions in research planning and implementation. Research coordinating or advisory teams have been employed in previous experimental trials as important mediums for strengthening the cultural adaptation and local relevance of the research instruments, design, and implementation while maintaining academic checks to protect the scientific standards of the study (e.g., Harris et al., 2001).

Figure 4.3. Research coordinating team



Setting and population

The trial involved six QS NFE centers in urban, low-income communities in Amman, Jordan. Four sites were male centers and two were female centers (no QS NFE centers are co-ed, corresponding with the public school system and cultural standards). Program and study age eligibility for males was 13 to 18 years old and 13 to 21 years old for females. In general, the participants were Arabs of Islamic faith. Although the program also enrolls large numbers of Iraqi refugees, only Jordanians were included in the study, given the unique circumstances of Iraqi youths and higher likelihood of moving during the course of the study for reasons related to their displaced status. Youths ineligible for the study but eligible for the program were simply recruited and enrolled in the program as normal.

Research design

A randomized controlled trial (RCT) was used to assess intervention impacts. The use of RCTs with social and development interventions is not uncontested. Critics have argued that using RCTs, among other claims, is inappropriate or impractical for complex social interventions, disingenuously tries to measure immeasurable social change concepts, lacks external validity, unethically deprives people of interventions with assumed benefits, and ignores the merits of other research methodologies (EES, 2007; Nutbeam, 1999; Prowse, 2007; Rodrik, 2009). Each of these critiques has been responded to exhaustively in the literature (e.g., Hawe et al., 2004; McCord, 2003; Morton, 2009a; Oakley, 2001; Oakley et al., 2006; Sackett & Wennberg, 1997). As asserted by authors of the MRC framework for developing and evaluating complex interventions, “randomization should always be

considered because it is the most robust method of preventing selection bias” (Craig et al., 2008, p. 981).

When feasible and appropriate, the methodological strengths of an RCT in reducing bias compared to non-randomized alternatives are well documented (Cochran & Cox, 1957; Glazerman et al., 2003; Jadad & Enkin, 2007). Most importantly, RCTs are best equipped to establish two or more groups of participants that are the same on both observable and unobservable characteristics. As a result, the investigator can compare changes in a group that receives the intervention of concern to a group that does not. Assuming adequate sample size and quality implementation of the RCT, this establishes a credible ‘counterfactual’—an ability to ascertain not only how participants exposed to the intervention group changed on specific outcomes over a certain period of time, but also what would have happened had they not participated in the intervention.

In designing an RCT of a complex intervention, one can consider the use of several types of comparator options. Conditions compared to the experimental intervention might comprise, for example, an alternative standard treatment, a placebo, or a waiting list (Anderson, 2008). Each option involves trade-offs and depends on the circumstances of the particular trial. Wait-list comparators, for example, enable all participants to eventually receive the intervention and can therefore improve enrollment and retention, but they also preclude longer-term tracking of outcomes with a counterfactual beyond the wait-list period. With respect to alternative interventions, researchers like Howick (2009) in the biomedical field

have made arguments to the effect that active comparison trials—those involving some form of alternative intervention to the experimental intervention—are methodologically superior to no-treatment or placebo controlled trials because they are likely to recruit larger sample sizes and because it is more valuable to understand how an intervention compares to alternative or established practice rather than to none at all.

The control group condition determined for this study involved a four-month wait-list with the offer of a basic recreational ‘club’ that provided two-hour activities on a biweekly basis. Immediately following posttest, control youth were invited and encouraged to begin participating in QS NFE. Clubs were implemented by trained Questscope volunteers—‘club coordinators’—from the University of Jordan. Club activities were designed to offer basic recreational opportunities organized by club coordinators with minimal integration of QS NFE methodology—participatory learning or activities designed to be asset-building. Typical activities included group meals, playing on computers, and simple games. Four clubs were based in community-based organizations, one club was based in a local non-governmental organization office, and one girls’ club was located in a school.

The primary purposes of the clubs were two-fold: to provide what local stakeholders considered to be a more ethical and marketable alternative to complete no-treatment control, and to help maintain communication with control youths, thus strengthening the likelihood of research retention and program enrollment with a population of youth with whom it can be very difficult to maintain

contact. Though little such research has been conducted in the social interventions arena, a review of antipsychotic drug trials comparing attrition in trials with active treatment controls to those with placebo controls found significantly higher dropout rates in the latter (Kemmler et al., 2005). Although the control offers a modest aspect of 'active comparison,' the club served more to minimize attrition through continued contact than to give a comparable basis for assessing an empowerment-based intervention against a non-empowerment-based intervention, given the relatively low frequency of the clubs (as well as overall low uptake of club activities, reported in chapter five).

The decision to have a four-month wait-list and period between pretest and posttest was a particularly difficult one for the research team. In a review on positive youth development in the USA conducted by Catalano and colleagues (2004), 80% of the twenty-five programs found to be effective lasted for at least nine months. Yet, in conducting the systematic review on YEPs discussed in chapter two, this author identified at least five experimental or quasi-experimental evaluations of youth development programs that detected statistically significant impacts on psychosocial outcomes in four months or less (all of which were completed after the Catalano and colleagues review) (Grolnick et al., 2007; Lakin & Mahoney, 2006; Olson-Merichko, 2005; Shelton, 2009b; Winkleby et al., 2004).

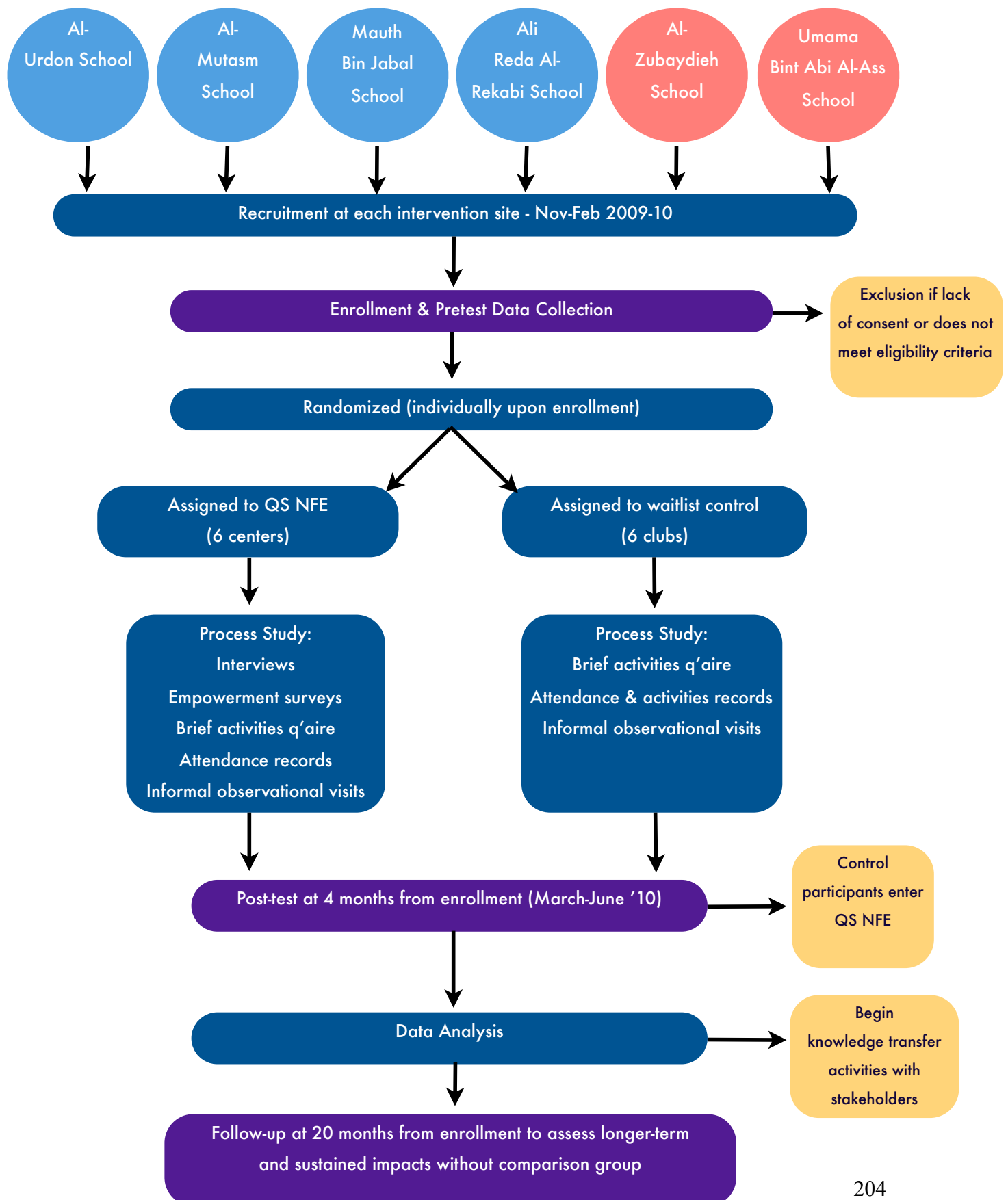
Moreover, Questscope conducted a non-randomized impact assessment of its mentoring program for youths referred from the Jordanian juvenile justice system in 2006. The evaluation ($n = 310$) compared two groups of youth who received

different levels of mentoring and found a statistically significant difference between high and low dosage groups in measured psychosocial outcomes by three months (Al-Khouli et al., 2006). Questscope expected comparable results for the QS NFE study, given higher overall dosage and added components of peer interactions, group activities, and educational programming.

The principal investigator also met with seven facilitators to inquire about the point at which they expected to see noticeable changes in the outcomes of interest based on previous experience; six out of seven expected to observe changes by four months. The facilitators and Questscope staff also projected significant attrition with a longer wait-list control group. Given the evidence indicating the plausibility of detectable effects within four months, the research team determined that the risk of premature posttest data collection was less than the risk of control group attrition and lack of facilitator buy-in.

Funding and staff permitting, further follow-up data collection is planned for twenty months following enrollment, but this will not involve a control group and will not be included in this dissertation. Nevertheless, further follow-ups can help to assess sustainability of any four-month effects as well as detect possible 'sleeping effects' (Eccles & Appleton Gootman, 2002, p. 168) that could materialize after participants have had more exposure to the program or simply more time for effects to materialize with or without further program exposure. A flow diagram of the overall research process is included in figure 4.4.

Figure 4.4. Research process diagram



Protocol

A research protocol for the RCT was published online prior to recruitment on ClinicalTrials.gov—trial identifier, NCT01001286—a service of the United States National Institutes of Health (Morton, 2009b).

Ethics

The research project was reviewed and granted approval by the University of Oxford Social Sciences & Humanities Inter-Divisional Research Ethics Committee in October 2009. Local approval was obtained from the Jordanian Ministry of Education. Youth and parental consent was obtained in a written form (in Arabic) given to all participants; the English-translated form is included as appendix 4.1. Prospective participants had the information verbally explained to them at least twice—once during recruitment (usually by a facilitator) and again at enrollment by a research staff person prior to administering the pretest survey.

All youths were informed that they would not be penalized for refusing participation and they could discontinue participation at any time without having to give a reason. All personal data is kept securely and confidentially. Participants were assigned a random four-digit personal identification number, which was used in place of names for outcome surveys in order to help ensure anonymity of data. All research staff and facilitators at research sites were given training and written guides for ethics procedures as well as other research implementation procedures.

Funding

This research was funded by a grant from Foundation Open Society Institute (Zug.). The funder had no known stake in particular outcomes and did not interfere in any way with the research process or place additional demands on the project from what was initially proposed by the research team.

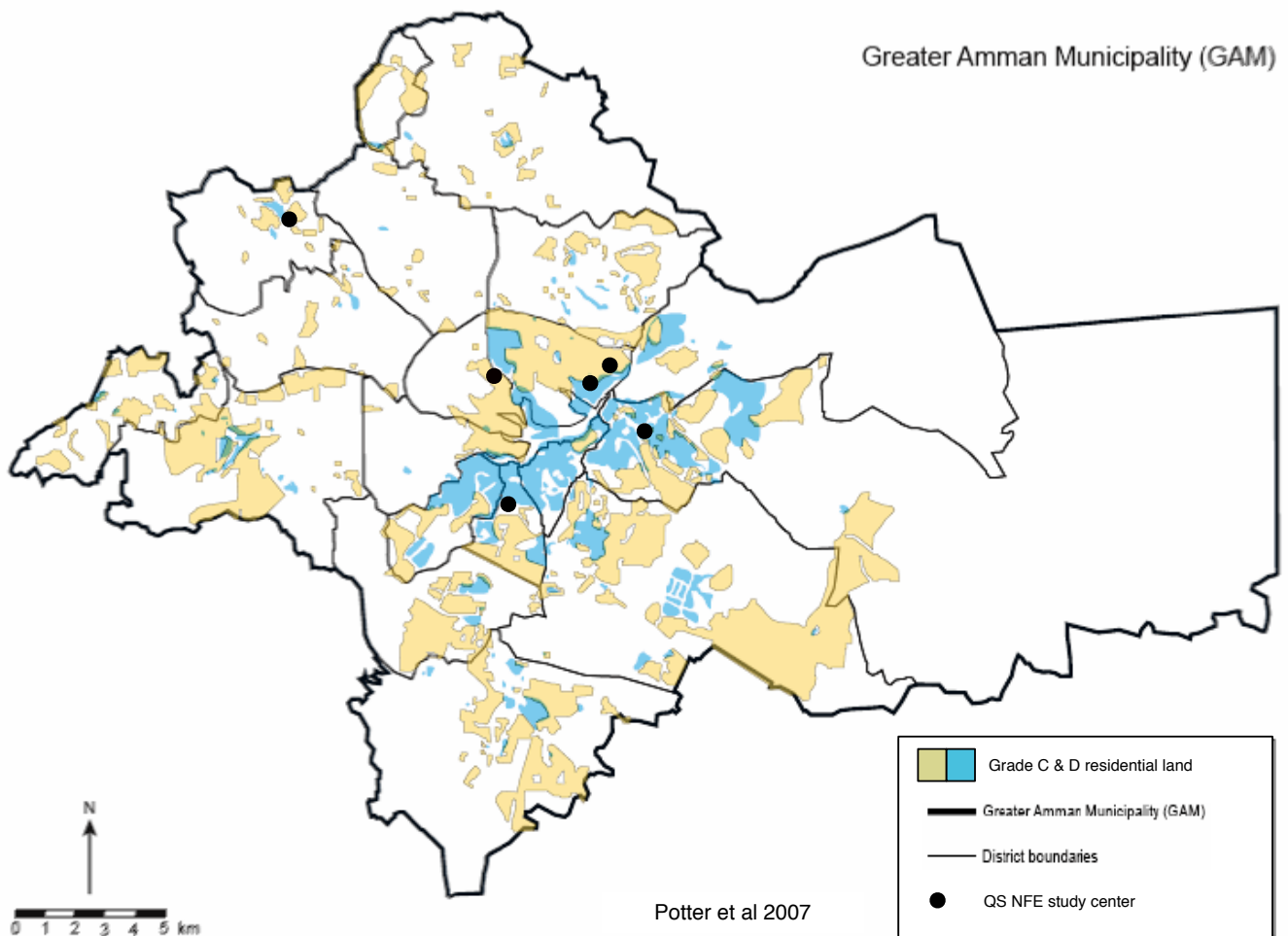
Effectiveness versus efficacy

The RCT was planned as an evaluation of intervention effectiveness rather than efficacy. Efficacy trials evaluate a treatment or intervention under ideal circumstances. This involves researchers in controlling (if not designing and staffing) the intervention in order to ensure strict adherence to intended intervention implementation and intended uptake (including dosage) among participants. Effectiveness trials, on the other hand, assess intervention effects under 'real world' circumstances (Gartlehner et al., 2006). Glasgow and colleagues (2003) argue that a reliance on efficacy trials inhibits knowledge transfer to policy and practice because the results of efficacy trials are less likely to match the results that would be produced by an intervention when it is actually diffused. The MRC framework reinforces the point with the assertion that "[a] key question in evaluating complex interventions is whether they are effective in everyday practice" (Craig et al., 2008, p. 980). As such, this research interfered as little as possible with normal intervention recruitment and implementation procedures so as to evaluate the program in its 'real world' state.

Multisite selection and preparation

This was a multisite impact study involving six QS NFE centers in Amman, Jordan: four male centers and two female centers. All centers were located in low-income communities with generally high concentrations of child labor. While credible, micro-level poverty data on income could not be identified, the relative deprivation of study site areas is reflected in the map provided in figure 4.5, which locates the six QS NFE centers in East Amman neighborhoods and other areas of Greater Amman in which lower grades of residential property are concentrated (indicated by color-shaded areas). Potter and colleagues (2007) explain that these areas are “characterized by low-income residents living at high population densities.”

Figure 4.5. QS NFE study center locations



Certain criteria were established in order to select the six centers from the forty possibilities; these criteria were based on research-related practicalities rather than choosing centers based on perceived program quality. The criteria determined by the research team to select QS NFE centers for the study included the following:

- *Location* – Only centers in Amman (the capital city and the country’s largest) were included in order to ensure easy and fast accessibility for the research team.
- *Setting* – While there are seventeen first-cycle programs based out of community-based organizations, the team determined only to include centers based out of school settings given that there are more of them, they involve an added assurance of long-term relevance and sustainability with the strategic support of the Ministry of Education, and it would reduce the heterogeneity of program implementation across sites. Moreover, the continuity provided by school-based centers, which host all three programming cycles, makes it easier to track youth outcomes beyond the first cycle.
- *Recruitment* - Several QS NFE centers at present predominately serve Iraqi refugees living in Jordan. The research team determined to include Jordanian participants exclusively in the study, given the volatility of Iraqis’ living situation and very different life circumstances. Consequently, centers that Questscope staff determined to be predominately focused on enrolling Iraqis were excluded.

- *Gender* – Although females constitute a smaller portion of dropouts than males in Jordan, resulting in a smaller number of female QS NFE centers, the research team decided it was important to assess the impacts of QS NFE on both gender groups. As such, two of the six centers that met the above criteria were selected as the only remaining female centers.

Multisite designs can offer important advantages with regards to increasing the external validity and rapid recruitment of randomized evaluations (Kraemer & Robinson, 2005; Weinberger et al., 2001). Yet, fractured RCTs can also pose threats of fractured research implementation, potentially producing misleading results (Kraemer & Robinson, 2005; Weinberger et al., 2001). This research took steps recommended by Weinberger and colleagues (2001) and Kraemer and Robinson (2005) to minimize such threats, including a standard research protocol and measures for all sites, a central data analysis system, an overarching research coordinating team, and centrally-based and trained staff responsible for all enrollment and data collection. Additionally, a full-day training on the research purpose and procedures was provided for all facilitators, staff, and partners involved in the study prior to recruitment.

Sample size calculation

A power calculation based on the study's primary outcome, general self-efficacy, was performed to determine a sample size goal. Figures were generated from a randomized controlled trial ($n = 347$) in Germany involving a ten-week cognitive behavioral intervention with eighth-grade student participants (Possel et

al., 2005). The study used the same ten-item self-efficacy scale used in this evaluation of QS NFE. To detect a mean intervention group increase of 3.57 points ($SD = 4.12$) in general self-efficacy based on a 10 to 40 point scale as observed in Possel (2005) with more at-risk youths compared with an assumed natural increase of 1.0 point ($SD = 6.17$) for control with two-tailed 5% significance level and a power of 80%, a necessary sample size of 72 participants per group (144 total) was estimated, accounting for an anticipated dropout rate of 10%. The Sampsize™ online calculator was used (Glaziou, 2003-2005). Notably, such power calculations involve significant assumptions, particularly concerning a program area and target population like the one in this study for which so little prior rigorous impact evaluation has been done (Schulz & Grimes, 2005). The calculation serves as a useful guide, but the actual sample size needed may vary according to different interventions, populations, and other circumstances.

Recruitment

Facilitators were asked to continue their normal recruitment procedures with those recruited as study participants with three exceptions: explaining the informed consent form (including the 50%-50% possibility of starting QS NFE right away or waiting four months to begin while having a basic recreational club); arranging a specific day to come to the center (usually within the week) for research enrollment, assignment, and pretest; and having extra support and encouragement from the research staff to meet sample size targets. Study inclusion criteria included the same criteria for QS NFE recruitment generally (must have been out of formal education for at least two years and 13 to 18 if male or 13 to 21 if female), as well as

Jordanian nationality, not having participated in QS NFE before, and youth and parent informed consent.

Prior to recruitment, individual sample goals were established for each QS NFE center. These goals were based on Questscope staff determinations and supported by previous recruitment trends and in-person conversations with facilitators at each of the centers included in the study. The research team gave four months for the recruitment period (November 2009 through February 2010). The total enrollment goal established was 200 participants—100 per group. The larger sample size goal was established in order to open the possibility of modest subgroup analyses (e.g., according to gender, age, and attendance). As the next chapter reports, 127 participants were ultimately enrolled.

Randomization

Each participant was randomized on-site immediately following enrollment and pretest at the school, in order to avoid biases associated with post-randomization pretest (Schochet, 2010). Unblinded randomization was conducted using an internet randomization program through SealedEnvelope™ (SealedEnvelope, 2000-2010). SealedEnvelope used random permuted blocks in order to ensure that the two groups were balanced periodically. Four-digit personal identification numbers were used to randomize each participant. The results of each randomization were automatically recorded online and emailed to the principal investigator to ensure that no assignment could be changed after randomization.

All participants recruited with a family member (sibling or cousin) were block randomized together with the family member(s). Block randomizing family members aimed to minimize contamination effects and avoid unnecessary restraints on recruitment given the cultural closeness of family ties and tendency for many young people to enter commitments with other family members. Such block randomization of siblings has been conducted elsewhere with similar justification to minimize contamination, such as in the experimental study of the Perry Preschool program (Schweinhart & Weikart, 1998). This contributed to a modest imbalance of numbers between the two groups (n of 67 for intervention and n of 60 for control). A total of twenty-three participants were block randomized with family members.

Data collection

A self-report survey was used to assess participants' baseline (pretest) scores on outcome measures, and the same survey was administered to assess changes in outcomes four months following enrollment (posttest). Each educational cycle is divided into two parts. The posttest corresponds with completion of the first part of the first cycle. This is intended to be a particularly important stage of the program for cultivating the social connections, competencies, and efficacy of marginalized youths.

Pretest surveys were conducted immediately prior to randomization. Randomized controlled trials of psychotherapy interventions and the youth development program, Teen Outreach, that have conducted surveys after group assignment, have shown baseline differences in outcome variables between

intervention and control groups favoring the intervention group despite random assignment and comparability on demographic variables (Allen et al., 1997). The investigators suggested that these differences could have been due to quick spurts in outcomes resulting from early experience with the intervention or even anticipation from knowledge of group assignment (Allen et al., 1997, p. 737). This problem can be avoided by changing the sequence to ensure conduct of pretest prior to assignment.

Given that many participants have low or no literacy, an audio-visual computer-based survey administration scheme was devised. The electronic survey was developed and administered using Creative Research System's The Survey System™ software. The full survey included eighty-eight items including three demographic items; all answer options were multiple choice. Survey completion time generally ranged from twenty-five to forty-five minutes. Each study participant received a small appreciative gift (e.g., football, wallet, scarf, or make-up kit) after pretest and posttest that ranged from five to ten US dollars in value.

The interview stations were based on portable netbook computers (miniature laptops). Participants were presented with each item in Arabic visually as well as in audio via headset. The system required minimal technical ability; participants could choose whether to select answer options using a mouse or a number pad. Almost all participants elected to use the mouse. This was encouraged because the mouse option required moving the cursor to select the answer options presented on each new screen. Because, in theory, a number pad button could be

clicked repeatedly for each option, the research team reasoned that the mouse option would increase the likelihood of thoughtful responses.

Using The Survey System enabled illiterate youths to complete the entire survey independently without the potential biases (e.g., telling the interviewer what the participant thinks s/he wants to hear or variations in how the survey is administered from one participant to the next) and confidentiality concerns involved with having the survey verbally administered by a research staff person. A research staff person was present to orient each youth to the system and provide any clarifications needed during the survey.

Previous studies comparing electronic surveys to paper surveys for public health trials in Sub-Saharan Africa have found electronic survey respondents to give more accurate data, to give fewer missing responses, and to be more likely to feel their responses were confidential compared with paper survey respondents (Mukoma et al., 2004; Palen et al., 2008). Further research found that, between electronic survey options, laptop-based surveys required less time for completion and involved even fewer entry errors than surveys administered from handheld devices (Haller et al., 2009), which informed the research team's decision to use netbooks.

Contamination

Contamination in a trial occurs when members of the control group are exposed to the intervention or intervention group or when members of the intervention group are exposed to the controls. The former situation is more often

problematic in experimental studies (Slymen & Hovell, 1997). Contamination can also occur when control group members participate in alternative interventions that could rival the effects of the intervention under study. Contamination can counteract the purpose of randomization, and sufficiently high levels of contamination effects can significantly reduce a trial's power to detect intervention effects (Torgerson, 2001). The target population for QS NFE has relatively less susceptibility to contamination effects than youth populations in many other program studies; out-of-school youths do not share an educational setting and generally lack other programming opportunities through which they can interact. Contamination was further minimized by block-randomizing family-related youths together. While contamination otherwise could not be controlled, it could be accounted for. On the date of posttest for each participant, a research staff person verbally administered and filled out a brief activities questionnaire consisting of two items:

1. Could you please briefly describe any work activities you have participated in over the last 2 weeks? How many hours per week have you spent in these activities?
2. Could you please briefly describe any community programs, trainings, or education activities (including QS Clubs or NFE) you have participated in over the last 2 weeks? How many hours per week have you spent in these activities?

The focus on the previous two weeks was intended to make it easier for participants to remember what they had done and therefore improve the accuracy of feedback. Responses were collected and examined to gauge whether any substantial contamination effects could have been present.

Primary analysis

Statistical analyses were conducted using PASW Statistics 18.0™ (formerly SPSS) and R Console™. Deviations in pretest scores from a normal distribution were assessed using the Shapiro-Wilk test and by observing Normal Q-Q plots. Kolmogorov-Smirnov is considered an especially tenuous, oversimplified test of normality (D'Aogostino, 1986). Baseline differences between groups were assessed using t-tests and corroborated by Mann-Whitney U tests, given suspect applicability of parametric tests to data for some measures.

Analysis of covariance (ANCOVA) was used to assess intervention effects. The procedure is frequently recommended for small, randomized trials, as the reduction of variability via adjusting for covariates tends to decrease error and increase statistical power, thus reducing necessary sample size for yielding significant effects where they exist (Norman, 1989; Tu et al., 2005; Van Breukelen, 2006; Vickers & Altman, 2001). Vickers and Altman (2001) champion ANCOVA as the preferred choice in randomized trials with baseline data, as it can adjust for random baseline imbalances. Recent randomized trials of psychosocial interventions published in academic journals have taken heed in their respective analysis strategies (e.g., Gardner et al., 2006; Hogarty et al., 2004; Hutchings et al., 2007). ANCOVA was

conducted separately for each outcome using PASW. With posttest scores as the dependent variable, group (intervention or control) constituted the fixed factor for assessing intervention effects, and baseline scores were included as covariates. No covariates' means were significantly different between groups (Miller & Chapman, 2001).

In order to ensure proper application of parametric tests, R Console was used to conduct diagnostic tests on normality of residuals and constant variance to determine whether the data met the requisite assumptions. Visual diagnostics of Normal Q-Q plots and Residual-Fit scatter plots were preferred over statistical tests of normality due to the demonstrated tendency of these statistical tests to exaggerate deviations from a normal distribution for larger sample sizes (e.g., above $n = 80$) (Field, 2009, p. 135; Lumley et al., 2002). Where normality was suspect, Box-Cox transformations were conducted on the data in R Console. Box-Cox uses the log-likelihood value to transform data to meet parametric assumptions. The Box-Cox transformation is defined by the equation $T(X) = (X^\lambda - 1)/\lambda$ where X is the variable being transformed and λ is the transformation parameter (Box & Cox, 1964). Transformation and primary analysis procedures were supported by independent statistical counsel provided by the University of Oxford Statistical Consultancy Service. Where transformations were inadequate or parametric assumptions otherwise suspect, tests were repeated using the non-parametric Mann-Whitney U test to compare group differences at posttest.

Tests were two-tailed and analysis followed a $p < .05$ significance level standard. Considering that past trials have found some youth interventions to cause unintended negative effects on participants, a one-tailed analysis would have been unjustified despite advantages for statistical power (Dishion et al., 1999; Wiggins et al., 2009).

The analysis is intention-to-treat (ITT). The investigator designed the survey so that each item must be answered to proceed to the next, meaning that there can be no missing data in terms of partially completed surveys. Any missing data for posttests due to withdraw of participants were replaced by the 'last observation carried forward' (LOCF) procedure. In other words, posttest scores were substituted by pretest scores for cases that withdrew from the study (either actively or passively). Notably, despite common use in other psychosocial trials (e.g., Flannery-Schroeder & Kendall, 2000; Hutchings et al., 2007) and the practicality of LOCF in randomized trials (Higgins & Green, 2009, section 16.1.2), the potential problems, particularly concerning bias that could contribute to a type I or type II error, with utilizing LOCF are prominently displayed in medical and psychiatric academic journals (Altman, 2009; Hamer & Simpson, 2009; Lane, 2008; Streiner, 2008). Indeed, any choice of method for accounting for missing data in analysis involves vulnerable assumptions, and that may be particularly true of LOCF.

Nevertheless, given a lack of longitudinal data collections and complex demographic variables for more sophisticated statistical imputations, and given the Cochrane Collaboration's preference of LOCF over 'completer analysis only,' the

investigator opted to use LOCF (Cochrane Schizophrenia Group 2010). Using mean scores to complete data would have been at least as presumptuous by assuming that those who withdrew from the study had the same outcome experience as those who did not. As Liu and colleagues (2006) point out, the best strategy for ITT is a proactive one, which includes continuing to collect data wherever possible for those who have been inactive in either group. To this end, our team went to great lengths to collect actual data for ITT regardless of participant activity.

Subgroup analysis and moderators and mediators

The possibilities for subgroup analysis considering the relationship of different variables to intervention effects were limited by a relatively small sample size. Rothwell (2005) explained, “[I]f a trial has 80% power to detect the overall effect of treatment (not uncommon), reliable detection of an interaction of the same magnitude as the overall effect (i.e., potentially clinically important) would need a four-fold greater sample size.” Given the impracticality of such a sample size standard for many trials, Cook and colleagues (2004) suggested that the results of subgroup analysis be treated more humbly as hypotheses, a notion that was followed in the present analysis and discussion thereof. Subgroup analyses were conducted on change scores, as statistical equivalence at baseline could no longer be assumed.

As Gardner and colleagues (2010) explained, in order to distinguish moderators from predictors, analysis needs to show that the moderator is associated differently with outcomes for the intervention compared to control

groups. To this end, subgroup analyses with the intervention group were compared with parallel analyses with the control group to examine whether or not any findings of interest were unique to the intervention group, and regression analyses that were performed to test moderation included group in the model and interaction term.

Subgroup analyses were conducted to explore possible relationships of age, implementation quality, and attendance to intervention effects. In terms of age, youths at different age levels may respond to the intervention differently, and the intervention may respond to them differently—e.g., facilitators might be more inclined to court older youth into empowering opportunities such as leadership roles. Either way, the age variable could have a moderating role with intervention effects. An RCT ($n = 457$) of the Woodrock Youth Development Project in the United States, for example, found positive intervention effects on interracial relations, self-esteem, and substance abuse knowledge for six to nine year-olds, but not for ten to fourteen year-olds (LoSciuto et al., 1997).

For the purpose of subgroup analysis, age was converted in the present study to a dichotomous variable based on the median split (16) of the sample, thus comparing a younger age group (13 to 15) to an older age group (16 to 21). Furthermore, the parameters of these categories are theoretically reinforced by previous research examining age differences in psychosocial functioning according to early-middle and late adolescence. In a sample of 312 American school youths, for example, Gavin and Furman (1989) found significant differences in responses to

behavioral survey scales between early-middle adolescents and late adolescents (the latter including 11th and 12th grades; the typical starting age for 11th grade is 16). In particular, the study found decreased need for group membership, conformity to group norms, and negative behavior in late adolescence coupled with increased positive behavior. Following earlier studies, the findings led Gavin and Furman to hypothesize the following:

Coupled with other research indicating that adolescents become more autonomous and self-reliant as they get older (J. C. Coleman, 1974; Steinberg & Silverberg, 1986), these findings suggest that it becomes less important for adolescents to take advantage of the status and provision of "belongingness" that a group provides. Instead, they may look more toward individual relationships (same or opposite sex) to fulfill their needs. (p. 832)

As such, older youths may respond less sensitively to a group-based intervention that could play a role in shaping peer group norms and dynamics. On the other hand, older youths may respond more sensitively to individualized, mentor-like relationships with facilitators as well as the opportunities for autonomy and challenge embedded in empowerment processes. In either respect, age group may constitute an important moderator variable for intervention effects.

To some extent, however, these parameters are nevertheless heuristic. The onset of stages of adolescence (and adolescence in general) from a physiological perspective can vary according to the timing of puberty, which can have repercussions for mental health (Kaltiala-Heino et al., 2003). Social implications of culturally-specific definitions (either explicit or implicit) of youth and adolescence can also render the meaning of age group boundaries malleable (Hill & Dennis

Fortenberry, 1992). No studies specific to Jordanian youth were found by this author in order to empirically refute or reinforce the age group demarcations delineated in the abovementioned research; it is conceivable that implications of specific age group categories could be different for Arab or Jordanian culture.

Despite low statistical power, for exploratory purposes, hierarchical multiple regression was also used to investigate potential moderation roles of age in program effects. Age was dichotomized into 'age group,' a binary variable including younger youth (aged 13 to 15) and older youth (aged 16 to 21). Notably, a number of statistical papers advise against converting continuous variables into categorical variables, as this loses information contained in the continuous measurement and tends to substantially reduce statistical power (Cohen estimates by about one-third) and effect sizes (Altman & Royston, 2006; Cohen, 1983; West et al., 1996). The reason age was dichotomized in this case was because participants were heavily skewed toward the median (16) and younger age group. Because regression assumes a normal distribution, this abnormal distribution (shown in appendix 4.2) without a large sample makes it especially difficult for the regression analysis to detect potential age effects, particularly in terms of potentially important differences between those near to median (e.g., 14 and 17-year-olds). Nonetheless, given that neither arrangement is statistically ideal, the regression was also run with age as a continuous variable to report any differences, but the emphasis is on age group as the moderator.

A separate regression was conducted for each outcome. The outcome change scores constituted the dependent variable. In Step 1, experimental group was included as the predictor followed by age group as the potential moderator. In Step 2, the same variables were included along with the introduction of the interaction term (experimental group x potential moderator). Dummy codes (0 and 1) were created for the experimental group variable, and weighted effects codes ($-n_2/n_1$ and 1) were created for the age group variable due to uneven numbers between the two groups (West et al., 1996). In order to reduce multicollinearity, continuous predictor variables were centered before including them in the regression or creation of the interaction term (Frazier et al., 2004). Effect sizes were measured by change in the variance (ΔR^2) due to the introduction of the interaction term (Frazier et al., 2004; Gardner et al., 2010). Any statistically significant interaction effects would be explored by comparing simple slopes using ModGraph-I™ (Jose, 2008) to investigate associations between the predictor and outcome at different levels of the moderator (Frazier et al., 2004).

Multisite RCTs also pose the possibility of relationships between variations in program implementation and intervention effects (Brieger et al., 2001). Utilizing scores from the implementation study's adapted Learner Empowerment Survey described in chapter six, this factor could be modestly explored in the analysis. Given the low sample sizes recruited by female QS NFE centers, the subgroup analysis on implementation compared outcomes for the male center with the highest average empowerment survey scores to the male center with the lowest average empowerment survey scores.

Finally, youth development programs have reported important dose-response relationships with analyses indicating a positive association between the frequency levels of youth participation and outcomes (e.g., Walker & Arbreton, 2001). This relationship was explored in the present study using attendance data from official records entered by QS NFE facilitators. Because the Ministry of Education and Questscope hold two days per week as the minimum expected attendance for participants, this analysis accordingly grouped attendance dichotomously (high or low attendance) by participants with less than two days per week average attendance between pretest and posttest and those with two days or more per week average.

Previous trials have also found that, in some instances, girls and boys can respond differently to youth interventions (Rhodes et al., 2008). Conceivably, this could be a particularly important issue in the Arab world, which tends to foster more distinct social roles and relationships for males and females (Booth, 2002). Due to low recruitment of females, however, little meaningful analysis, even on an exploratory level, could take place with respect to the gender variable.

Questionnaire

The order of questionnaire items was changed where authors of the instruments recommended that items be randomly ordered (the General Self-Efficacy scale and Strengths and Difficulties Questionnaire items were randomly ordered and interwoven). The questionnaire involved a three-option Likert scale for all items. The measures included originally provided answering scales with three to

five response options. Previous Questscope impact evaluation with mentoring used measures with outcomes providing only two response options (yes/no). Scales with more options enable respondents to give more accurate answers whereas too many options can confuse respondents with less survey competency.

In consultation with local staff and QS NFE youths, answer scales with four or five answer options were considered to be too complicated. Three, however, were not considered too many and were preferable to two options, which were described as too limiting. Consequently, all items were conformed to a scale of three answer options (for some measures, answer scales provided the options: not true, somewhat true, or certainly true; for others, the options included: hardly ever, sometimes, or most of the time).

In the questionnaire, twelve items from the Strengths and Difficulties Questionnaire and nine items from the Social Support Appraisal Scale were reverse-scored as intended by the instruments' authors. Responses for each item were coded to a uniformed score range from 0 to 2 points with higher scores indicating positive responses (i.e., not true = 0, somewhat true = 1, and certainly true = 2; scores were reversed for reverse-score items).

Translation and adaptation of measures

The measures were independently translated and back-translated into Arabic by a professional translator and checked by a bilingual research staff member. This process follows procedures recommended in the literature for improving linguistic equivalence of measures across cultures (Church, 2010).

Instruments that had already been translated into Arabic for use in previous studies (the General Self-Efficacy Scale and the Strengths and Difficulties Questionnaire) were retranslated and compared in order to ensure consistency with Jordanian dialect. The translated questionnaire was reviewed on paper with a convenience sample of thirty youths currently participating in QS NFE. The research team used feedback from these youths to refine the questionnaire to its final version.

Arabic language encompasses a great deal of lingual heterogeneity both in terms of formality and informality as well as regional and cultural differences (Bateson, 2003, p. viii). The youth feedback was particularly important for refining the instrument for 'semantic equivalence' (Suhonen et al., 2010) of the items to ensure that the given Arabic clearly communicated the intended meaning to the target population of youth. With respect to content relevance, the Questscope staff and youth feedback indicated that three items, one on the SDQ emotional symptoms subscale and two on the SDQ hyperactivity subscale, were difficult for the youth to relate to and were not reflective of the intervention's theory of change; these were excluded in the final instrument. Excluded items included the following: "I get a lot of headaches, stomach-aches or sickness;" "I am restless, I cannot stay still for long;" and "I am constantly fidgeting or squirming." The modified version was field-tested once more with eleven QS NFE youths, this time on the computer-based system. The research team considered feedback at that stage an indicator of a satisfactory instrument. The final questionnaire is included in English as appendix 4.3.

Measures

This study used a variety of measures to evaluate psychosocial constructs included in the intervention's theory of change. Rather than only assessing a single primary outcome, the MRC's Framework for Complex Interventions suggests using a range of theoretically relevant outcomes in order to capture effects in different constructs that may be impacted, as well as to detect possible unintended consequences (Craig et al., 2008). A summary table of the instruments is provided in table 4.1. More detailed descriptions and justifications follow.

Primary outcome

General Self-Efficacy

The General Self-Efficacy (GSE) scale is a ten-item measure (Schwarzer & Jerusalem, 1995). The scale was translated, adapted, and validated in twenty-five countries, including neighboring Syria. Cronbach's α for Syria was 0.79, and 0.88 and 0.86 for the United Kingdom and United States, respectively (Scholz et al., 2002). The mean age for this international study was twenty-five. The scale was designed for the "general adult population, including adolescents" (age 12 and above)(Schwarzer & Jerusalem, 1995). Self-efficacy refers to "beliefs in one's capabilities to mobilize the motivation, cognitive resources, and courses of action needed to meet given situational demands" (Wood & Bandura, 1989). Example items in the GSE scale include the following: "I can always manage to solve difficult problems if I try hard enough" and, "It is easy for me to meet my needs and accomplish my dreams."

Table 4.1. Instruments

Instrument	Purpose	Description	Adaptation
General Self-Efficacy scale (Schwarzer & Jerusalem 1995)	Assesses belief in one's ability to meet challenges and demands across a wide range of situations.	10-item self-report version with 4-option Likert scale. Intended for general adult population, including adolescents (ages 12+).	Minor wording in translation, answer scale reduced to 3 options.
Social Skills Questionnaire (Spence 2003)	Measures range of social skill areas including cooperation, self-control, initiative, communication and expression.	30-item self-report (pupil) version with 3-option Likert scale, intended for youths.	Minor wording in translation.
Social Support Appraisals Scale (Dubow & Ullman 1989)	Assesses quality of supports from friends, families, teachers, and classmates.	41-item self-report measure with 4 subscales using 5-option Likert scale. Intended for youths.	Minor wording in translation, answer scale reduced to 3 options, only used 2 subscales (friends and family; 22 items).
Strengths & Difficulties Questionnaire (Goodman 1997)	Brief behavioral screening that assesses psychological difficulties in peer problems, conduct, emotional symptoms, and hyperactivity as well as prosocial strengths.	25-item self-report (student) version with 5 subscales and an overall difficulties scale using 3-option Likert scale, intended for youths.	Minor wording in translation, 2 items omitted from hyperactivity subscale and 1 from emotional symptoms subscale.
Other Adult Connectedness (Blum & Ireland 2004)	Brief scale indicating how much youths believe neighborhood adults, friends, teachers, and religious leaders to care about them.	4-item self-report measure using 3-option Likert scale, intended for youths.	Minor wording in translation, only 1 item used (neighborhood adults).
Learner Empowerment Measure (Fymier et al., 1996)	Implementation measure assessing youths' perceptions of meaningfulness, competence, and impact in a class or program environment.	30-item self-report measure using yes/no answer options. Original intended for college students, adapted by Lakin & Mahoney for adolescents.	Used adapted version by Lakin & Mahoney (2006), minor wording in translation, added 4 items.

By engaging young people as valued partners, empowering programs aim to improve young people's belief in their own abilities to shape their lives and environments (Jennings, 2006; Wong et al., 2010). Non-empowerment-based programs may 'teach' adolescents valuable skills or concepts, but if the intervention fails to cultivate a participatory environment in which young people are challenged and supported in enacting skills and ideas for themselves, they may be less likely to develop the sense of efficacy to employ them in daily life. Bandura (1997) argued that individuals with problems generally have the cognitive knowledge of what they need to do, but they lack the non-cognitive sense of ability to carry out the desired action.

High self-efficacy has been shown to predict better performance in work, academics, and sports; increased happiness, job satisfaction, and persistence; improved safe sex practices; and successful smoking cessation and prevention (de Vries et al., 1988; Judge & Bono, 2001; Kalichman & Nachimson, 1999; Martin & Gill, 1991; Multon et al., 1991; Natvig et al., 2003). A meta-analysis conducted by Stajkovic and Luthans (1998) found that self-efficacy accounted for a 28% improvement in work-related performance. Lower self-efficacy, on the other hand, appears to predict higher levels of depression among young people (Bandura, 1999) and is associated with higher alcohol consumption (Taylor, 2000).

The topic of self-efficacy measurement has drawn a rift between those espousing domain or task-specific self-efficacy and those championing generalized or global self-efficacy. Proponents of specific self-efficacy argue that judgments of

self-efficacy are context dependent and caution against general measures that allegedly fail to correspond to a particular set of performance tasks (Bandura, 1986; Pajares, 1996). General self-efficacy, however, offers validated instruments that have shown GSE to predict performance and satisfaction outcomes, particularly when the performance of concern involves broader and more complicated tasks (Chen et al., 2001; Gilad et al., 2004). Chen and colleagues (2001) found that past comparisons between predictive powers of general and specific self-efficacy scales have been misleading, and that general self-efficacy measures can be more useful given that jobs and life roles are increasingly broad, complex, and demanding.

Secondary outcomes

Social Skills

The Social Skills Questionnaire (SSQ) uses a thirty-item instrument to measure social skills areas including cooperation, self-control, initiative, communication, and expression (Spence, 2003). Testing of the SSQ demonstrated a Guttman split-half reliability of 0.83 and coefficient α of .85. The instrument may serve as a better measure of social skills than the popular Social Skills Improvement System (SSIS) Rating Scales, as the SSIS includes many items that are less directly related to social skills than to more general aspects of behavior (Spence, 2003). Example items in the SSQ included the following: “I do kind things for other people without having to be asked,” “I tell a parent or older person if I have a problem or need help,” and, “I look people in the eye when we are talking.”

While QS NFE has an obvious mandate to improve participants' academic competencies, the program is principally a social activity. By facilitating supportive relationships with caring adults, positive peer interactions, and opportunities to exercise leadership and communication, QS NFE expects youths to improve in critical social skills. In day-to-day life, these skills help young people effectively navigate relationships, achieve tasks, work in groups, and embark upon personal goals. Notably, previous longitudinal research has found that young people's social skills and competencies explained social and economic success (e.g., wages, employment, and educational attainment) as well as, or better than, academic competencies (e.g., reading, writing, and mathematical abilities) (Heckman et al., 2006). Involving young people in decision-making and positive activities with adults is considered by youth empowerment literature an important mechanism for developing young people's transferable skills (Wong et al., 2010).

Social Supports

The Social Support Appraisals Scale (SSAS) assesses the presence and quality of adolescents' supports from families, peers, classmates, and teachers (Dubow & Ullman, 1989). Because, as out-of-school youths, control participants in this study did not have teachers or classmates, the latter two domains did not apply. Accordingly, the study only included the subscales measuring supports from friends (ten items) and family (twelve items). Cronbach's α for the subscales ranged from .78 to .83 (Dubow & Ullman, 1989). The SSAS examines a youth's perceived availability of support from families and peers, and the extent to which he or she is

esteemed and valued in such networks (Dubow & Tisak, 1989). Example items include, “Some youths and their friends do a lot of things for each other, but other youths and their friends don’t. Do you and your friends do a lot of things for each other?” and, “Some youths are an important member of their family, but other youths are not. Are you an important member of your family?” for friend and family supports, respectively.

At-risk youths tend to lack the positive social bonds that promote resilience and healthy development through the formative years of adolescence. Youth empowerment programs aim to forge stronger social supports around young people by facilitating prosocial interactions between youths and their peers, connecting youths with community adults, and promoting stronger ties with the family. A number of studies over recent decades have documented the role of family supports as a protective factor against problematic youth behaviors and mental health problems (Barnes & Farrell, 1992; Barrera & Garrison-Jones, 1992; Wills et al., 1992). The link between adolescent-family connectedness and better mental health was repeated in the results of a survey administered to 2,893 adolescents representing eight Arab societies (Dwairy et al., 2006).

Yet, while family connections remain important—even more for Arab than American adolescents (Dwairy et al., 2006)—empirical research suggests that the role of peer supports becomes increasingly important relative to family supports as youths get older. Furman and Buhrmester (1992), for example, found in a study of 549 American youths that, while parents were the most important providers of

support to children in the fourth grade, same-sex friends were reported as being as supportive as parents in the seventh grade and were even more frequent providers of support in the tenth grade. Furthermore, peer social support has been shown to be a significant predictor of subsequent mental health improvements (Hirsch & DuBois, 1992). Hurrelmann (1990) found the quality of adolescents' relationships with parents and friends to predict delinquency, aggression, drug abuse, and psychosomatic complaints.

The questionnaire also included a single-item measure to assess social support in the form of connectedness to local adults included in a connectedness scale by Blum and Ireland (2004). The item asked, "How much do adults in your neighborhood care about you?" with answer options, "a lot," "somewhat," and "very little." Local research team members felt the construct was important to the intervention's theory of change, so this item was included as a gauge for possible changes rather than as a measure for which the investigator expected statistically significant effects given low sensitivity. Improving community and local adult connection is a central pillar for several youth empowerment and youth development models (Jennings, 2006; Kim et al., 1998; Villarruel et al., 2003, p. 329; Wong et al., 2010). International research has found higher neighborhood connectedness among youth to correlate with higher self-esteem, improved health indicators, academic achievement, and lower depression (Blum, 2003; Witherspoon et al., 2009).

Strengths and Difficulties

The Strengths & Difficulties Questionnaire (SDQ) is a twenty-five-item behavioral screening that assesses five domains of positive and negative psychological attributes (Goodman, 1997). Domains (subscales) include emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. Data analysis assessed outcomes according to each subscale as well as an overall 'difficulties' measure that encompasses all items aggregately except for the prosocial scale items, which are conceptually different from psychological difficulties (Goodman, 1997). An example of a strength item is, "I try to be nice to other people; I care about their feelings," and an example of a difficulty item is, "I fight a lot; I can make other people do what I want."

Consistent with a study that assessed mental health with children in Gaza, only the one-sided version of the SDQ was used rather than also using the 'impact supplement' with questions on the overall extent, duration, and burden of difficulties on the family (Goodman & Scott, 1999; Thabet et al., 2000). Like Thabet and colleagues (2000), the research team considered this supplement too complicated for the target population in the context of this research and more susceptible to cultural bias than the items in the primary instrument component.

The SDQ has been shown to consistently perform as well or better than both the Rutter questionnaires and the Child Behavior Checklist, despite using fewer items for faster completion (Goodman, 1997; Goodman & Scott, 1999). The self-report version of the SDQ has compared favorably with the teacher- and parent-rated SDQ scores in psychometric testing (Goodman et al., 1998). In testing internal

reliability with 11 to 16 year olds in the United Kingdom, the self-report SDQ yielded the following satisfactory Cronbach's alphas for the included scales: 0.82 for total difficulties, 0.75 for emotional symptoms, 0.72 for conduct problems, 0.69 for hyperactivity, 0.65 for prosocial behavior, and 0.61 for peer problems (Goodman et al., 1998). The instrument also successfully predicted adjustment and psychiatric disorders among children and adolescents (Goodman, 2001).

The price of neglecting early development of prosocial behavior and prevention of antisocial behavior can have alarming consequences for society. A longitudinal study in the United Kingdom, for example, found long-term costs to the public (e.g., via crime and state benefits) of antisocial children with conduct disorder as much as ten times higher than costs associated with youths without such problems (Scott et al., 2001). This is not to mention the lost economic opportunity in youth rendered less effective in contributing productively to the local economy and public good as a result of poor life preparation. Moreover, the SDQ's focus on positive attributes as well as problematic behavior is more appropriate for YEPs like QS NFE, which aim to develop the strengths of adolescents rather than merely minimize problems.

As alluded to above, the SDQ has been tested with Arab youth in Gaza, though the sample tested for the adolescent self-report version ($n = 57$) was much smaller than the samples tested for parent-report versions for younger children (Thabet et al., 2000). Moreover, the primary analysis (for which Cronbach's alphas were provided) was only conducted on the parents' questionnaire forms, mostly for

younger children, and the authors detected cultural differences with the scale to be particularly important for pre-school children. As such, the applications of the results to use of the SDQ in the present study are limited. These caveats aside, the study nevertheless found Cronbach's alphas for the parent version of the SDQ to be "substantially lower than those established in the Swedish sample using the SDQ" (Thabet et al., 2000). Ultimately, Thabet and colleagues (2000) considered the SDQ a "promising" rating scale for other cultural populations, but the study notably highlighted the need to "identify and use indigenously meaningful constructs."

Thabet and colleagues' concluding counsel highlights an important consideration and potential limitation for all of the measures used in this study. While this project aimed to identify some of the most credible and appropriate instruments available for the outcomes of concern in this research, they have all, nevertheless, been created and largely validated in Western societies. Though beyond the scope of this research, further action to develop and test culturally meaningful self-report measures with Arab adolescents (including those of low socioeconomic background) for psychosocial constructs is urgently needed in order to create a more amenable situation for further robust impact evaluation in the region.

Implementation measure

Mixed methods were employed to investigate program implementation, and these are described in greater detail in chapter six. Because implementation was examined as a factor in subgroup analyses for the RCT, this section provides a brief

overview of the empowerment measure on which subgroup analysis for implementation was based.

Among the included and relevant excluded studies discussed in this dissertation's systematic review, only one study reported using an instrument specifically intended for measuring empowerment as an aspect of assessing program implementation. This study was Lakin and Mahoney's (2006) in which an adapted version of the Learner Empowerment Measure (LEM) was used to gauge young people's feelings of empowerment in the classroom-based intervention. An adaptation of the LEM was administered to each of the six QS NFE centers as an assessment of implementation quality focused on empowerment, the primary theoretical basis for the program.

The quantitative, self-report measure is comprised of thirty items devised by Schultz and Shulman (1993), adapted by Fymier and colleagues (1996) for college students, and modified again by Lakin and Mahoney (2006) for adolescents. The measure draws from Thomas and Velthouse's (1990) definition of empowerment as consisting of four dimensions: *meaningfulness* ("the value of a task in relation to one's own beliefs, ideals, and standards"), *competence* (feeling "qualified and capable to perform the necessary activities to achieve the goal"), *impact* (the perceived difference one's contribution makes in the overall group or activity), and *choice* ("the degree to which individuals self-determine their task goals or methods for accomplishing them") (Fymier et al., 1996, p. 183).

Based on factor analysis findings, Fymier and colleagues (1996) combined the impact and choice constructs into one. The overall empowerment measure has demonstrated an alpha of .90 with a college student sample; the reliabilities of the meaningfulness, competence, and impact constructs were .89, .83, and .81, respectively (Fymier et al., 1996). Example items for the three constructs include: “The things I have to do for the group mean a lot to me personally” (meaningfulness), “In the group, I feel confident that I can do what I am expected to do” (competence), and “My participation is important to the success of the group” (impact).

The adapted instrument by Lakin and Mahoney involves yes or no answer options. ‘No’ responses were coded as zero, ‘yes’ responses were coded as one, and five items were reverse-scored. Four items were added by this study’s research team that addressed youths’ feelings of empowerment in matters specific to QS NFE. The English instrument is included in appendix 4.4. The survey was administered verbally by a research staff person to QS NFE sessions with participants following on an anonymous paper version to record responses. Each QS NFE center had one survey administration, which occurred on a day of the week on which facilitators anticipated relatively high attendance.

Chapter 5: Randomized Controlled Trial: Results

Introduction

This chapter describes the results and lessons of a multisite randomized controlled trial (RCT) of a youth empowerment program (YEP) in Amman, Jordan. Questscope non-formal education (QS NFE) is the product of a unique governmental-non-governmental partnership between Questscope for Social Development in the Middle East (QS) and the Jordan Ministry of Education (MOE). Questscope non-formal education serves out-of-school youth typically from low-income communities, ages 13 to 18 for males and 13 to 21 for females. The two-year program consists of three eight-month education cycles based on participatory methodology. The intervention emphasizes supportive relationships between youths and trained adults and a prosocial, youth-driven program environment.

While this impact study includes six QS NFE centers (two female and four male) in Amman, the program operates in forty schools and seventeen community-based organizations spanning eight governorates in Jordan and has enrolled over 7,000 youths since 2004. Graduates receive a tenth grade alternative certificate that enables youth to participate in vocational training and receive government business loans. According to the program model, the first cycle (a.k.a., ‘informal education’) is particularly important for building youths’ trust, confidence, and enthusiasm for learning. This study assesses outcomes for a sample of young people in Jordan ($n = 127$, including 67 intervention and 60 control) after the first four months of enrollment.

Enrollment, follow-up, and dosage

The study had a 58% uptake among those directly invited by a facilitator or research staff member to participate. Approximately 31% of refusals gave reasons related to the study (declining random assignment), and the remaining 69% declined the intervention in general, primarily due to work and family reasons. Recruitment goals for all but one center were unmet (see figure 5.1), highlighting in some cases significant present challenges for the program with recruiting Jordanian youths, especially females. These challenges and potential solutions are explored in the process and implementation study discussed in chapter six. The final sample size approached the minimum goal established from the power calculation ($n = 144$) but nevertheless left the possibility of underpowered results and type II errors in the event of effects that were close to statistical significance.

In total, 127 youths were enrolled in the study and randomly assigned to either begin the intervention, QS NFE, immediately ($n = 67$) or to a four-month wait-list control ($n = 60$) with basic biweekly recreational 'club' activities available to them. Participants were given the same survey at enrollment and again four months following their respective enrollment dates to follow outcomes. Eight participants (6.3%) from the overall sample were lost to follow-up. A flow diagram is provided in figure 5.2.

On average, control youths participated in 1.3 'club' activities during the waiting period, reinforcing the treatment of the comparator primarily as a wait-list control rather than an active comparison. Youth feedback indicated that clubs had

low overall participation because, unlike QS NFE, the lack of daily activities in the club made it difficult for working youths to attend according to their own shifting schedules. Additionally, some control youths preferred to wait for an educational program rather than take time away from work or family demands for purely recreational activities. The clubs arrangement with involvement of volunteers, however, did help to maintain contact with control participants. The intervention group also had lower-than-expected attendance overall, with 52.3% of the group attending less than the two-day per week average expected by the program.

Implications of relatively low attendance are considered in this chapter's subgroup analysis section, and likely reasons and solutions for low attendance are explored in chapter six.

Figure 5.1. Recruitment summary

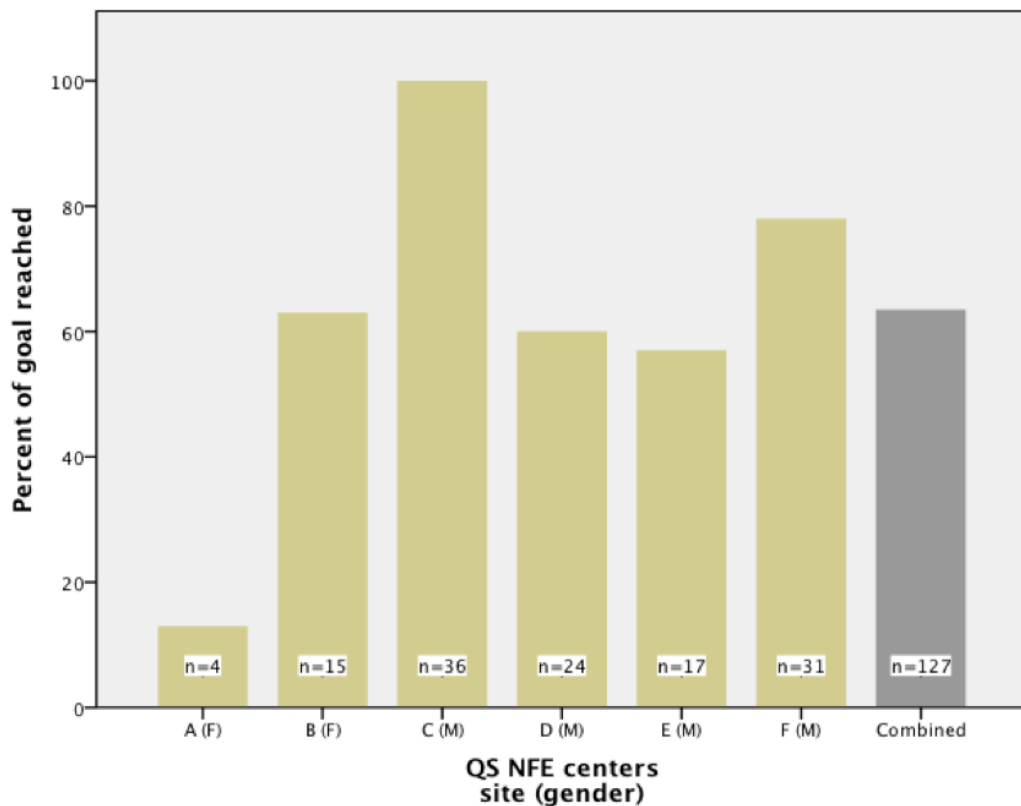
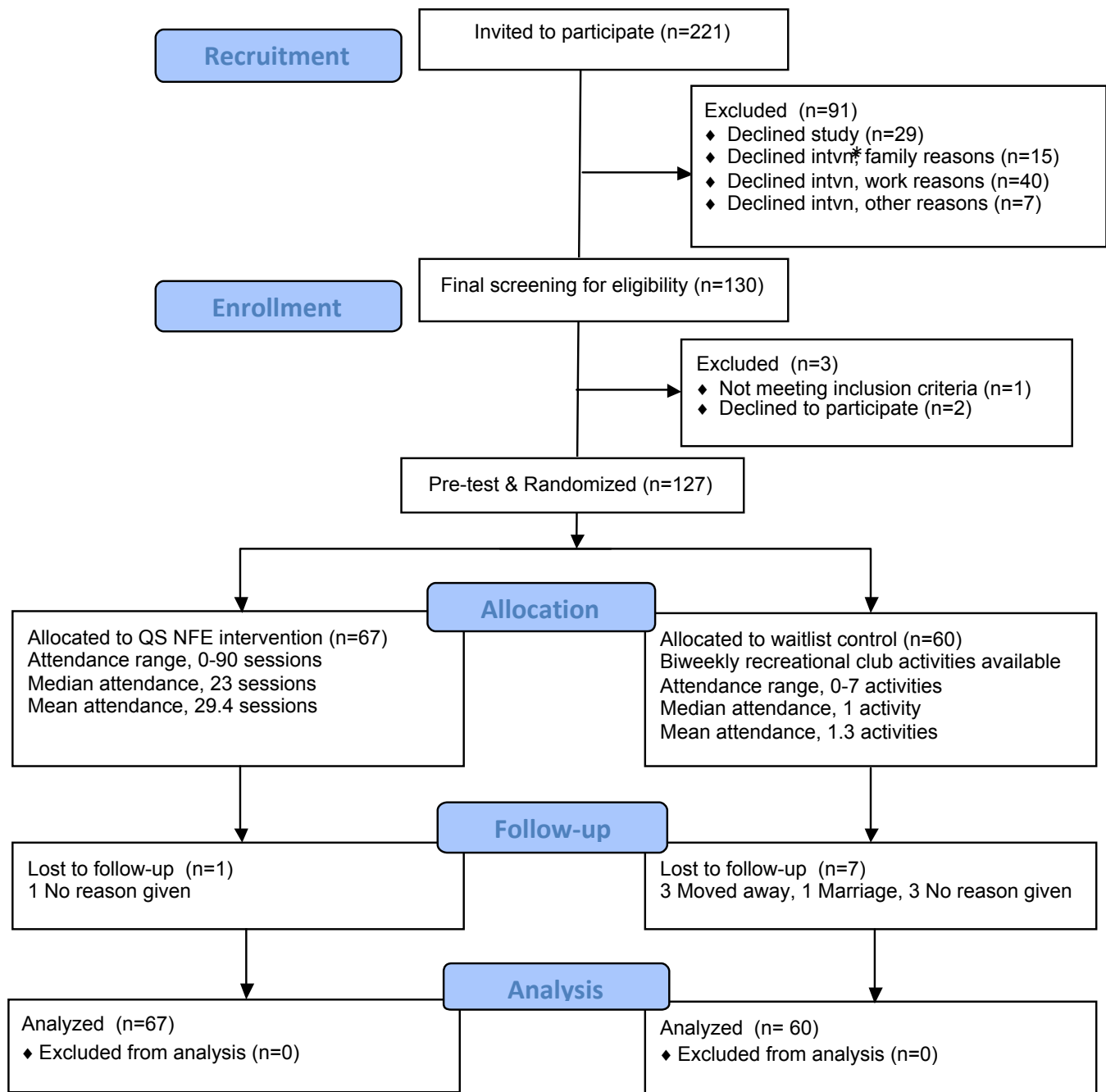


Figure 5.2. RCT flow diagram



*Intvn - Intervention

Preliminary analysis

Baseline comparisons

Although participants were randomized to establish two statistically equivalent groups, a comparison of means was conducted as a precaution to confirm the success of randomization, as recommended by the CONSORT statement for reporting RCTs (Boutron et al., 2008). T-tests showed no statistically significant differences on recorded baseline demographic variables or outcome measure scores. Chi-square tests were conducted for categorical data and also showed no significant differences between groups. The same was true in comparing groups on baseline scores using the Mann-Whitney U test on rank-order data as a non-parametric verification (Mann-Whitney U results are available in appendix 5.1). As such, the preliminary analysis results indicate successful randomization. Descriptive baseline data and *p* values produced by the t-tests are included in table 5.1.

Contamination

Responses to the 'brief activities questionnaire' administered at posttest gave little cause for suspicion of contamination effects. Exceptions comprised of one female control youth who reported working at the school nursery in the same school in which the QS NFE center operated and one male control youth who reported attending classes at the QS NFE center.

Table 5.1. Comparison of groups by baseline demographics and outcome measure scores

[*]	Intervention Group (<i>n</i> = 67)	Control Group (<i>n</i> = 60)	<i>p</i> Value
Demographics			
Age, yr	15.8 (1.6)	16.0 (1.6)	.47
No. (%) female [♣]	10 (15)	9 (15)	.99
No. (%) working [♣]	35 (52)	32 (53)	.90
Baseline scores ^{**}			
GSE	13.3 (4.2)	12.7 (3.9)	.42
SSQ	44.4 (11.6)	42.9 (12.9)	.50
SSAS-Frd	12.9 (3.8)	12.3 (3.7)	.38
SSAS-Fam	16.0 (5.1)	15.6 (4.1)	.67
SDQ:			
Overall difficulties	22.6 (5.3)	21.7 (5.1)	.34
Conduct problems	7.3 (1.8)	6.8 (2.1)	.15
Emotional symptoms	4.7 (2.1)	4.5 (1.9)	.51
Hyperactivity	4.2 (1.2)	3.8 (1.1)	.09
Peer problems	6.4 (1.9)	6.7 (1.9)	.52
Prosocial behavior	7.7 (2.0)	7.9 (1.8)	.44
Adult Connect	0.9 (0.8)	1.1 (0.8)	.26
[*] Unless otherwise indicated, data are given as: mean (SD) [♣] <i>p</i> value for chi-square test given for categorical variables ^{**} GSE – General Self-Efficacy scale, SSQ – Social Skills Questionnaire, SSAS – Social Support Appraisals Scale, SDQ – Strengths and Difficulties Questionnaire			

Internal consistency of measures

Given that none of the measures had been psychometrically tested in previous studies specifically with a Jordanian youth population, all of the study's outcome measures were tested for internal reliability using Cronbach's alphas with pretest scores ($n = 127$). The results from PASW Statistics 18.0 are summarized in table 5.2. The General Self-Efficacy (GSE) scale and the Social Skills Questionnaire (SSQ) showed satisfactory internal reliability (α of .78 and .94, respectively), though a very high alpha for social skills could be largely attributable to the large number of items in the scale (Cortina, 1993; Field, 2009, p. 675). Moreover, a very high alpha (exceeding .90) could be an indicator of redundancy in the questionnaire, rendering some items unnecessary (Streiner, 2003). Alphas for social supports of friends (.66) and families (.73) subscales were lower than Dubow and Ullman's (1989) original reliability test alphas (ranging from .78 to .83) but identical to their retest reliability alphas (ranging from .66 to .73).

The Strengths and Difficulties Questionnaire (SDQ) showed substantially lower alphas than were found with the self-report version in a European sample (Goodman et al., 1998). The overall difficulties measure produced an acceptable alpha of .71, but subscales ranged from .02 to .56. The results for the SDQ corroborated comparable alphas found with the parent-report version tested in Gaza (with a sample that was heavier with younger children). The Gaza sample produced alphas of .32, .50, .55, .18, and .65 for the conduct, emotional symptoms, hyperactivity, peer problems, and prosocial subscales, respectively. The reliability results from the two studies suggest a need for development of more culturally

appropriate behavioral screenings and measures of psychiatric difficulties for Arab youth. The overall difficulties scale did show satisfactory internal reliability, but this could again be largely attributable to a large number of items (Streiner, 2003).

Cortina, for example, demonstrated that a scale consisting of eighteen items from three uncorrelated subscales still produced an alpha of .64 (Cortina, 1993). As such, alpha should be considered with some caution when assessing the reliability of large scales.

Table 5.2. Cronbach's alphas for outcome measures

Scale	Cronbach's Alpha	Items whose omission improves Alpha (Alpha if item deleted)
General Self-Efficacy	.78	none
Social Skills	.94	none
Social Supports - Family	.73	none
Social Supports - Friends	.66	picked on and teased (.68)
SDQ Conduct subscale	.42	do as I am told (.59)
SDQ Emotional Symptoms subscale	.56	none
SDQ Hyperactivity subscale	.02	easily distracted (.54)
SDQ Peer Problems subscale	.29	get along better with older people than with people my own age (.36)
SDQ Prosocial subscale	.54	none
SDQ Difficulties overall	.71	do as I am told (.73), easily distracted (.72)
Local Adult Connectedness	n/a	n/a

Alphas for the hyperactivity and peer problems subscales were particularly low. The hyperactivity subscale may have been compromised with the omission of two items in the cultural adaptation process. One remaining item, however, has a sizeable impact on the subscale's reliability, and, notably, if removed, the alpha is equivalent to that found in Gaza. Peer problems reliability is problematic in both the present study and the Gaza study, and the same item in both studies is particularly so: "I get along better with older people than with people my own age." This could be a neutral indicator. Lack of social connection to peers could be considered negative whereas connection to adults could be positive. Respondents may have been confused on how to answer. An alternative outcomes analysis was conducted without the item and provided in appendix 5.5. The conduct subscale's reliability fared higher than in the Gaza sample but lower than in the European sample. One item whose omission improves alpha suggests that the concept of obedience could be further explored in terms of its cultural application in the Arab world, which places a high premium on authority (Al-Faleh, 1993), as well as in the context of a youth empowerment framework, in which case obedience might not be considered in the same light.

Primary analysis

ANCOVA was conducted to assess intervention effects in PASW. Linear regression models in R Console were corroborated with ANCOVA in PASW, and the results were the same. A summary of primary analysis results is given in table 5.3. Confidence intervals for ANCOVAs are supplied in appendix 5.2. Appendix 5.3

provides plots for visual diagnostics on data adherence to parametric test assumptions and transformations. The analyses were also corroborated with Mann-Whitney U tests comparing groups at posttest (given successful randomization at baseline) in order to verify results with a non-parametric test (see appendix 5.4). While ANCOVA provided more precise statistics by reducing variability, the Mann-Whitney U test results nevertheless produced similar basic findings, albeit with generally modestly smaller p values.

Table 5.3. RCT outcomes

	Baseline		Posttest		ANCOVA*	
	Mean	SD	Mean	SD	<i>F</i> , <i>p</i> -value	Effect size ^a
General self-efficacy						
Intervention	13.3	4.2	14.0	3.8	.83	.24
Control	12.7	3.9	13.2	3.6	.36	
Social skills						
Intervention	44.4	11.6	45.5	10.3	1.45	.26
Control	42.9	12.9	43.1	9.6	.23	
Friends social support						
Intervention	12.9	3.8	13.0	4.2	.05	.16
Control	12.3	3.7	12.5	3.6	.82	
Family social support						
Intervention	16.0	5.1	15.3	5.3	.01	.05
Control	15.6	4.1	15.1	4.7	.97	
SDQ - Overall Difficulties						
Intervention	22.6	5.3	23.2	5.2	1.06	.33
Control	21.7	5.1	21.9	5.1	.30	
SDQ – Conduct						
Intervention	7.3	1.8	7.6	1.9	5.21	.57
Control	6.8	2.1	6.5	2.4	.02*	
SDQ – Emotional symptoms						
Intervention	4.7	2.1	5.2	2.1	1.88	.34
Control	4.5	1.9	4.6	2.0	.17	
SDQ – Hyperactivity						
Intervention	4.2	1.2	4.0	1.3	.01	.08
Control	3.8	1.1	3.9	1.3	.94	
SDQ – Peer problems						
Intervention	6.4	1.9	6.4	1.9	1.45	.31
Control	6.7	1.9	6.9	1.7	.23	
SDQ – Prosocial*						
Intervention	7.7	2.0	7.9	1.8		.02
Control	7.9	1.8	7.9	1.8	.78	
Local adult connectedness*						
Intervention	0.9	0.8	1.10	0.8		.06
Control	1.1	0.8	1.05	0.8	.69	

Main analyses used ANCOVA except: * where data did not meet parametric assumptions and Mann-Whitney U tests were used to compare groups at posttest.

^aCohen's *d* effect size: small (.15-.40), medium (.40-.75), large (>.75).

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

Primary outcome

General self-efficacy

Univariate analysis demonstrated no significant intervention effect on *general self-efficacy* after adjusting for baseline scores, $F(1, 124) = .83, p = .36, \text{partial } \eta^2 = .01$.

Secondary outcomes

Social skills

Univariate analysis on Box-Cox transformed posttest data demonstrated no significant intervention effect on *social skills* after adjusting for baseline scores, $F(1, 124) = 1.45, p = .25, \text{partial } \eta^2 = .01$.

Social supports

Univariate analysis demonstrated no significant intervention effect on *family social support* after adjusting for baseline scores, $F(1, 124) = .01, p = .97, \text{partial } \eta^2 = .00$. Univariate analysis demonstrated no significant intervention effect on *friends social support* after adjusting for baseline scores, $F(1, 124) = .05, p = .82, \text{partial } \eta^2 = .00$. The Mann-Whitney U test demonstrated no significant intervention effect on *adult connectedness* based on differences between groups at posttest, ($Mdn. \text{ control} = 1.00, mdn. \text{ intervention} = 1.00$), $U = 1933.00, z = .396, p = .69, r = .04$ (transformed data did not adequately approximate a normal distribution to justify a parametric test).

Strengths and difficulties

Univariate analysis on Box-Cox transformed posttest data demonstrated no significant intervention effect on *overall difficulties* after adjusting for baseline scores, $F(1, 124) = 1.06, p = .30, \text{partial } \eta^2 = .01$. Univariate analysis on Box-Cox transformed posttest data demonstrated a significant positive intervention effect on *conduct* after adjusting for baseline scores, $F(1, 124) = 5.21, p = .02, \text{partial } \eta^2 = .04$. Univariate analysis on Box-Cox transformed posttest data demonstrated no significant intervention effect on *hyperactivity* after adjusting for baseline scores, $F(1, 124) = .01, p = .94, \text{partial } \eta^2 = .00$. Univariate analysis on Box-Cox transformed posttest data demonstrated no significant intervention effect on *emotional symptoms* after adjusting for baseline scores, $F(1, 124) = 1.88, p = .17, \text{partial } \eta^2 = .02$. Univariate analysis demonstrated no significant intervention effect on *peer problems* after adjusting for baseline scores, $F(1, 124) = 1.45, p = .23, \text{partial } \eta^2 = .01$. The Mann-Whitney U test demonstrated no significant intervention effect on *prosocial behavior* based on differences between groups at posttest, ($Mdn.$ control = 9.00, $mdn.$ intervention = 8.00), $U = 1953.00, z = .281, p = .78, r = .02$ (transformed data did not adequately approximate a normal distribution to justify a parametric test).

Subgroup analysis

Age

Mean change scores for eight out of eleven measures favored the younger intervention age group. When t-tests were conducted to assess differences in change scores between age groups within the intervention group ($n = 32$ in the younger age group and $n = 35$ in the older age group), five measures produced less than a .20 p

value, and SDQ overall difficulties yielded a significant difference ($p < .05$) between age groups. The results are summarized in table 5.4. When the same analysis was conducted for the control group ($n = 19$ in the younger age group and $n = 41$ in the older age group), no differences between age groups produced p values less than .20. In other words, there appears to be a systematic difference between changes in younger intervention youth (age 13 to 15) and older intervention youth (age 16 to 21) from baseline to posttest. The absence of such notable differences in the control group suggests that these differences may, in fact, be related to age interactions with intervention effects rather than simply reflecting natural ageing effects unconnected to the intervention.

Table 5.4. Comparing change scores between age groups in QS NFE

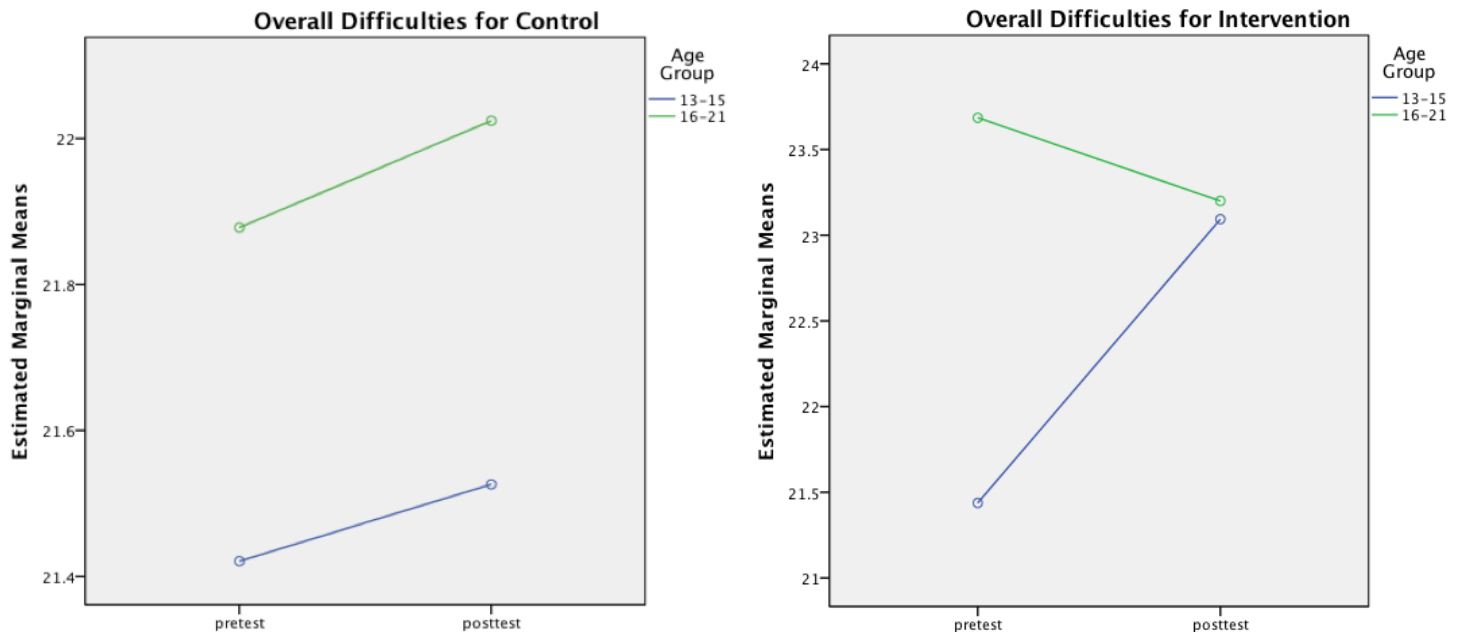
Measure	Age Group	Mean	SD	p value
General self-efficacy	13-15	1.44	4.64	.17
	16-21	.03	3.54	
Social skills	13-15	3.56	13.88	.09
	16-21	-1.29	8.30	
Conduct	13-15	.63	1.77	.13
	16-21	-.06	1.83	
Emotional symptoms*	13-15	1.09	2.28	.06
	16-21	-.09	1.85	
Overall difficulties	13-15	1.66	3.99	.04
	16-21	-.49	4.42	

Analyses conducted with T-test.

*Mann-Whitney U test used where data violates parametric test assumptions.

Plots of estimated marginal means for pre- and posttest scores separated by age group for both control and intervention are provided in figure 5.3 for overall difficulties, the measure with the most significant difference in intervention group change scores between age groups ($p < .05$). Both plots reveal that younger youths entered the trial with lower baseline scores than older youths. Both plots also reflect very little change from pretest to posttest for older youths. Notably, however, while younger control youths also show very little change, younger intervention youths show a relatively steep improvement in psychosocial difficulties.

Figure 5.3. Estimated marginal means* for pre-post age group differences



*Higher score means lesser problems.

Exploratory moderator analysis was also conducted on age group using multiple regression. No statistically significant interaction effects were detected. As Kraemer and colleagues (2006, p. 1287) posited, however, with hypothesis-

generating moderator analysis in pilot studies like this one, “effect sizes, not p values, are the focus of interest.” Effect sizes for moderator effects on three outcomes exceeded .01: SDQ emotional symptoms ($\Delta R^2 = .024, p = .08$), SDQ overall difficulties ($\Delta R^2 = .015, p = .16$), and social skills ($\Delta R^2 = .013, p = .19$). All of these would be considered small effect sizes by Cohen’s (1992) conventional standards. Yet, in a 30-year review of 636 moderation analyses from 106 articles across multiple social and psychological journals, Aguinis and colleagues (2005) found that the median observed effect size for moderators was only .002. As such, the authors suggested lower (though unspecified) standards for what constitutes a potentially meaningful moderator effect size. When age was regressed as a continuous rather than binary variable, similar results emerged. Conduct problems and hyperactivity, however, replaced emotional symptoms and overall difficulties in terms of effect sizes exceeding .01. Full moderator analyses results for age as both a continuous and categorical variable are provided in appendix 5.7.

Implementation

As a means for exploring relationships between quality of program implementation and intervention effects, this analysis compared the change scores of the intervention group ($n = 19$) from the male QS NFE center with the highest implementation score average according to the Learner Empowerment Survey results to the male center intervention group ($n = 16$) with the lowest implementation score average. Mean change scores for nine out of eleven measures favored the center with higher implementation scores. Moreover, the mean change

scores for the high implementation center's intervention group were higher than the means for the center's control group for ten out of eleven measures. Conversely, the mean change scores for the low implementation center's intervention group were lower than the means for the center's control group for eight out of eleven measures. The high implementation intervention group had no negative mean change scores whereas the low implementation intervention group had negative mean change scores for four measures (local adult connectedness, social support of friends, prosocial attitude, and hyperactivity). With little statistical power and modest differences, however, no differences achieved or approximated statistical significance.

Additionally, and more importantly, the data associated with both centers was reanalyzed as separate RCTs. Independent samples t-tests were used to compare differences in mean change scores. This approach was preferred over the ANCOVA method used in the primary RCT analysis and comparison of posttest scores since the much smaller sample sizes increased the likelihood of random baseline imbalances. As expected with a substantial reduction in power, there were no statistically significant findings. Nonetheless, a few results showed enough difference between groups to merit attention (a summary is supplied in table 5.5).

In the high implementation center, differences between the intervention and control groups in mean change scores approximated statistical significance on one measure: prosocial attitude, favoring the intervention ($p = .06$, $d = .65$). Two other outcomes produced $p < .2$ significance levels favoring the intervention: hyperactivity

($p = .17$, $d = .37$) and overall difficulties ($p = .12$, $d = .53$). As for the low implementation center, differences between the intervention and control groups in mean change scores approximated statistical significance on two measures: local adult connectedness ($p = .09$, $d = .64$) and hyperactivity ($p = .07$, $d = .67$), both of which favored the control with negative mean intervention group change scores. One outcome, emotional symptoms, produced a $p < .2$ significance level favoring the intervention group ($p = .17$, $d = .51$). Given the reduction in sample size and low sensitivity of the measure, the local adult connectedness finding for the low implementation center is intriguing.

Table 5.5. Separated outcomes for male centers with highest & lowest empowerment scores

	Baseline		Posttest		t-test [⋄]	
	Mean	SD	Mean	SD	<i>T, p-value</i>	Effect size ^a
Highest implementation center						
SDQ – Prosocial [⋄]						
Intervention	7.7	.85	8.4	1.3	1.98	.65
Control	8.3	1.8	7.8	1.9	.06*	
SDQ – Hyperactivity						
Intervention	4.2	0.9	4.5	1.2	1.09	.37
Control	4.1	1.1	3.9	1.0	.17	
SDQ - Overall Difficulties						
Intervention	22.1	4.7	24.7	5.2	1.62	.53
Control	22.7	5.1	23.2	5.0	.12	
Lowest implementation center						
Local adult connectedness [⋄]						
Intervention	1.0	0.8	.6	0.6		.64
Control	1.2	0.9	1.3	0.8	.09*	
SDQ – Hyperactivity						
Intervention	3.9	1.4	3.6	1.5	1.87	.67
Control	3.5	1.2	4.1	1.4	.07*	
SDQ – Emotional symptoms						
Intervention	4.3	2.3	5.2	1.9	1.41	.51
Control	4.7	2.0	4.4	2.1	.17	

Analyses used t-test to compare groups on change scores except: [△] where data did not meet parametric assumptions and Mann-Whitney U tests were used to compare groups on change scores

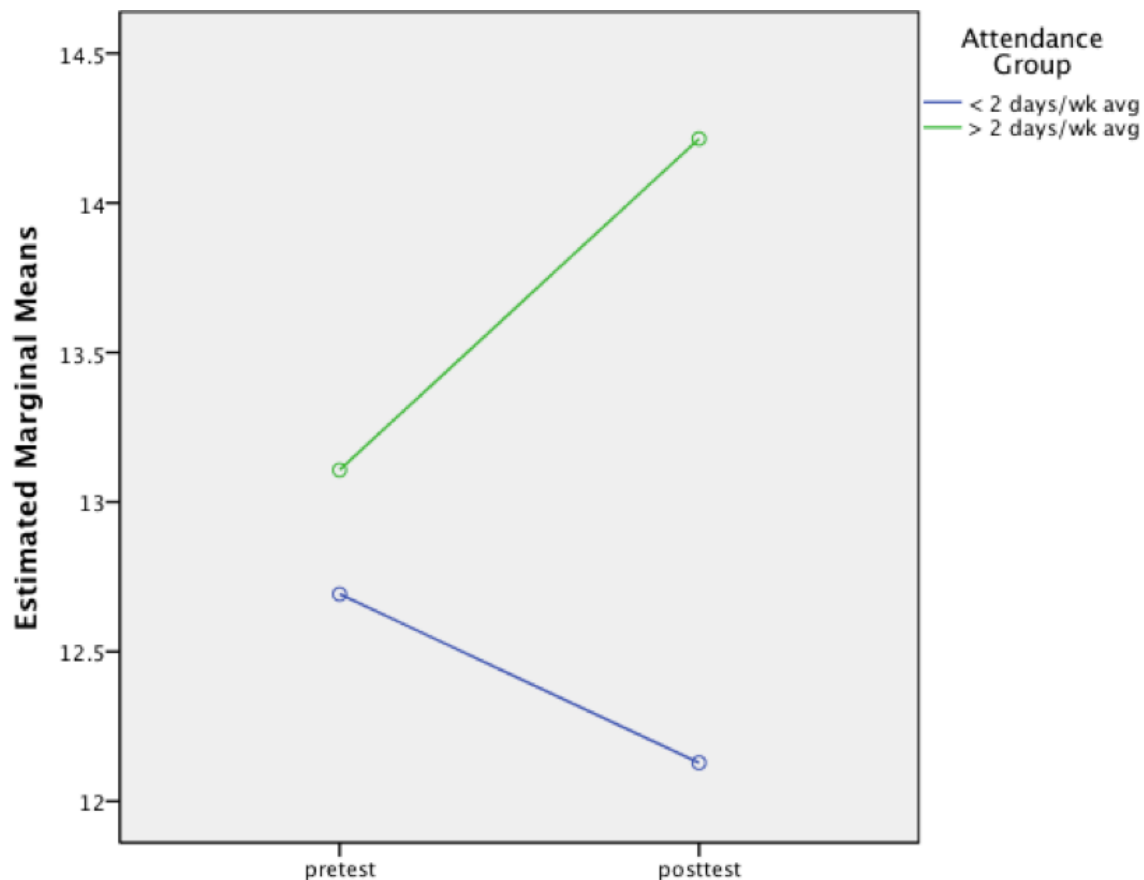
^a Cohen's *d* effect size: small (.15-.40), medium (.40-.75), large (>.75).

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

Attendance

Analysis on attendance was conducted using t-tests to assess differences between high attendance participants (weekly average of two or more sessions, $n = 28$) and low attendance participants (weekly average less than two sessions, $n = 39$) based on change scores. While mean change scores favored the high attendance group on seven out of eleven measures, statistical power was weak and p values in most cases were negligible. The one exception was social support of friends. There was a near-significant difference in change scores between intervention group participants with high attendance ($M = 1.11$, $SD = 3.01$) and those with low attendance ($M = -.56$, $SD = 4.04$); $t(65) = -1.8$, $p = .07$. The plot given in figure 5.4 illustrates the differences in changes between attendance groups.

Figure 5.4. Social supports of friends by attendance



Discussion

While closer inspection of the data gives some encouraging insights, overall trial results generally did not empirically demonstrate the effectiveness of QS NFE in changing psychosocial outcomes among out-of-school youth over a four-month period. The impact study failed to reject the null hypothesis for all but one outcome measure (conduct problems). The following discussion considers existing theory and evidence along with implementation and demographic factors in order to explore possible explanations and implications for the results, beginning with the primary outcome.

Outcomes

Self-efficacy

The study's primary outcome, general self-efficacy (GSE), was not significantly affected by the intervention. In order to explore the possible reasons for this null result, the discussion first merits closer attention to the mechanisms underlying self-efficacy. Pajares (2006a, p. 344) explained that individuals "form their self-efficacy perceptions by interpreting information from four sources: mastery experience, vicarious experience, social persuasions, and physiological reactions." According to Pajares, the most influential of these sources is generally mastery experience—the "interpreted result of one's own performance." Perceived successes build young people's beliefs in their ability to succeed further. While the first three sources are reflected in the theory of change for youth empowerment

programs (YEPs), the contribution of mastery experience is theoretically what makes empowerment such a particularly powerful instrument for constructing young people's beliefs of self-efficacy (Zimmerman, 1995).

Properly implemented, YEPs provide young people with feasible but challenging experiences in exercising skills through supported opportunities to lead and contribute. This active approach to youth development can enable young people to develop a psychological portfolio of achievements that forms the basis of self-efficacy. Through the process of empowerment, youths can gain mastery experiences not only in a particular task (e.g., performing well on a test or providing a community service), but also across a wide range of areas, such as communicating, helping, managing relationships, problem-solving, and achieving goals.

One possible explanation for a lack of robust impact on general self-efficacy by QS NFE could be insufficient opportunities for mastery experiences available to all participants in the first programming cycle. The implementation study data provides some insight to this end. One of the basic criteria for YEPs in this research is regular involvement of youths in program decision-making. Although setting a defining base standard for youth empowerment, this criterion does not represent the upward potential for dynamic engagement through youth participation.

The results of the empowerment survey indicate that 85.9% of respondents felt that their opinions mattered in program decision-making with respect to determining activities. When asked whether they felt they had the power to make a difference in how things are done in the program, however, positive responses

dropped to 69.7%. Similarly, 70.7% of respondents felt that the program gave them a chance to do something about the problems in their neighborhood and country. The approximately 15% distance between items indicates that while the program appeared to meet the base standard of involving youths in program decision-making, fewer youths could attest to the more active forms of empowerment (e.g., leadership and change agency) that are likely to comprise opportunities for mastery experience. Qualitative data reinforced this hypothesis; most youth struggled in interviews to identify concrete examples of how they or their peers had participated in activity leadership, though several testified to a more limited empowerment indicator: that their voices were heard in decision-making.

This interpretation of the data emphasizes the importance of distinctions between levels of youth participation. If verified by further research, it could have important implications for the intensity of empowerment required to change perceived self-efficacy in marginalized adolescents. Mohajer and Earnest (2009) reviewed YEPs for vulnerable populations and, like QS NFE, approached empowerment through the framework of Freire's (1972) participatory pedagogy. Of the five reoccurring themes Mohajer and Earnest found in YEP evaluations and Freire's works, two themes—skills development and community involvement—emphasize more active forms of empowerment conducive to mastery experience.

Further to the point, recent developments in youth empowerment have advanced the notion of 'youth-adult partnerships' (Anderson & Sandmann, 2009; Zeldin et al., 2008). In some respects, the QS NFE model embraces the tenets of

youth-adult partnerships, including the consideration of both parties as valued contributors in the decision-making process and a commitment to mutual learning (Anderson & Sandmann, 2009). A full characterization of youth-adult partnerships, however, is perhaps best exemplified by Jones and Perkins' (2009, p. 1160) definition: "A fostered relationship between youth and adults where both parties have equal potential in making decisions, utilizing skills, mutual learning, and promoting change through civic engagement, program planning and/or community development initiatives." The Jones and Perkins definition represents a high bar standard for youth empowerment, calling youths and adults into relationships of equal contribution through which young people do more than give input—they share responsibility with adults for program execution and leadership.

Such shared responsibility likely exposes youths to more challenging and engaging experiences through which they can build a sense of mastery encompassing a range of skills and problems, supported by attentive adult partners. Similarly, Cargo and colleagues described challenge as essential in the youth empowerment process, as it "allows people to actualize their potential through formation of 'salutary adaptive reaction patterns'...provid[ing] youth with repeated opportunities for rising to and overcoming challenges" (Cargo et al., 2003, p. 77). Cargo and colleagues suggested that this process of progressive mastery experiences leads to development of youths' sense of "enablement," which we could also interpret as a sense of self-efficacy. Further commitment to a level of empowerment that promotes mastery experiences through the process of young people's involvement could be an important step toward advancing the significance

of changes in self-efficacy. One facilitator, for example, recommended that QS NFE centers be supported with time and resources to establish rotating leadership positions among youths in the program in order to ensure that every youth has the opportunity to practice leadership skills and work more closely with adult staff to plan and execute the program. The opportunities for such mastery experiences with marginalized young people, however, depend on intentional participatory facilitation and individualized support from trained program staff and adequate time and resources for the program to give space for young people to gain such experiences. The concentration of training and activities to promote self-efficacy through mastery experiences appears to have lessened with reductions in implementation quality in the program's recent history (discussed in greater detail in chapter six).

Moreover, the failure to reject the null hypothesis for the general self-efficacy outcome could bear relevance to the academic debate between those favoring generalized or global measures of self-efficacy and those advocating domain or task specific self-efficacy measures. Because GSE is considered to be a more enduring trait-like version of self-efficacy, it may also be less sensitive to short-term change (Chen et al., 2001; Eden, 1988). Prominent self-efficacy researchers such as Bandura (1997) and Pajares (1996) have contested the applicability and reliability of general self-efficacy measures. Pajares, for example, posited that “[b]ecause judgments of self-efficacy are task and domain specific, global or inappropriately defined self-efficacy assessments weaken effects” (Pajares, 1996, p. 547). Yet, despite criticisms, more recent research has found respectable psychometric properties and broader

applications for GSE measures and disparagements based on reliability to be overstated (Chen et al., 2001; Scherbaum et al., 2006). Chen and colleagues described the merits of GSE as “enabling individuals to adapt effectively to novel and adverse environments,” particularly as “jobs become broader and more complex” demanding a transferable basis of self-efficacy (Chen et al., 2001, p. 77). Moreover, while assessing academic self-efficacy in an impact study in which all participants are participating in an academic setting may be sensible, the application of academic self-efficacy to an RCT in which the control group is not involved in any academic activities is arguably imprudent.

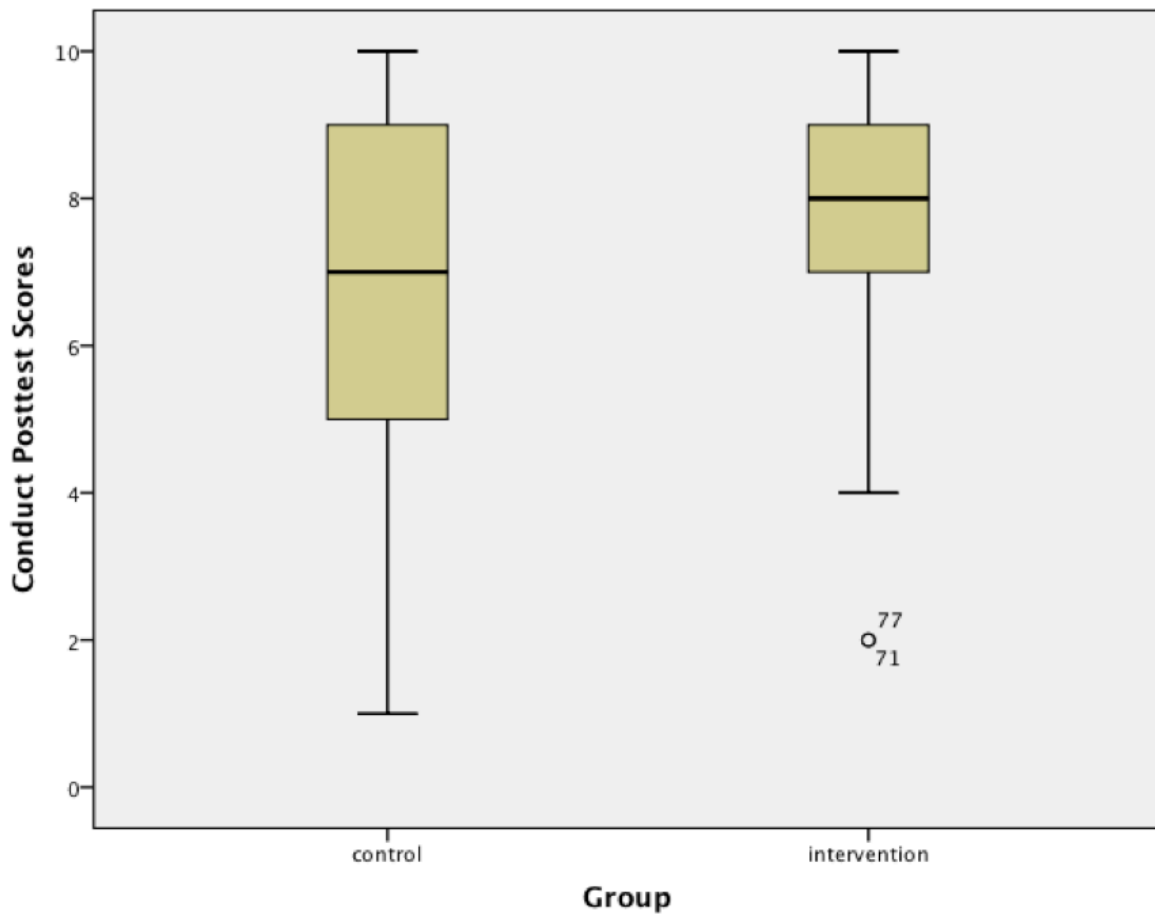
In actuality, both sides of the dispute likely call attention to important issues with important implications. Specific self-efficacy can be more informative when the primary interest lies in a particular domain of performance goals and is likely to measure effects more sensitively in such situations. Conversely, specific self-efficacy may be an inadequate or inappropriate measure for interventions such as many youth development programs with broader life preparation goals and can fail to capture the importance of young people’s ability to navigate a range of daily tasks and hurdles characteristic of most young people’s reality. See chapter seven for further discussion on the applications of self-efficacy in youth empowerment research.

Conduct

The intervention did demonstrate a significant impact of medium effect size on conduct problems ($p < .05$, $d = .57$). Conduct was a subscale measure of the SDQ.

A longitudinal study of 177 boys in the United States found conduct disorder to be a significant predictor of psychopathy into young adulthood (Burke et al., 2007). Using a Structural Equation Model, Patterson and colleagues (2000) have found antisocial behavior in youths ages 9 to 14 to significantly predict early onset of delinquency. Without intervention, Capaldi and Stoolmiller (1999) have demonstrated conduct problems to be durable in males from early adolescence to adulthood and to predict a range of adjustment problems, including problems with peer relations, substance abuse, self-esteem, relationships with parents, non-completion of education, and unemployment. Among females, conduct disorder at age fifteen was a significant predictor of antisocial personality disorder, substance dependence, illegal behavior, dependence on multiple welfare sources, and physical partner violence at age twenty-one (Bardone et al., 1996). This evidence suggests that by improving conduct in at-risk youth, QS NFE could be forming an important basis for resilience into adulthood. As illustrated by the figure 5.5 box plot of conduct posttest scores according to groups, the differences between lower quartile tails suggests that QS NFE played a particularly important role in lifting more extreme low scores.

Figure 5.5. Box plot: Differences between groups at posttest on conduct problems*



*Higher score means lesser problems.

It is interesting that conduct was significantly affected by the four-month posttest without significant intervention effects detected for other outcomes. Positive youth development programs are characterized by a common emphasis on developing young people's strengths rather than fixing perceived deficits (Catalano et al., 2004). As such, one could expect to see more immediate changes in asset areas such as self-efficacy and social skills. While the present study seems to reinforce the notion that focusing on young people's strengths is an effective way to address their

problems, it does not support the idea that measureable changes in strengths will therefore necessarily precede changes in problems.

In retrospect, the finding that QS NFE demonstrated significant positive impact on conduct even when other outcomes were not significantly affected is not necessarily surprising when considering the demographics of the sample. Various studies have demonstrated higher prevalence of conduct disorders with boys than with girls (Anderson et al., 1987; Cohen et al., 1993; Lahey et al., 2000; Offord et al., 1987). Similarly, Dwairy (2004) found higher behavior disorder scores with Arab adolescent males than females. More specifically among boys, in a longitudinal study of 10 to 20 year-olds, Cohen and colleagues (1993) determined conduct disorders to be highest with *younger* boys.

Previous research has found measures of psychosocial constructs such as general self-efficacy to be particularly precise with those populations that are lower on the construct spectrum (Scherbaum et al., 2006). As such, it is perhaps no coincidence that conduct was the singular outcome significantly affected, primarily due to changes in the younger age group, in a sample that was 85% male. Incidentally, as previous research would have predicted, the lower age group did also begin with lower baseline conduct scores than its older counterpart. Given that conduct problems are particularly prevalent in young adolescent boys, it appears that QS NFE effectively addresses this problem in its most critical developmental period and target group. Conversely, however, further research is merited to

determine the extent to which results are generalizeable to an Arab female youth population.

The precise mechanisms in QS NFE driving the intervention effect on conduct problems are largely speculative. Interviews with several boys in the program suggested that QS NFE provided a positive alternative to idleness and negative interactions outside of the center. One participant, for example, noted that the facilitators “have you enjoy sports... They fill your free time. Too much free time is not good.” Similarly, another boy noted, “I used to get in trouble. After I came here, I started staying away from trouble.” A third interviewee was more explicit about susceptibility to risky influences prior to the program:

I mean, I was just sitting at home. I would go out on the street and sit with the boys. Of course, they are useless boys. I did not used to smoke, but I learned how to smoke. Thank God, I kept away from many bad things out of my own free will. I never drank alcohol nor smoked hashish. They used to bring these things... Now that I met the boys from the center, I stay with them. There is no messing around with things like that!

It may be that an impact on conduct is one of the most immediate consequences of changing an out-of-school youth’s environment, even if that environmental change only occurs for several hours a week. The results may also indicate that the program has successfully managed to establish positive group norms early in the programming cycle (Henry, 2008), particularly norms that socially invalidate concepts captured by the SDQ’s conduct subscale items: fighting, lying, cheating, stealing, aggression, and disobedience. More specifically, inspection of change score data suggests that gains in the intervention group’s conduct shows

that gains were mostly due to improvements in youths' reports of fighting (mean change score = .16, $SD = .90$) followed by being accused of lying or cheating (mean change score = .09, $SD = .79$); mean change scores for the other three items ranged from -.03 to .06.

While the results for conduct are encouraging, they also should not be overstated. Firstly, multiple significance testing can lead to chance significant findings in secondary outcomes (Feise, 2002). Secondly, the intervention effect observed, while statistically significant, does not necessarily reflect substantial change from clinically significant risk to normal levels of conduct. Testing of the parent-version SDQ with children and adolescents in Pakistan and the self-report SDQ with youths in the United Kingdom gives the following score parameters for normal (unlikely to be clinically significant), borderline (may reflect clinically significant problems), and abnormal (significant risk of clinically significant problems) conduct scores: 8-10, 7, and 0-6, respectively⁴ (Meltzer et al., 2003; Samad et al., 2005). Notably, the control group declined as much as the intervention group improved in conduct scores. Consequently, the results reveal a partly protective effect of QS NFE, which, while important, would further limit the extent to which one might observe dramatic behavioral changes in program participants.

Furthermore, the low internal reliability alphas for all SDQ subscales, including conduct problems ($\alpha = .42$), should be taken into account. The only other identified study that tested the SDQ for Arab children revealed an even lower

⁴ Scores were reversed from the cited studies where high scores indicated high problems in order to match this study where high scores indicated low problems.

conduct problems alpha with the parent-report version ($\alpha = .32$). This low demonstration of internal consistency draws attention to a limitation in the overall outcome finding and further underscores a note of caution toward interpreting the positive result too assertively. Further development and testing specific to Arab adolescents of instruments to assess psychological and behavioral strengths and difficulties are needed.

Strengths and difficulties

The study did not reveal statistically significant intervention effects on overall difficulties nor on SDQ subscales apart from conduct problems. The intervention group appeared to experience a slight (non-significant) decline in peer problem scores, indicating very mixed experiences between QS NFE youths with peer groups during the first four months of the intervention. Possible reasons for this are considered in the next section discussing social supports of friends.

Previous research has shown hyperactivity to be a childhood predictor of conduct problems and adjustment in adolescence (Tayler et al., 1996), but less evidence appears to establish empirical links between adolescent hyperactivity and adjustment into adulthood. This construct was not identified as a central aspect of the intervention's theory of change. Nevertheless, exploratory subgroup analysis indicated that a poorly implemented QS NFE center seemed to have had negative intervention effects on hyperactivity that trended toward statistical significance.

Emotional symptoms tend to be particularly problematic for adolescent girls (Bolognini et al., 1996; Dwairy, 2004; Rosenthal et al., 1991). As such, this outcome might have been more sensitive to change in a more heavily female sample.

With respect to overall difficulties (the aggregate of the conduct problems, emotional symptoms, hyperactivity, and peer problems subscales), no significant overall intervention effects were detected. There appears to have been, however, a significant age group relationship with this measure. Early adolescents (ages 13 to 15) in QS NFE experienced significantly more improvement in overall difficulties than their older peers (ages 16 to 21).

The prosocial subscale was a measure of strengths, which was not conceptually appropriate for combining into an aggregate score with other SDQ subscales focused on difficulties (Goodman, 1997). Interestingly, when analyzing the QS NFE center with the highest empowerment survey ratings in isolation, even with sharply reduced sample size, differences in changes between the intervention and control groups were nearly significant with a medium-sized effect favoring the intervention. This finding reinforces the theoretical emphasis on developing strengths over addressing perceived difficulties in the youth empowerment and youth development literature (Perkins, 2009; Wilson et al., 2006). As expected, when an empowering program climate is better implemented, this strengths-based impact appears to materialize.

Social supports

None of the social supports or connectedness outcomes were significantly affected by the intervention. Both intervention and control groups reported slight declines in family supports from pretest to posttest. Lack of short-term intervention effects on family supports is not necessarily surprising. While the program might consider more ways to reach out to parents, this outcome may well just take a longer amount of time to change. In qualitative interviews, several youths and facilitators described a challenging relationship between the program and parents. As a female facilitator put it, “One of the obstacles we face is the parents themselves.” Several youths described family relationships as unchanging and unsupportive through the QS NFE experience. One female explained, “My family does not care. One day [my father] is here, and the next day he travels. They do not ask. I do not feel they care.”

Moreover, as considered elsewhere in this discussion, a potential ‘cultural double standard’ between an empowerment-based program and an authoritarian-based household may, at least in the short-term, create further tensions between some youths and their families (Dwairy, 2004). Conversely, some young people did describe positive family interactions during the course of their QS NFE involvement. One of the male participants, for example, posited that his parents started liking him more, and other participants reported feeling encouraged by their families. Nevertheless, these supports were either already positive going into the intervention or null or negative effects experienced by some statistically muted positive changes experienced by others.

The lack of intervention effects on social supports of friends is perhaps more concerning. The peer group is immediately connected to the program and a vital element of the theory of change. Given that behavioral norms are generally established early in the life of a group (Henry, 2008), it is all the more important that positive peer group dynamics are facilitated early in the program experience rather than hoping these changes take place later on during or after participants' exposure to the intervention. Interestingly, the intervention group change scores for the SDQ's peer problems subscale averaged exactly zero, and the mean intervention group change score for friend supports was nearly so (.13). In other words, while there was positive change in peer relations for some QS NFE participants, there was comparably as much negative change experienced by others.

This dichotomy took shape in the qualitative data as well. Some participants expressed feeling bothered or teased by their colleagues; one female participant even described peer groups in the center as 'gangs,' referring to tense dynamics between Iraqi and Jordanian youths. (Both Jordanians and displaced Iraqis participate in the non-formal education program, but this trial only included Jordanians in the study due to the lack of stability for Iraqi participants.) Meanwhile, others articulated their interactions with peers in the program as family-like (even transcending differences between Jordanian and Iraqi participants) and the most important part of their experience.

It is unsurprising that no significant intervention effects were discovered for connectedness to local adults, given low sensitivity of the single-item measure.

Nevertheless, the intervention group did fare modestly better than the control group at posttest. Notably, although the difference was statistically non-significant, connectedness to local adults constituted the only measure by which mean change scores for the older youths in the intervention group outperformed younger youths in the intervention group ($mean = .23, SD = 1.0$ versus $mean .09, SD = 1.2$). While only speculative, this could reinforce the theoretical distinction between age-related needs, with early adolescents believed to be more influenced by peer groups and older adolescents placing higher value in intimate and adult relationships (Gavin & Furman, 1989). The data may also indicate that facilitators more readily foster empowering experiences and friend-like bonds with older participants. While natural to respond to older youths in this way given generally higher levels of maturity, the program may need to work to be more intentional toward ensuring that younger participants have access to these kinds of bonding and leadership experiences as well.

Finally, it is compelling that separate data analysis of the QS NFE center with the lowest empowerment survey scores (and average attendance) revealed a negative, medium-sized intervention effect on local adult connectedness that trended toward significance ($p = .09, r = .30$). As such, we have some cause for hypothesizing that the extent to which youths feel empowered/disempowered by adults does influence the extent to which they feel connected/disconnected to them. This interpretation is consistent with previous (primarily qualitative) research, which suggests a reciprocity involved in increasing youth participation in which adults feel more connected to youths (Camino, 2000; Kirby & Bryson, 2002, p. 21-

22; Zeldin et al., 2000) and youths feel more connected to adults (Kirby & Bryson, 2002, p. 24; Mitra, 2005).

Social skills

The literature frequently ties skill building to the youth empowerment process (e.g., Cargo et al., 2003; Mohajer & Earnest, 2009; Nygreen et al., 2006). Yet, a significant intervention effect on social skills was not found in this study. Like self-efficacy, the development of social skills in young people is generally associated in the literature with meaningful leadership opportunities and challenging experiences in the empowerment process, such as enabling youths to facilitate group activities or help organize community projects (Cargo et al., 2003; Nygreen et al., 2006; Ozer et al., 2008). If the theory is valid, then a null effect could indicate that the program needs to integrate more of these kinds of supported opportunities for more participants in its empowerment model, specifically in the first four months of their involvement.

Notably, subgroup analysis revealed a difference in social skills change scores between older and younger youths that approached statistical significance ($p = .09$). As with the overall difficulties measure, figure 5.8 shows that the younger age group started with noticeably lower baseline scores than the older age group. As such, younger adolescents may have been more sensitive to intervention effects, as they had more room to change in this area. Alternatively, the program model or implementation may have been more developmentally appropriate for younger adolescents. Further meaningful leadership and challenge experiences suitable for

older participants could be necessary to promote improvement of social skills across the age range of QS NFE's target population.

Lastly, this analysis would have benefited from more specific subscales within the social skills measure. Factor analysis in previous testing of the instrument indicated three correlated factors assessing conflict resolution/avoidance, warmth and empathy, and social involvement (Spence, 1995, p. 8). In a personal communication, however, the author indicated that only the total scale has been used, and subscales according to themes, such as those mentioned above, have not been delineated in the instrument (Spence, 2010). Further development of the scale (or others measuring self-report social skills in youth) to assess categories of social skills should be a research priority.

Subgroup variables

Age

Analysis revealed a significant difference in change scores between younger and older youths in the QS NFE group on the SDQ's overall difficulties scale ($p = .04$), and most other measures, though statistically non-significant, followed this direction. The fact that younger youths in the control condition showed little change appears to rule out a simple regression to the mean explanation for changes from lower baselines. It is possible, however, that a ceiling effect could be represented. Because older youths entered QS NFE reporting fewer problems in difficulties as measured by the SDQ, there may have simply been less room for them to improve,

particularly in the span of four months. If changes occur at all with older youths, they may be more gradual.

It is also plausible that the intervention's methodology or implementation is more age-appropriate for younger than older youths. Such a hypothesis reinforces research findings by Gavin and Furman (1989) indicating that younger youths respond more sensitively to peer group interactions and belongingness. Questscope non-formal education may be particularly important for early adolescents to this end by providing a positive peer group environment, removing them from negative peer environments associated with life on the streets, and filling a void from a lack of peer group interactions altogether associated with a lifestyle of child labor. Notably, 89% of the empowerment survey respondents reported that their participation taught them how to work well with others, bolstering the role of the QS NFE experience in facilitating positive peer group interactions.

The age range of QS NFE participants, 13 to 21, encompasses a fairly wide scope that captures more than one developmental period (Arnett, 2000; Cohen et al., 1993; Marsh, 1989). Thus, increasing autonomy and leadership opportunities may be a particularly important way to increase the intervention's engagement of and impact on older youths. Buttressing the point, a two-year longitudinal study of 2,210 youths before and after transition to high school in the USA found that young people's desire to participate in classroom decision-making processes increased as they transitioned into older adolescence (Midgley & Feldlaufer, 1987).

Presently, QS NFE does not include explicitly distinct approaches or programs specific to different age groups. Although the systematic review conducted in this dissertation identified no studies that tested the effects of variations of youth empowerment models specific to different age groups, this kind of research merits further development. Interventions like QS NFE might be able to strengthen impacts by targeting approaches to youth empowerment that are sensitive to young people's needs at different stages of maturity, including avoiding over-generalized programmatic clustering of older youths with younger youths.

Gender

Although low enrollment of females prohibited the possibility of subgroup analysis based on gender, the role that gender might have played in study results cannot be overlooked. Adolescent gender differences are well documented and could have implications for intervention effects. In a study of 7,829 10th-grade students in the USA, for example, Fagan and colleagues (2007) found that girls had significantly less exposure to risk factors and more exposure to protective factors that predicted serious delinquency, and that such factors had stronger association to delinquency for boys than girls. Comparisons between Arab and non-Arab girls indicate that Arab girls are more "strictly controlled and are expected to adhere completely to the paternally dictated norms of the Arabic family" (Dwairy, 2004, p. 236; Weller et al., 1995). The potentially stark contrast of emphases on critical consciousness and individual agency in QS NFE to obedience-driven home life could render the intervention particularly influential in the lives of Arab girls.

On the other hand, the contrast could leave girls confused and conflicted. At least in the short-term, development of empowerment in Arab young women could unintentionally spawn tensions between the young woman and her family if the family implicitly or explicitly rejects such awakening of agency (assuming the intervention to be effective to this end). Such tensions conceivably could induce mental health problems such as stress and anxiety. Either way, further formative qualitative research, as well as impact evaluation capable of illustrating outcomes for Arab young women participating in YEPs, is merited.

Gender-specific attendance patterns could also have implications for results. As several facilitators observed, although it is generally more difficult to recruit girls than boys into the program, girls enrolled are conversely more likely than boys to attend frequently. This assertion materialized in the data: while female centers struggled more than male centers to meet recruitment targets during the course of the study, female participants had 40% higher average weekly attendance than their male counterparts (2.9 days versus 1.7 days). The qualitative data indicates that Jordanian females can be particularly difficult to recruit into QS NFE due to confinement to the home, strong paternal family control (female facilitators frequently mentioned parental concern of the risks that girls could interact with boys or be exposed to negative peer influences outside of the home), and the challenges for female facilitators to recruit as aggressively as male facilitators because of social pressures for female facilitators to spend hours outside of the program at home and cultural barriers for women compared to men in visiting

homes, businesses, marketplaces and neighborhoods to actively seek out prospective participants.

On the other hand, once parents or husbands are persuaded to permit females to participate in QS NFE, their attendance may exceed that of males as a result of fewer variables to compete with program attendance. Considering that Arab females tend to have fewer opportunities than males to succeed, they may also be more inclined to make the most of the program experience and education, thus increasing the commitment to attending. One 21-year-old female participant poignantly emphasized this point:

This program is my main goal; it is the most important thing in my life because there is no other way to get me to where I want except the program. How can I explain it to you? The only thing in my life that will take me to where I want to go is the program. My only hope is the program.

Furthermore, boys are more likely to work outside of the home to help earn income for the family, introducing employer and family demands on labor, and, because boys are less confined to the home, they may have more exposure to competing peer pressures that tempt youths away from programming. Notably, 58% of males in the RCT sample reported working at the time of pretest compared to 26% of females; for some male centers, this report of working climbed to as high as 83%. Given attendance disparities, if YEPs can improve recruitment of Arab females, a trial with a more heavily female sample might expect stronger outcomes as a result of higher actual program dosage.

Gender sensitivities to particular outcomes should also be further explored. Research by Eccles et al (1997), based on longitudinal data, has found feelings of

connectedness to, and emotional support from, parents and schools to explain variance in both psychological and behavioral indicators of positive development, particularly for girls. Moreover, adolescent males repeatedly report higher self-esteem and lower levels of depression—the two are typically correlated—than adolescent females (Bolognini et al., 1996; Rosenthal et al., 1991). This gender trend was repeated with Arabs in Israel; adolescent girls reported higher anxiety disorder and depression than boys (Dwairy, 2004). The SDQ's emotional symptoms subscale items relate to depression and anxiety and therefore might have been more pertinent and perhaps more sensitive to change with a sample consisting of a larger proportion of females. Considering that affective traits associated with self-esteem can be particularly low for adolescent females, future impact studies of YEPs with larger female representation in the samples should also consider the inclusion of a self-esteem measure along with self-efficacy.

Attendance

The most obvious alarm that arose from the implementation data was attendance records. Only 48% of participants met the minimum expected attendance (two days per week average) between pretest and posttest. Though the evidence is limited, unsurprisingly, youth development programs appear to have an important relationship between dosage and outcomes. Evaluation (without a prospective control) of the San Francisco Beacon Centers, for instance, observed that youths who participated in the program for three or more sessions per week were 33% less likely to report a drop in self-efficacy by the 18-month follow-up

period than peers who participated in fewer or no sessions (Walker & Arbreton, 2001). Similarly, mean change scores suggested similar trends for most outcomes in this study's data set, though subgroup analyses lacked statistical power.

The clearest difference between attendance-level groups was found in the social supports of friends measure, which trended toward statistical significance. Participants who attended at least two days per week on average reported increased supports among friends whereas data from lower attending youths reflected a slight decrease in perceptions of friendship supports. Subgroup analysis in this study did not permit determination of directions of variable relationships. In other words, it is unclear as to whether lower attendance contributed to lower outcomes (a dose-response effect) or whether some of the lower outcomes better provide insights into factors that contributed to lower attendance. Plausibly, to some extent both relationships are true.

If the intervention has positive impacts on those who attend regularly, participants who do not attend simply do not have exposure to the program components believed to affect change in psychosocial outcomes. Facilitators spoke of the essentiality of cultivating trust and confidence in relationships between adults and youths as well as between youths and their peers as the bedrock for the empowerment process. This assertion echoes themes in youth empowerment literature (Camino, 2000; Cargo et al., 2003). Low attendance could have precluded the amount of contact conceivably required to adequately develop these aspects of

the participants' relationship to youths and adults at the center and therefore stall or thwart positive effects anticipated by the theory of change.

Implementation quality

Chapter six discusses quality of implementation of QS NFE and implications for youth empowerment in greater detail. For the time being, this section draws attention to the possible roles of adherence (in terms of participants' rating of the intervention's adherence to its empowerment-based theory of change) and program quality (presence of a supportive, participatory program environment) in explaining outcomes.

The subgroup analysis highlighted apparent differential outcomes between the male QS NFE center with the highest rated implementation quality according to empowerment survey results and the male center with the lowest rated implementation quality. Some results, even with small sample size, trended surprisingly close to statistical significance. Namely, the QS NFE center with higher empowerment scores produced an approximately significant intervention effect on the SDQ's prosocial subscale ($p = .06$, $d = .65$) while the center with lower empowerment scores appeared to affect feelings of local adult connectedness ($p = .09$, $d = .64$) and hyperactivity ($p = .07$, $d = .67$) negatively at a trend significance level.

Though underpowered subgroup analysis only enables hypothesis-development, we have cause to surmise that, when implemented well, particularly according to the main theoretical components of the empowerment process—

supportive adult interactions that promote competence, meaningful activities, and the ability to choose and have an impact in the program through participation (Fymier et al., 1996)—YEPs like QS NFE can have positive impacts on psychosocial outcomes of highly marginalized young people. If no such patterns could have been identified, it might have been more sensible to suggest that implementation was irrelevant and the theory of change failed—i.e., that the intervention model simply did not work. The trends, however, are encouraging (though still speculative) for QS NFE's effectiveness when the empowerment process is well implemented. Further rigorous impact studies of YEPs with comparable target populations and adequate sample size to more precisely differentiate effects according to implementation quality are needed, but the near-significant intervention effect on prosocial behavior and less statistically robust differences between control and intervention groups on other outcomes are promising indicators.

Different factors could explain the apparent link between lower implementation and null or negative intervention effects. Centers with inadequate implementation could unintentionally place young people in destructive peer environments. Dishion and colleagues (1999) have illustrated how well-intended youth programs have in some cases *contributed* to delinquent behavior by aggregating deviant young people together and in so doing reinforcing antisocial behavior through negative peer interactions. Behavioral norms are generally established early in the lives of groups (Yalom, 1995, p. 111). Adult staff play an essential role in facilitating prosocial norms early on and sustaining them through the class or program (Henry, 2008). Lackadaisical implementation could forfeit

critical opportunities to foster positive group norms in a YEP leaving room for unhelpful group norms to root themselves instead. On the other hand, strong implementation of youth empowerment could reinforce positive group values, ownership, collective efficacy, and a sense of safety (Wong et al., 2010).

Notably, the conclusions that can be drawn on implementation's relationship to outcomes are limited by the lack of external observer ratings based on standardized instruments of intervention implementation. Future research could benefit from such measurement in tandem with impact study results, and YEPs like QS NFE should, as much as feasible in the context of an empowerment framework, develop clear manuals and curricula for programs—in partnership with young people—so that more accurate assessment of implementation could be made possible.

Study strengths

This research had a number of strengths. The overall dissertation and the conduct of this trial in particular are based on the MRC's framework for developing and evaluating complex interventions (MRC, 2008). The framework plays an important role in advancing the relationship between evidence-based practice and community-based, social intervention. When possible and appropriate, the MRC framework encourages the use of RCTs for less biased estimates of intervention effects. The present study appears to have been the first successfully completed RCT of a youth empowerment or non-formal education program in the Middle East. The undertaking has demonstrated that rigorous impact evaluation is achievable in an

Arab youth context. The effort could help raise the profile of and confidence in rigorous impact study designs for social and educational interventions in the region, as well as enable others interested in impact evaluation to learn from the practical challenges and limitations of the present study so as to improve on it.

The inclusion of multiple intervention sites in the trial was an advantage for this study. Multisite RCTs can have more external validity than a single-site study since any one program location could have unique programming or contextual factors that could influence results (Weinberger et al., 2001). The inclusion of several intervention sites also enabled comparisons of implementation and exploratory analysis into potential relationships between implementation characteristics and intervention effects. This is particularly important in studying YEPs, an intervention category in which the process of a young person's involvement is believed to be so critical to his or her development (Blanchard et al., 1996).

The integration of an RCT with mixed-methods research augments the value of the project. The importance of joining RCTs with complementary process and impact study components is becoming increasingly recognized (Mayo-Wilson, 2007; Oakley et al., 2006). This recognition transcends extraneous debates over the superiority of quantitative or qualitative methods when, in fact, the merits of particular methods are better judged according to particular research questions and contexts (Sackett & Wennberg, 1997). The application of other methods can help explain RCT findings and illustrate a fuller picture of the intervention so as to

increase the usefulness of the research to stakeholders. The next chapter focuses on the process and implementation components.

Self-administered questionnaires have traditionally been literacy-dependent (Weinhardt et al., 1998). The use of audio-visual, netbook-based survey administration was an important application of technology to overcome challenges with tracking self-report outcomes in largely illiterate youths in a developing context. Additionally, although attendance (compliance) was generally low, at 6.3% attrition, the study nevertheless managed a high follow-up rate with posttest completion.

Finally, substantial stakeholder involvement had important advantages, especially with the assembly of a practitioner-academic research coordinating team and the contribution of program staff in the execution of the research project. This approach helped to increase local buy-in of rigorous impact study, promote better cultural relevance and appropriateness of research methods and instruments, and support the development of stakeholders' skills and understanding of evaluation through the process of their training and involvement.

Study limitations

While an important step forward, this research is nevertheless a relatively small and pilot experimental study. The research has limitations that should be considered in interpreting the results and advancing further impact evaluation of youth programs in the region. The length of time between pretest and posttest could constitute one study constraint. Some youth development programs have required

up to several years to reveal significant intervention effects on at-risk youth (e.g., an oft cited evaluation of the Quantum Opportunities Program, Hahn et al., 1994). In a review on positive youth development in the United States conducted by Catalano and colleagues (2004), 80% of the twenty-five programs found to be effective lasted for at least nine months.

However, in conducting the systematic review discussed in chapter two, this author identified at least five experimental or quasi-experimental evaluations of youth development programs that detected statistically significant impacts on psychosocial outcomes in four months or less (all of which were completed after the Catalano and colleagues review: Grolnick et al., 2007; Lakin & Mahoney, 2006; Olson-Merichko, 2005; Shelton, 2009b; Winkleby et al., 2004). This later evidence, along with stakeholder feedback described in chapter four, suggests that intervention effects in four months are plausible. It is nevertheless possible that a longer follow-up period could have revealed different outcomes. Planned follow-up without a control group at twenty months could shed light on this possibility.

Sample size may have limited the study's ability to detect effects. This RCT fits into a growing body of small experimental trials conducted to assess the effectiveness of youth development programs (HFRP, 2006). Although a sample size calculation was conducted, in a field such as youth empowerment with a meager previous research base from which to draw data for estimations, assumptions used render such calculations particularly tenuous. Moreover, with the suboptimal attendance rates, a larger sample may have been required to detect intervention

effects. Though the sample size may have been sufficient for assessing overall intervention effects, a larger sample would be needed for more robust analysis according to subgroup variables. Additionally, the numbers of subjects across intervention sites were unequal, which means that intervention or contextual characteristics of some sites were overrepresented in the sample and others underrepresented (Kraemer & Robinson, 2005).

This impact study only used self-report measures for outcomes. Some studies have observed susceptibility of self-report measures to social desirability response bias (Arnold & Feldman, 1981), response-shift bias (Howard & Dailey, 1979), and recall bias (Weinhardt et al., 1998). These issues were somewhat addressed by using well-reputed instruments that have been psychometrically tested and corroborated with parent, teacher, and/or professional ratings in previous studies (Weinhardt et al., 1998). While the data offers limited evidence of possible response-shift bias given some extreme scores at baseline, the use of ANCOVA based on posttest scores in the primary analysis rather than basing primary analysis on change scores minimized the impact of such potential data contamination. Furthermore, precautions such as the use of a computer-based self-administered survey along with careful efforts to ensure that participants understood the full confidentiality of their responses were taken to improve sense of privacy and candidness. Nevertheless, while inappropriate or impractical for the current study, future trials might strengthen impact studies of YEPs by supplementing self-report measures with reports by other raters or institutional data (e.g., education or criminal justice records).

Some of the measures had unsatisfactory reliabilities with the observed sample. Respected instruments for measuring the outcomes of interest in this study have largely been developed and tested outside of the Arab world. Church (2010) highlighted a number of measurement biases that can arise in the application of foreign measures to different cultural contexts. Negatively phrased, reverse-scored items, for example, have, in some cases, been shown to contribute to lower scale reliability and validity, especially in non-Western cultures (Church, 2010; Smith, 2004). Incidentally, the scales that produced the lowest internal consistency scores in this study—SDQ subscales—also included several reverse-scored items.

Future research should develop and test measures that are culturally appropriate for Arab adolescents. Further testing of the self-report SDQ as well as development of more sensitive instruments with good psychometric properties to measure connectedness to adults are particularly needed in Arab contexts. Notably, during the course of this dissertation, another research team has developed, and begun reliability and validity testing of, a new 21-item instrument to measure mental health of Arab youth (Makhoul et al., 2010). Early psychometric results have been promising (Chronbach's $\alpha = .90$). The high alpha, however, could be partly attributable to simply testing reliability on a large number of items, and the usefulness of the instrument might be limited for intervention research in that it lacks subscales to assess more precise aspects of mental health. Mental health in the abstract may not be a priority for YEPs' theory of change. Nonetheless, the measure could support future research in the region, and the intensive process used to create

and validate the instrument could be informative for developing other culturally valid measures with Arab youth.

Furthermore, the investigator reviewed the literature, and the youth development field in general appears to lack good, validated measures that could assess changes in the adult connectedness construct with more sensitivity. This echoes a similar call for more “sensitive, developmentally appropriate measures for determining...community connectedness” made by the Berg and colleagues (2009) youth empowerment trial included in this dissertation’s systematic review.

Generalizability

Randomized controlled trials inevitably carry strengths and limitations for external validity, or ‘generalizability’ (Rodrik, 2009). Generalizability concerns the extent to which study results apply to other circumstances (Altman et al., 2001). As the revised CONSORT statement put forward, generalizability is “a matter of judgment and depends on the characteristics of the participants included in the trial, the trial setting, the treatment regimens tested, and the outcomes assessed” (Altman et al., 2001). To this end, reporting information on eligibility criteria, setting, and location; the interventions and how they were administered; the definition of outcomes; and the recruitment and follow-up periods are essential to improve readers’ ability to make informed judgments about the extent to which findings apply to other circumstances of interest (Altman et al., 2001). These details were reported in chapter four, except for how the intervention was administered, which was assessed by the implementation study reported in chapter six.

Expanding on the CONSORT directions for reporting external validity, Bonell and colleagues (2006) recommended that trials include assessment of the intervention's coverage, feasibility, acceptability, and for whom and what the intervention is effective. *Acceptability* is examined in-depth in the context of chapter six's implementation study, particularly with respect to participant responsiveness to the intervention. The most salient issue of generalizability is the extent to which results apply to other populations. In part, this is a 'coverage' question—i.e., who was not included in the study, and to what extent could one expect that results would generalize to those populations? It is also a matter of 'for whom and what'—i.e., of those included in the study sample, did the intervention work differently for subgroups of the sample population with particular characteristics? Addressing this question helps to avoid ecological fallacies—inappropriate generalizations of aggregate study results to individuals for whom findings might have been different than they were for the overall sample.

With regards to *coverage*, the study's sample comprised Jordanians and was largely male, Muslim, urban, uneducated, and low-income. Populations characterized by different demographics might well have responded to the QS NFE model differently. At the macro-level, the study lacked the coverage to ensure applicability of results, for example, to other countries, particularly those with sharply different contextual differences. Youth in rural areas within Jordan in which the program is implemented could also respond differently. At the micro/meso-level, the trial's coverage could also limit generalizability to youths in the same community and even in the same program. For example, QS NFE has a substantive

Iraqi refugee constituency due to grant funding responding to the crisis of a large number of Iraqi youth in Jordan displaced by war. Iraqis come to the program often under starkly different circumstances from Jordanian youths—e.g., with more education, less child labor experience, and higher levels of war-related traumatic exposure. A sample that included Iraqi youths may have produced different results.

Additionally, more heavily female samples could also respond differently to the intervention. Qualitative feedback from youths and facilitators, for instance, suggests that once females are in the program (recruitment is typically a challenge due to family preoccupations about girls taking part in extracurricular activities), they may be more committed to attendance since they often prize limited time with peers out of the house and do not have as many variables, such as paid labor or neighborhood peer groups, competing for their time. Intervention trials in other cultural contexts have observed gender differences in outcomes (Arnett, 1999; Leventhal & Brooks-Gunn, 2004).

Further to the micro-level, the ‘for whom and for what’ aspect of generalizability was examined in exploratory subgroup analysis according to age, but, as a pilot study, this analysis was limited by sample size, and additional potential moderator variables should be included in larger trials in the future. Exploratory subgroup analysis findings reported in this chapter suggest the hypothesis that the program is more effective in changing outcomes for younger youth (ages 13 to 15) than older youths (ages 16 to 21); this hypothesis should be further tested to determine whether the finding is repeated with greater power and

generalizable under other youth empowerment programming circumstances in the Arab world.

In terms of *feasibility*, the trial had some advantages for generalizability, as it was conducted on an existing program implemented according to natural conditions and within its natural setting (West et al., 1996). As such, research-related funding or researcher control did not artificially uphold the intervention studied. On the other hand, the intervention was enabled through a unique non-governmental organization (NGO) and Ministry partnership that significantly reduced operational costs for the NGO. This factor, along with limited manualization of the intervention methodology by the implementing NGO, could constrain the feasibility of transporting the intervention more broadly, particularly under different NGOs or in other countries with different political contexts.

There are no hard and set rules about the extent to which findings from a particular, stand-alone trial transfer to another setting of interest. A researcher's task is to provide the best information possible to help readers make informed decisions about interpreting findings for their needs. As the systematic review reported in chapter two makes clear, however, further experimental studies, complemented by process and implementation studies, are lacking and needed across a broader range of demographic and cultural contexts to inform the applications of YEPs. The ability to appraise this trial's generalizability improves if it can be analyzed along with more trials of comparable interventions in other circumstances.

Summary and conclusions

This research involves the first completed randomized controlled trial of a youth empowerment program (YEP) in the Middle East and one of the few globally. The undertaking demonstrates that rigorous impact study designs can be integrated into youth development efforts in the region in order to build a stronger evidence base for what works and provide practical data that can direct stakeholders toward key areas for improvement.

Study results demonstrated a significant, medium-sized intervention effect on reducing and protecting against conduct problems ($p = .02$, $d = .57$) over the course of four months. Previous research suggests that improved conduct can serve as an important proximal indicator for resilience in other aspects of adjustment into adolescence and early adulthood. Significant intervention effects, however, were not detected by four-month posttest for any of the other observed outcomes, including the primary outcome, general self-efficacy.

A lack of significant effects on other outcomes was likely connected to gaps in program implementation and generally low attendance levels, but limitations on statistical power render these conclusions largely speculative. Notably, unlike QS NFE, the YEPs evaluated in two other studies included in this dissertation's systematic review had explicit youth mobilization and social action components. Emphasis on similarly challenging and active leadership opportunities in QS NFE *might* help to strengthen youth engagement in the program and impacts on

psychosocial wellbeing, but this kind of youth empowerment still lacks evidence in the region.

Closer inspection of the data revealed some encouraging signs for QS NFE. The intervention appeared to have a stronger impact on younger adolescents (ages 13 to 15) for most outcomes; differences between age groups were especially pronounced for overall difficulties, emotional symptoms, and social skills. Young people with higher program attendance experienced more positive changes in social supports of friends. The QS NFE male center with high youth ratings of empowerment outperformed the center with the lowest empowerment ratings on the majority of measures and appeared to have been especially more effective in strengthening participants' prosocial behavior. Conversely, however, the center that struggled the most with implementation and attendance produced null intervention effects and even modest indications of harm.

The findings should provide some tepid encouragement for the positive role of YEPs such as QS NFE in the lives of marginalized Arab adolescents. On the other hand, results should also serve to warn stakeholders of the possibility of unproductive use of resources when implementation quality is inadequately realized to ensure that youths have sufficiently supported, meaningful experiences to develop their skills, competencies, and sense of efficacy. This pilot impact study provides useful insights for program development and future research, but the evidence for YEPs remains largely inconclusive.

Focuses on quality of implementation, strategies to promote attendance, and further evaluation of YEPs in the Middle East and North Africa (MENA) should be priorities for funders and practitioners. The next chapter describes the results of a mixed-methods approach to assessing the state of QS NFE implementation and exploring factors that promote and prevent the diffusion of empowerment processes across program sites.

Chapter 6: Implementation and Process Study

Introduction

This chapter describes the methods and results of a process and implementation study of Questscope non-formal education (QS NFE) coinciding with the randomized controlled trial (RCT) reported in chapters four and five. The transmission of findings from RCTs to informing theory and practice requires information contextualizing why and how results come about. This context helps researchers better understand RCT findings and more accurately interpret their implications (Bonell et al., 2006; Mayo-Wilson, 2007; Oakley et al., 2006). Eccles and Appleton-Gootman asserted that, at minimum, a comprehensive program evaluation should not only address the question, "was the program effective?" but also, "how well was the program implemented?" and, "if effective, why?" (Eccles & Appleton-Gootman, 2001, p. 8). The importance of conducting and reporting implementation research is amplified with multisite RCTs in which implementation can vary within the same study (Weinberger et al., 2001). The integration of a process and implementation study component is consistent with the MRC Framework for Complex Interventions, which asserts that implementation assessment can help explain why interventions fail or succeed and how they can be optimized (Craig et al., 2008).

Two main questions directed this phase of the research. These were:

1. *Implementation* - How well was the program implemented?

2. *Process* - What are the process factors that influence the extent to which youth empowerment is implemented and sustained in local programs?

Methods

Implementation study

The first and primary question addressed implementation. It is predominantly descriptive (or 'confirmatory') in nature (Palinkas et al., 2011). The program's theory of change articulates hypothesized essential elements: youth participation in decision-making, supportive relationships with trained adult facilitators, asset-building activities, and a pro-social environment. Guided by key areas of implementation fidelity (Mihalic, 2009), this implementation study aimed to clarify: (a) the extent to which the program adhered to these components in operation, (b) the quality of their delivery, (c) quantity of program exposure participants had, (d) the participants' responsiveness to the program, and (e) the extent to which program components were differentiated between each other and other programs.

Mixed methods were employed to investigate program implementation. The use of both qualitative and quantitative methods enabled a degree of triangulation in order to contribute to more valid methods and provide a check on qualitative research conclusions (Mays & Pope, 1995).

Quantitative measures

The Learner Empowerment Measure, a 30-item instrument devised by Schultz and Shulman (1993), adapted by Fymier and colleagues (1996) for college

students, and adapted again by Lakin and Mahoney (2006) for adolescents, was administered to each of the six QS NFE centers to gauge participant perceptions of empowerment in the program environment (Fymier et al., 1996). Chapter four (pp. 238-240) gives a description of the measure and its properties. The measure went through the same translation, adaptation, and piloting process described for outcome measures in chapter four (pp. 226-227), except that the measure was not piloted or administered using the computer-based system. Assisted by youth and staff feedback, wording was adjusted where needed to increase clarity of item meanings (linguistic equivalence) in language the youths understood.

Four items were added to the end of the instrument to assess aspects of empowerment specific to the QS NFE model. These supplementary questions asked whether participants felt their opinion mattered in determining program learning topics, rules and expectations, and activities, and one question asked whether facilitators said positive and affirming things to them. The English language version of the administered empowerment measure is included in appendix 4.4. A research staff person administered the survey verbally to QS NFE groups with participants following on an anonymous paper version. Each program center had one survey administration, which occurred on a day of the week on which facilitators anticipated relatively high attendance.

Attendance was monitored through standard institutional data collected and used by the Ministry of Education and Questscope for the program. This system involved facilitators tracking daily attendance of participants and inserting the data

into a computer-based system, which Questscope or research staff could monitor or extract data from using a central online network.

Qualitative methods

The implementation study also involved a series of semi-structured interviews with an initial sample of QS NFE youths ($n = 12$) and facilitators ($n = 12$) (12 stakeholder interviews were added in the process study for a total n of 36). Participants were identified from the six QS NFE centers included in the RCT: two youths and two facilitators from each. Purposive sampling was used. The research staff coordinated with the centers to arrange for private interviews with two of the four facilitators in each center team according to willingness and availability. For the same visit, facilitators were asked to arrange private interviews with two willing and available youths included in the intervention group of the RCT who had most recently completed the four-month posttest. The sampling method was intended to gather a range of youths representing the sample included in the impact study rather than selecting participants based on a particular characteristic.

Interviews were conducted in Arabic by a trained research team consisting of a lead Arabic interviewer, a simultaneous translator, and the principal investigator. Signed informed consent forms were collected for all interview participants; the form is included in appendix 6.1. A set of starting questions was predetermined by the research team for interviewees to explore implementation areas such as the extent to which youths were involved in program decision-making, the nature of youth-facilitator relationships, peer dynamics, facilitator approaches, quality and

frequency of training and support of facilitators, quality and frequency of engaging activities, characteristics of the general programming environment, and barriers to recruitment and retention.

In total, the qualitative research included 36 interviewees representing approximately 22 hours of recorded interview time. The mean adult interview time was approximately 57.5 minutes, and the mean youth interview time was approximately 18.5 minutes. All interviews followed a semi-structured interview format assisted by question guides specific to each group and designed with input from Oxford University and Questscope research team members. The question guides for all interviewee groups are included in appendix 6.2. All but three interviews were conducted in Arabic. Coding was conducted from translated English transcripts. Qualitative data analysis was assisted by the NVivo 8™ software package.

Considering methodology in the context of Mihalic's (2009) implementation framework, each dimension could be at least partially evaluated through the aforementioned methods. *Adherence* to the program components was assessed by qualitative interviews with youths and facilitators primarily in terms of understanding program adherence to implementing participatory-methodology (to what extent youths felt they had opportunities to contribute to program decision-making) and facilitator training intended to support the methodology. Adherence to

the participatory-methodology was further assessed with the quantitative empowerment survey, especially with its impact domain items.

Quality of delivery was assessed through qualitative investigation of the youths' perceptions of the activities and facilitators' efforts, as well as through interviews with facilitators exploring their own expressions of commitment to the role and how they saw their role. *Exposure* was assessed using attendance records, and issues affecting exposure were explored through qualitative interviews.

Participant responsiveness was assessed through qualitative interviews with youth participants concerning the extent to which they felt engaged in the program and their perceptions concerning the value of the program to them; responsiveness was also assessed quantitatively by the empowerment survey's meaningfulness domain items. Finally, *program differentiation* was explored qualitatively by assessing whether the main program components in QS NFE's theory of change were clearly identifiable from youth and facilitator feedback.

Process study

The qualitative research based on the same aforementioned interviews explored the secondary research question for this chapter regarding the *process* factors that promote or obstruct implementation of the empowerment process at the program level. To this end, a modified *grounded theory* approach was applied insofar as the study generated new interviewees, questions, and codes from data accumulated through the qualitative research process (Bryant & Charmaz, 2007). Grounded theory has useful applications in this kind of study due to its explanatory

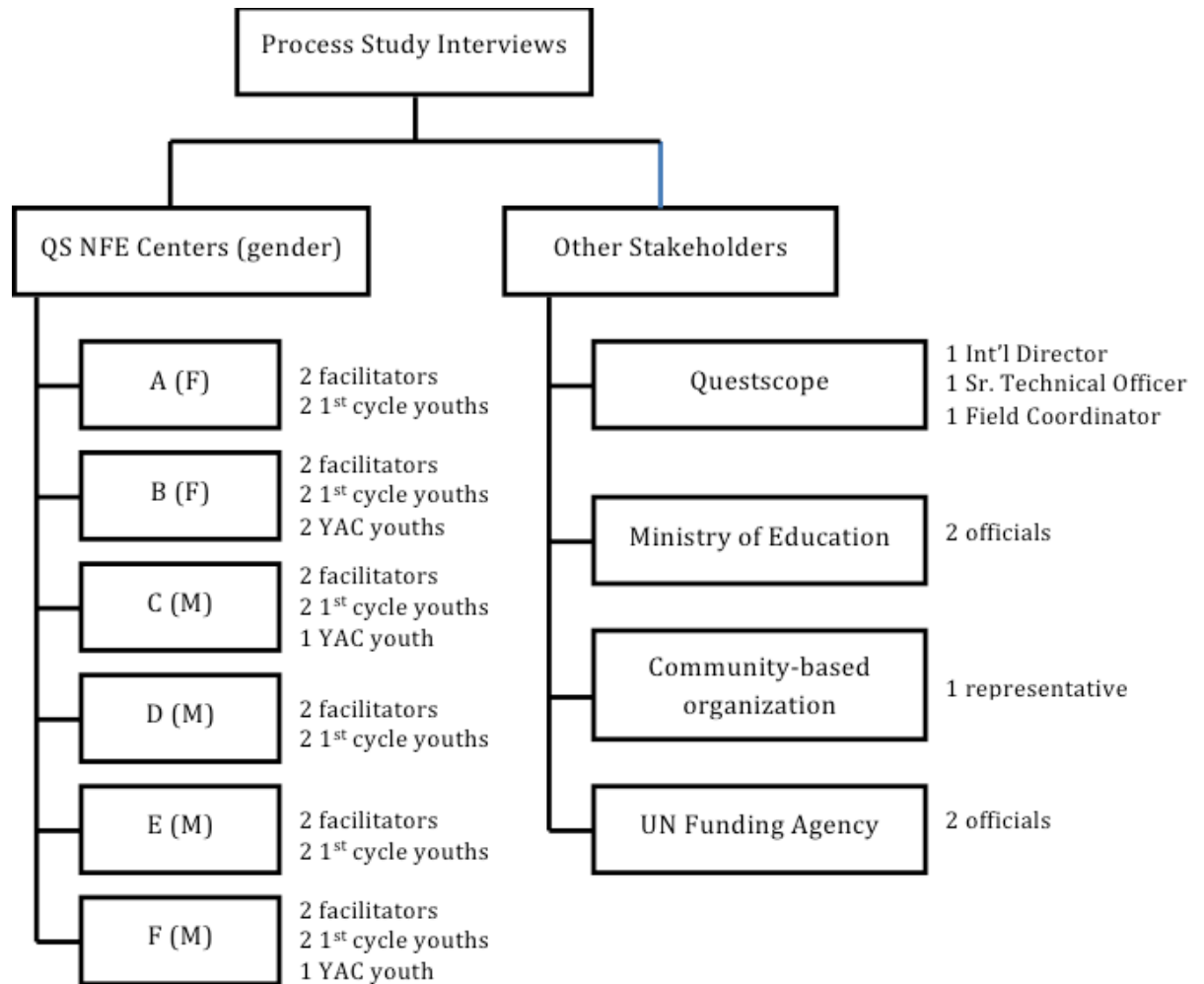
power. It contrasts with deductive hypothesis-testing approaches that begin with a theory and try to test its validity (Mills et al., 2006).

Twelve interviewees were added to the research (included in the $n = 36$ mentioned above), which included a staff director at one community-based organization that implements Questscope's informal education, two Ministry of Education (MOE) officials designated to QS NFE, four third-cycle or graduated QS NFE youths serving on Questscope's Youth Advisory Committee (YAC), three Questscope staff members centrally involved in the design and implementation of the QS NFE program, and two staff representatives at a United Nations agency office in Jordan that has provided one of the main sources of funding to the program. These additional stakeholder perspectives aimed to provide a more rounded understanding of experiences with empowerment (or lack thereof) in the program and factors that influenced the empowerment process. A diagram of interviewees is given in figure 6.1.

Grounded theory can include both *positivist* and *social constructionist* influences (Mills et al., 2006). The present methodology has some postpositivist elements in the sense that some mechanisms were employed to systematize the process and *reduce* (not eliminate) bias, such as semi-structured interview guides, systematic data coding processes, and a beginning list of relevant interviewees to help produce more systematic lessons about organizational processes. Positivist (or 'postpositivist') orientation is particularly influential in assessing implementation fidelity, which inherently assumes that there are discernable standards to which

program practice is expected to adhere, and the qualitative research is designed to help assess that adherence.

Figure 6.1. Interviews diagram



The process study was oriented toward a social constructionist viewpoint for qualitative inquiry to the extent that it worked to understand behaviors and ideas related to central concepts for implementation in their social context. As Patton explained, qualitative research from the social construction paradigm investigates people's perceptions, explanations, beliefs, and world-views and the consequences

of these constructions on their behaviors (Patton, 2002, p. 96). This approach contrasts, for example, with a strictly *positivist* approach (stronger in its assertions of scientific certainty than postpositivism) to studying organizational phenomena in which the researcher believes there are objective truths to be discovered apart from social constructions of knowledge and believes himself or herself as objectively capable of uncovering them (Patton, 2002, pp. 92-93; Willig, 2001). Social constructionism also differentiates from *relativism*, which “rejects concepts of ‘truth’ or ‘knowledge’ altogether” (Willig, 2001, p. 3).

As Schwandt (2000, p. 198) explained, social constructionism assumes that knowledge is in some sense ideological, political, and permeated with values. Ungar (2001), for example, conducted qualitative case studies on the construction of ‘resilience’ among youths in out-of-home placement. The data were used to explain how care providers had a negative or positive impact on at-risk youths’ identities depending on whether they contributed to the construction of ‘problem-saturated’ or ‘health-enhancing’ identities. Likewise, the nature of youth-adult relationships and youth involvement in YEPs could be better understood with careful attention to how these phenomena are socially influenced.

This orientation to qualitative research can contribute to better understanding of why particular aspects of the program methodology were, or were not, implemented or why they were implemented in particular ways. This social constructionist approach to qualitative research was aided by the researcher’s prolonged interactions and multiple unstructured observations with program

centers and staff over the course of a year, which helped the researcher to better understand qualitative interview data in its social context.

Implementation study results

Implementation fidelity refers to the degree to which an intervention is delivered as intended (Breitenstein et al., 2010). Although the literature on conducting and reporting implementation research is less developed than that on outcomes research, papers in recent years have outlined dimensions of implementation fidelity in order to aid investigators (e.g., Montgomery et al., 2007; Schoenwald et al., 2011). The implementation framework described by Mihalic (2009)—adapted from Dane and Schneider (1998)—was devised specifically for studying youth interventions. As such, the categories of Mihalic’s implementation fidelity framework—adherence, exposure, quality of delivery, participant responsiveness, and program differentiation—were used to guide the reporting of this implementation study. Table 6.1 presents a summary of implementation study methods and findings according to Mihalic’s framework.

Table 6.1. Summary of QS NFE implementation fidelity			
Category	Methods	Results	Illustrative quotes (+) positive; (-) negative
<i>Adherence</i>	- Quantitative questionnaire with impact domain - Qualitative interviews with facilitators and youth participants	- Only partial (1/3) of facilitator training implemented for new staff since 2006 - 80.3% overall positive survey responses on <i>empowerment</i>; variations between centers (61.9%-87.2% range for center means) - Indications of youth participation at all centers but extent of empowerment varied between centers	(+) “We implement any suggestions or idea they have... We record everything; everyone must participate.” – facilitator (-) “The youth participation is adequate to a certain level, but we as facilitators usually direct them most of the time; it’s not necessary to go with all of their suggestions.” – facilitator (-) “The training is getting less. I do not know why. It is not the same as before.” – facilitator
<i>Exposure</i>	- Institutional attendance records - Qualitative interviews with facilitators and youth participants	- 48% overall participants met expected attendance; variations between centers - More trips, engaging activities, outreach to families and employers, and transportation support could increase attendance - Frequency and duration of program implementation met expectations	(-) “Really, it is easy to get a student, but <i>maintaining</i> them in the program is what’s hard. There is a lot of family pressure – for example, the mother telling the son he has to work.” – facilitator (+) “Yes, the activities and the entertainment; this makes talking to the boys easier, and they will be motivated.” – youth
<i>Quality of delivery</i>	- Qualitative interviews with facilitators and youth participants	- Generally supportive youth-adult relationships with some exceptions (some youth expressing neglectful or demeaning interactions) - Generally positive facilitator attitudes toward youth participation with some exceptions (some facilitators expressing controlling attitudes) - Feedback suggests high quality of facilitator training (but inadequate frequency) - High commitment among facilitators to mission	(+) “The facilitators are very good with us and they do not spare any effort to help us.” – youth (-) “Our relationship with facilitators is okay; it’s neither good nor bad... Sometimes they do not cooperate.” – youth (+) “We did not join NFE because of financial gains. I used to work in exams, and I used to make good money. My community needs me, and I need my community. I honestly need my community to become the way I want it to be, because I live in it.” – facilitator
<i>Participant responsiveness</i>	- Quantitative questionnaire with meaningfulness domain - Qualitative interviews with facilitators and youth participants	- 84.2% overall positive survey responses on <i>meaningfulness</i>; variations between centers (66.6%-88.6% range for center means) - Generally positive youth expressions of engagement with some exceptions	(+) “When I first came, I was not comfortable because I did not know everyone, but now I am comfortable. I participate in the lesson; I was not like that before!” – youth (-) “Maybe because we are too many; they do not allow you to participate or give your opinion or participate in an activity.” – youth
<i>Program differentiation</i>	- Qualitative interviews with facilitators and youth participants	- Components (youth participation, asset-building activities, trained adult support, and prosocial environment) differentiated, though less so in some centers with lower adherence - Generally differentiated positively from formal education experiences with some exceptions	(+) “I am comfortable in the center; it is not like the school. The teachers in the school do not understand like the teachers here.” – youth (-) “It’s all the same; it’s all normal; it’s all studying.” – youth

Adherence

Adherence refers to whether the intervention is “delivered as it was designed or written”—for example, with all core components being delivered to the appropriate population, staff trained appropriately, using intended techniques and materials, and in the locations or contexts prescribed (Mihalic, 2009). Unlike some psychosocial interventions, YEPs often are not highly prescribed. A participatory methodology is inherently flexible so as to foster youth engagement through dynamic processes of involvement and shared decision-making. Nevertheless, youth empowerment does depend on well-trained staff or volunteers who can facilitate participatory processes (O'Donoghue et al., 2002), and, as stipulated by the theory of change, certain process components and types of activities characterize YEPs and must be made as explicit as possible and assessed in the context of implementation research.

Facilitator preparation

All teachers selected as ‘facilitators’ to implement the Questscope non-formal education (QS NFE) program are expected to participate in a special training curriculum designed and implemented by Questscope. The full training involves approximately 85 hours of sessions and exercises divided among three levels. The training curriculum is not gender-specific, but trainers try to make content relevant for participants’ particular settings and needs. The first level of training is generally conducted in an outside environment with facilitators from multiple centers (e.g., in a hotel or conference center), and subsequent levels of training are typically

conducted in the program context either at the center or in the local community.

The levels include the following content:

Level 1: Introduction to participatory methodology (5 days, approximately 30 training hours) - The first level orients facilitators to participatory and dialogue-based learning, adolescent development, and skills to prepare for program sessions.

Level 2: Participatory rapid appraisal and recruitment (6 days, approximately 30 training hours) - The second level of the training puts special emphasis on 'participatory rapid appraisal' (Chambers, 1994), a community engagement technique to help facilitators connect to the community, understand context in which out-of-school youths are living and working, and build local channels for recruiting out-of-school youths into the program.

Level 3: Advanced participatory facilitation skills (5 days, approximately 25 training hours) - The third level advances further into topics of empowerment-based educational and activity tools and social change and emphasizes skill-building for group facilitation and participation.

The qualitative research appeared to reveal a low level of implementation fidelity with respect to facilitator training in the period during and leading up to the present study. All interviewed facilitators did report participation in the first of the three levels of training, which takes place at an off-site location. The *full* training curriculum, however, had not been implemented since 2006. As such, only facilitators who worked with the program in 2006 or earlier were exposed to all

three levels of training. Additionally, facilitators and Questscope staff generally indicated that little to no training update ("refreshment") was taking place to help facilitators who had been with the program for longer to maintain or advance their skills in the program methodology.

Eight out of twelve facilitators did report receiving a three-day (approximately 25-hour) training update at an off-site location provided by Questscope in 2008, which focused on program issues unique to the growing Iraqi constituency, discussing general implementation problems, and involved some revisiting of the participatory methodology skills trained in level one. This, however, was generally perceived as inadequate compensation for the training desired and did not include advanced levels of empowerment facilitation that are features of the training curriculum for levels two and three. Seven interviewees described the amount of training as inadequate, and no facilitators spoke favorably of the *amount* of training provided. One facilitator, for example, recounted, "The training is getting less. I don't know why; it's not the same as before." Calls like the following for more regular training and refreshment were common among facilitator interviews: "I would like Questscope to train me more in non-traditional methods for teaching, especially in focusing on dialogue in the teaching process. Questscope should give periodic training."

Although the amount of training exposure was suboptimal, most facilitators did report favorable perceptions of the quality of training that they did receive. As one facilitator explained of the training, "It is positive. You can tell that a lot of effort

was put into it. We benefited from it with our work; we could not have communicated with the girls without the training. [We learned] participation—learning through participation.” Several facilitators ($n = 9$) even indicated that they incorporated the methodologies they learned for QS NFE into their teaching in formal education. In particular, interviewees positively highlighted the competence and enthusiasm of the trainers as well as how the methodology of the training departed from traditional educational methodologies. The advanced trainings (levels two and three) received during or prior to 2006 and the partial training (level one and update training) provided since may have been adequate to sustain a certain level of implementation adherence to the participatory methodology, but the overall low adherence to the full training design marks a potentially important constraint.

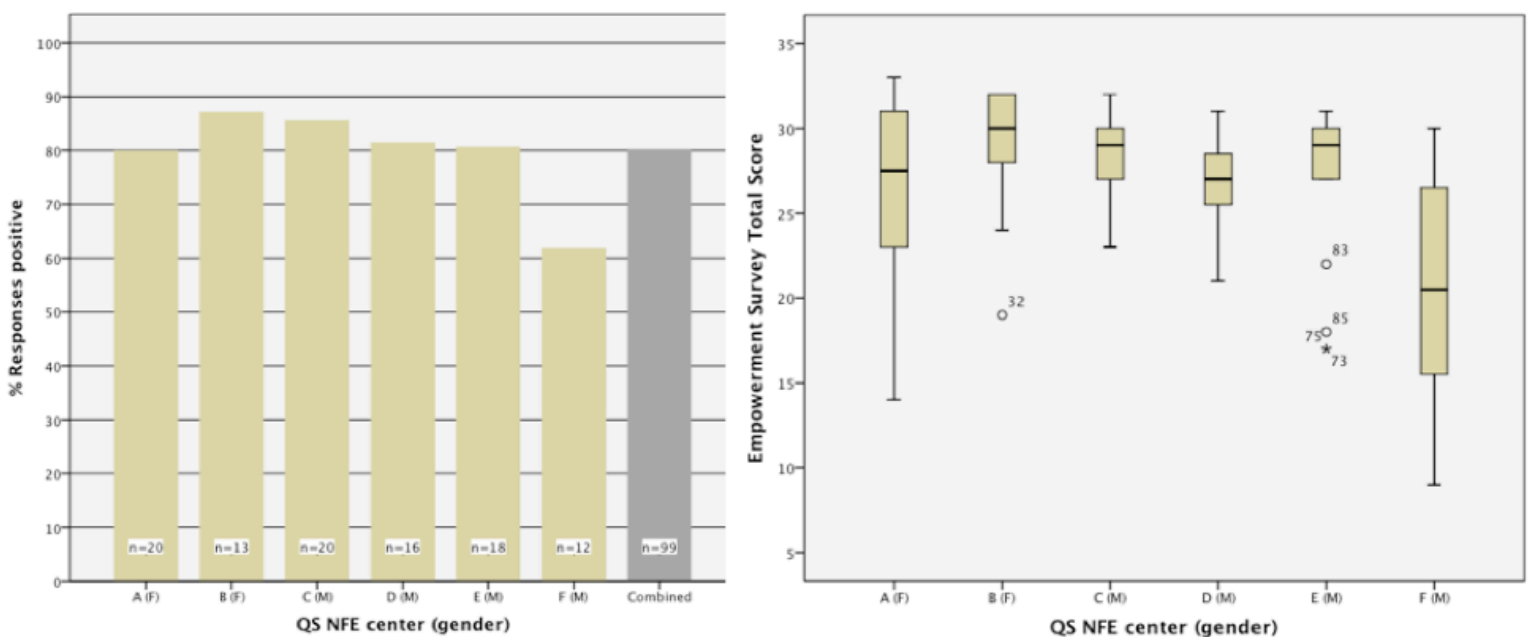
Youth involvement in decision-making

The second area of adherence addressed in this chapter is the application of the intended program methodology. Questscope non-formal education is based on a methodology in which adult staff facilitates learning and social processes involving young people regularly in program decision-making. Youth empowerment was dually assessed by the Learner Empowerment Survey and qualitative interviews.

For the Learner Empowerment Survey, the study obtained 100 survey responses of center attendees. One survey response was omitted from analysis due to missing data for the majority (over 50%) of items. On average, respondents ($n = 99$) answered 80.3% positively of the program’s implementation according to the

empowerment survey. The survey results indicated generally high feelings of overall empowerment in the program context. There were notable differences, however, between centers (see figure 6.2). A regression analysis indicated that center was significantly correlated with empowerment survey results ($t(98) = -2.89, p = .005$), suggesting heterogeneous implementation across sites.

Figure 6.2. Overall empowerment survey results (bar graph, left, and box plot, right)

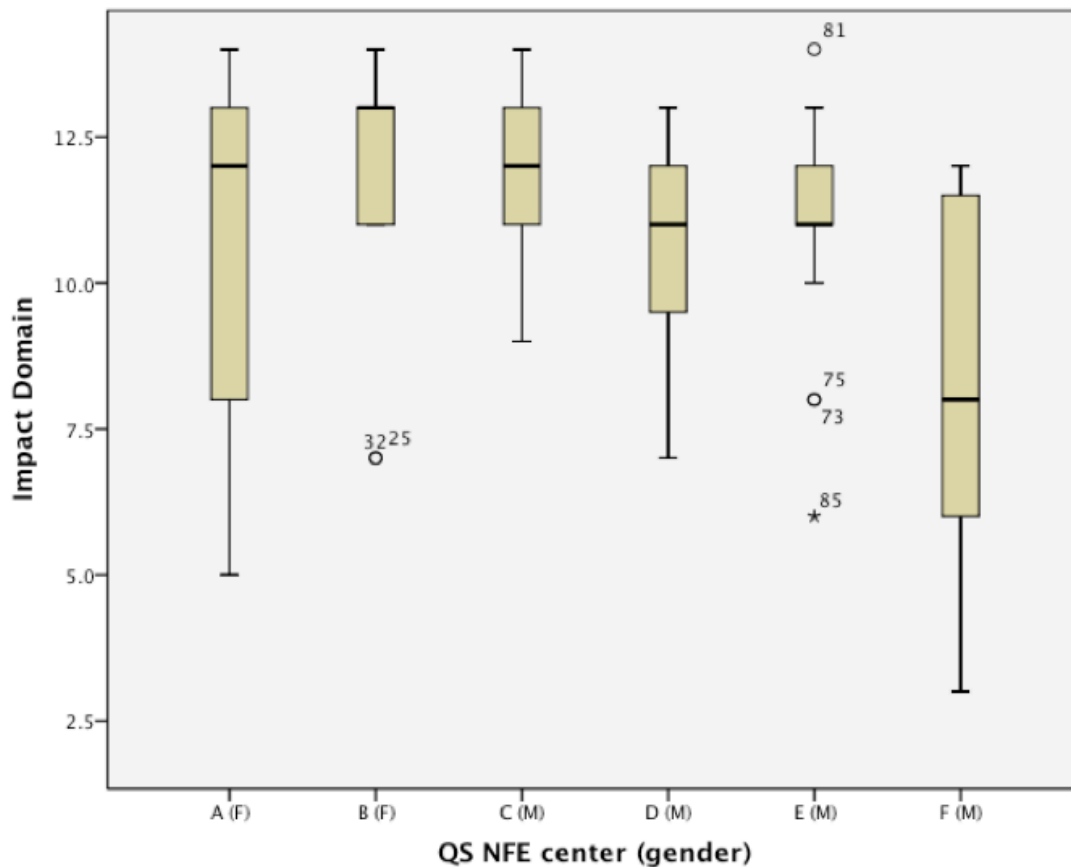


Average percentages of positive responses according to domains included the following: 84.2% for *meaningfulness*, 77.9% for *impact*, and 79.8% for *competence*. The items included in the *impact* domain are specific to young people's feelings of ability to make choices in the program and influence the way the program is run. It includes items such as "I have the power to make a difference in how things are done in the group" and "My opinion matters in determining

activities.” The box plot provided in figure 6.3 again shows heterogeneity between centers on youths’ ratings according to the *impact* domain and relatively low scores in one male center (F), and a wide range of responses in one of the female centers (A). Regression analysis confirmed a significant relationship between center and impact domain scores ($t(98) = -2.414, p = .02$). Qualitative interviews largely corroborated these results but also provided further understanding of the extent to which adherence to empowerment-based methodology was realized.

All twelve of the facilitator interviewees reported using participatory methodologies in the implementation of the program. In most cases, this involved soliciting young people’s opinions and ideas on day-to-day learning topics and activities and implementing programming based on that feedback. One facilitator explained, “Implementing the program in this way, that is part of the methodology—for them to choose subjects and discuss them.” Another facilitator described the integration of leadership opportunities into the participatory process: “Yes, because when we decide on a trip or a camp, where to go, we consult them and they make the plans. We just convey their opinion, where they like to go. Today there is a lesson; we ask them ‘Who would you like to be the leader?’ Sometimes we forget we are the facilitators, and we give them the role.” Furthering the leadership theme, another facilitator explained: “In the (formal education) classroom, the children were marginalized. In my group, I give them leading roles, because this class of society is marginalized. Our role is to build the self-confidence.”

Figure 6.3. Impact domain scores



While several facilitators described the participatory process in more active terms such as leadership roles and deciding or choosing activities, some appeared to view the participatory process more passively. This usually involved welcoming input from the youths, which the facilitators would consider, but the decision about whether to integrate the idea and the action in implementing it was essentially recognized as the adult's role. For example, when asked about the youths' ability to participate in decision-making, one facilitator stated, "Yes, I listen to them. If they

ask for something logical, I help them.” Another facilitator suggested that the youth could give input because “they might have an idea that we’ve missed; they might have an idea we haven’t thought of.” Neither of these interviewees described more substantive ways they involved youth in decision-making with further follow-up questioning.

These approaches better align with what Hart’s ladder described as the “consulted and informed” level of participation – i.e., “The project is designed and run by adults but children are consulted” (Hart, 1992). This orientation falls below the standard of youth empowerment outlined by the review criteria described in chapter two and represented by the top three levels of participation in Hart’s ladder. Seeking youths’ input, while potentially important, by itself does not foster the engagement, sense of ownership, and mastery experiences that characterize the youth empowerment process when youths are directly involved in program decision-making and leadership. This creates an adherence gap in the program’s theory of change, and it reinforces heterogeneous experiences with the *impact* domain reported by youth through the empowerment survey.

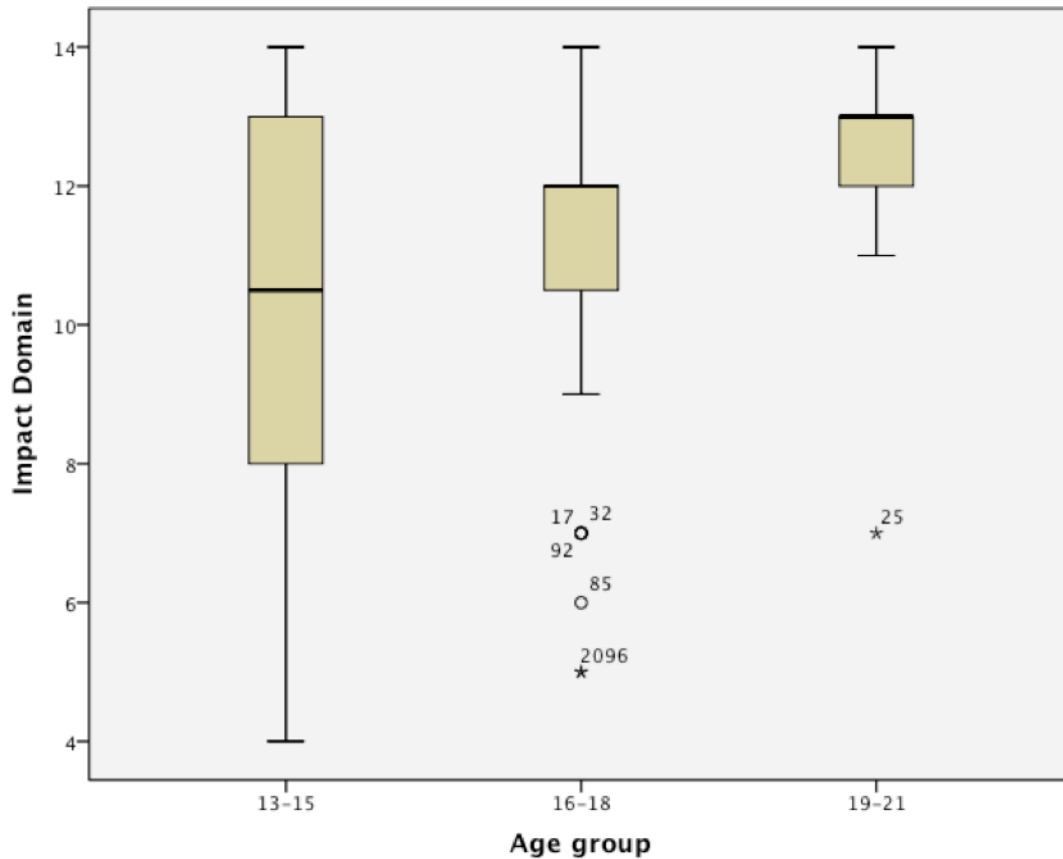
Youth interviews also suggested mixed experiences of empowerment. Highlighting positive methodology adherence, one of the youth, for instance, explained that participants were regularly involved in determining activities and that all the youth in the program contributed to the establishment of group rules and expectations. This matches the intervention’s stated intention to foster participation through its program decision-making and ‘educational contract’

(group rules and expectations). Another youth described how staff facilitated sessions in which youths provided ideas for programming, the ideas were recorded, and then the group made activity decisions. The same interviewee gave an example: “Something simple: we did not used to play tennis here, okay? I asked them if we could play tennis. Immediately, they gave us a tennis course and we started playing tennis. Something simple, immediately.” Mechanisms such as recording youths’ ideas and showing immediate follow-through reinforce adherence to the participatory component of the program model. These techniques assure youths that their views are recognized and integrated in how the program is run.

Not all youth participants, however, reported feelings of involvement in program decision-making. When asked whether he felt like he had influence on how the program was run, one youth responded, “Not always. Sometimes. When the facilitator asks me, for example, he suggests something to all the students—it depends on what he suggests. The facilitator does what he is convinced is right.” Another youth’s feedback indicated a stronger departure from adherence to empowerment methodologies: “Sometimes, I want to express an opinion or participate, but [the facilitators] do not allow me... Maybe because we are too many. They do not allow you to participate or give your opinion, or participate in an activity.” When asked if he felt the facilitators listened to him and his peers, another youth stated, “No, they all ignore us.” It is possible that such comments merely reflect adolescent angst, but such perceptions could also highlight instances of legitimate implementation gaps; notably, such comments did tend to come from youths in centers with lower empowerment survey ratings. Of the eighteen youth

interviews, six interviewees' comments indicated problems in adherence to the empowerment-based program methodology.

Figure 6.4. Impact scores by age group



As such, while there did appear to be general adherence to the participatory methodology intended by the program design, youth and facilitator feedback indicated that there was not universal adherence. Gaps could dilute intervention effects predicted by youth empowerment's theory of change. Notably, regression analysis indicated that, while the participants' program cycle did not significantly predict impact scores ($t(90) = .329, p = .74$), age did ($t(90) = .2.928, p = .004$). As

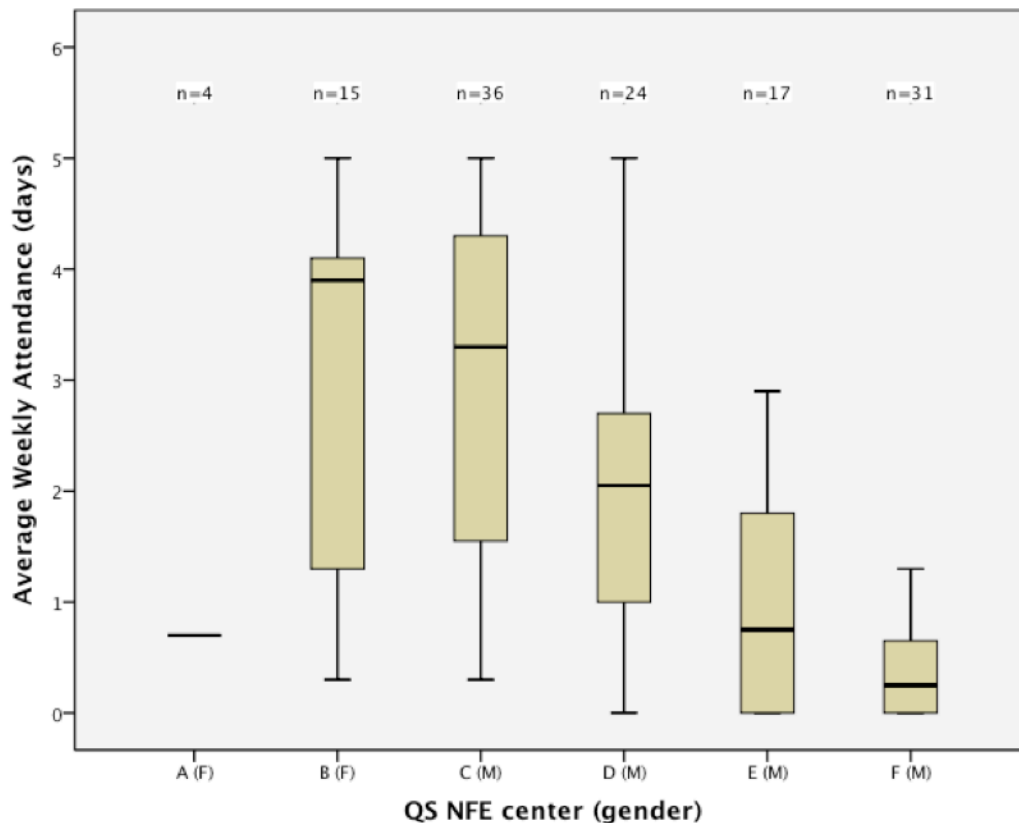
shown in the box plot in figure 6.4, older youths were more likely to report higher impact scores than younger youths, who reported more widely varying experiences with participation. This may have been due to facilitators' propensity to foster more leadership roles and participation among older youths. Some facilitators indicated this propensity in interviews. As such, facilitator training may need to focus on skills for facilitating developmentally appropriate opportunities for younger adolescents to participate as well.

Exposure

Exposure, or "dosage," describes the amount of program delivered and received in relation to the amount prescribed by the program model and can include the number of sessions or contacts, attendance, and/or the frequency and duration of sessions (Dane & Schneider, 1998). Exposure was primarily documented in this study according to participant attendance, as has been done in other youth intervention fidelity studies (e.g., Sanchez et al., 2007). Attendance may be measured by four indicators: (1) whether participants spend *any* time in a program, (2) the intensity of their attendance, (3) the duration of their attendance, and (4) the breadth of their attendance (Weiss, 2004). This study assessed attendance by intensity—days per week of program participation. Additionally, research staff kept note of disruptions in the program administration. Brief public holidays or weather circumstances disrupted usual program operations for one to two weeks in the included program centers. These were accounted for in analysis of participant attendance by omitting weeks without programming from calculations of weekly

attendance averages. Otherwise, the programs appeared to operate as expected in terms of frequency and duration of sessions.

Figure 6.5. Attendance by program sites



Exposure, as measured by attendance, revealed one of the most significant implementation gaps. The study showed overall attendance problems, with 52.3% of the sample attending less than the two-day per week average expected by the program. The box plot in figure 6.5 shows that median attendance averages for half of the six centers met or exceeded the two-days-per-week expectation. The majority of participants in the other three centers did not meet this exposure expectation, and, in two of these centers, no participants met the expected attendance level.

Notably, the center with the lowest attendance was also the center with the lowest participant responsiveness according to the empowerment survey.

Qualitative data analysis showed that the youths' single most prominent program recommendation was to increase the quantity and quality of entertaining and engaging activities. Ten youths alluded to increased engaging activities as a way to improve the program and participation. As one young person explained, "The first thing that attracts [youths] is the entertainment; when you tell them that they will have fun, they'll come... Most of the girls like to come because it is more entertaining than educational." More specifically, youths were particularly enthusiastic about increasing daytrips in the program—e.g., to local museums, historical sites, or parks. Many target youths lacked the resources and opportunities to independently explore beyond their immediate work places or communities. As such, trips were considered a major program attraction and opportunity for increasing attendance and recruitment. Several facilitators reinforced this point. When asked about how to increase attendance, a facilitator stated, "Yes, a trip—anything, however small. It does not have to be Aqaba (a beach destination in southern Jordan). If we take them to the park every two months they would feel it is great!" Other participants focused on increasing activities within the center, including more experiences with arts, music, and other forms of play and recreation.

Several youths and facilitators also described transportation as a barrier for attendance. At the time of this study, Iraqis in the program received 2 JD (\$2.82 USD) per day for transportation from funding arrangements specific to that

population, but Jordanians did not receive any transportation support. One youth explained the challenge this posed for him and his peers: “We all live far away, and we take various means of transportation that we have to pay for ourselves.” Some youths and facilitators presented the issue not only as a technical challenge for helping youths to participate regularly in the center, but also as a toll on morale given the discrepancy in benefits offered between two groups in the same intervention. Frequent comments regarding this implementation challenge were captured by one of the facilitators:

What is happening now is that the Iraqi students are getting rewarded. They get 2 JDs while the Jordanian students do not get anything. This issue is demoralizing for the students. I wish that at least they would be given 1 JD, or even half a JD, for transportation. Why are they being treated differently? Financially, I mean. Most of them dropped out of school because of their financial-economic situation... They continue, but they do not come regularly—even if they really want to come.

Transportation has been identified in other youth empowerment literature as a common structural barrier preventing young people from participating in youth-friendly environments that needs to be acknowledged and removed (Scheve et al., 2006).

Notably, some QS NFE centers—generally those with higher implementation quality in the empowerment methodology—maintained relatively high attendance despite lack of transportation support for participants. As such, it is possible that if program quality is high enough, many young people will sacrifice a great deal in time and personal resources to participate, but lack of transportation support nevertheless poses a clear attendance challenge for youths coming to the centers

from distant homes or workplaces and left to draw on very limited personal financial means to do so.

Quality of delivery

Quality of program delivery refers to the “manner in which a teacher, volunteer, or staff member delivers a program (e.g., skill in using the techniques or methods prescribed by the program, enthusiasm, preparedness, attitude)” (Mihalic, 2009). Quality was principally considered in the present study in terms of facilitators’ attitudes toward adolescents and toward program methodologies and the quality of youth-adult relationships as reflected by qualitative data.

Though this was not universally true, interviews with youths and facilitators indicated generally high quality of youth-adult relationships. In fact, seven out of twelve facilitators and eight out of sixteen young people described the youth-adult relationship as a friendship or family-like bond. One facilitator explained, “I have a rule: always go down to the level of the students... They are my brothers.” These feelings were reciprocated by many youths—e.g., “I feel that the facilitators are my friends.” Open and informal communication and positive feedback appeared to constitute particularly important elements of high-quality youth-adult relationships. In many cases, open and informal communication was expressed through playfulness and humor between youths and facilitators. For other youths, this meant that the facilitator could be trusted with discussions about personal problems and offered advice and support to deal with them. Positive feedback was described by

youths in terms of encouragements and recognition for little efforts or accomplishments.

Several youths differentiated between the quality of relationships they had with facilitators in the QS NFE program and those they experienced with teachers in the formal education system. One youth explained, “I am comfortable in the lesson. The teachers in school do not understand like the facilitators here.” When another youth was asked if he felt the facilitators were controlling, he responded, “No, no. I left school because of this... I consider him almost my older brother. The most important thing is he is my friend. If we have a problem, we tell it to him.” The strength of these relationships implied a positive degree of quality in terms of many facilitators’ commitment and approach toward young people.

Nevertheless, just as other aspects of implementation have indicated heterogeneous fidelity, so too did quality of delivery. A minority group of interviewees suggested poor quality of delivery in terms of facilitators’ willingness or competence to maintain positive and empowering relationships with all youth participants. One young person described how facilitators in the center would embarrass participants with lower reading levels and threaten to send them to the first educational cycle if they were in the second. More often, youths did not have specific complaints about the dynamics between youths and adults, but even with follow-up questions they were unable to give examples or specific indications of a positive relationship; they would simply and phlegmatically describe the relationship as *ady*, meaning 'normal.'

Notably, analysis presented another more subtle variation in quality of delivery with respect to the attitudes that facilitators brought to the delivery of the program. For example, this section included descriptions of youth-adult relationships as friend or family-like as a positive indicator of quality of implementation delivery, but nuances in the ways adults expressed such attachment could be insightful. Most facilitators described a positive youth-adult relationship in parental terms (e.g., “These students could get lost, so I consider them as my own children.”), whereas youths were more likely to describe a positive youth-adult relationship as a friendship or a rapport between siblings (e.g., “I do not feel she is a facilitator; I feel she is my friend.”). While both perspectives imply a strong level of concern and bonding in the relationship, the subtle distinction may have implications for empowerment. The sibling or friendship perspective may imply a more peer relationship, whereas the parent-child view of the relationship may reinforce disempowering attitudes toward youths that maintain a distance of authority between youths and adults.

Some facilitators expressed attitudes toward adolescents that more clearly departed from youth empowerment values. Notably, four out of five of the interviewees that expressed orientations explicitly countering the values of the methodology worked in the two program centers with the lowest empowerment survey scores. Two facilitators described youth-adult interactions in terms of obedience. One of whom expressed frustration with participants: “The girls are stubborn and do not obey us.” The other facilitator explained that when they told participants to be quiet, they obeyed them. In both cases, facilitators appeared to

measure the quality of the relationship in terms of obedience—the young people’s compliance with adult authority and rules.

The other three facilitators recognized young people’s contributions in the process but minimized their importance in shaping decisions affecting the individual youths or the program. This indicates attitudes encumbering quality of delivery. One facilitator, for example, described youths’ contributions or ideas as expressions of selfishness that could be entertained enough to appease their ego. Others simply acknowledged little commitment to involving youths in program decision-making or realizing ideas proposed by the participants. One of the facilitators, for example, explained how she tried to talk one of the young people out of her desire to learn to dance, while the other contended that he would direct the youths most of the time and did not see it as “necessary to go with all of their suggestions.” These statements did not coexist with others by the same interviewees representing more empowering attitudes, suggesting that the comments accurately reflected their basic orientations toward youths in the program at that time.

Although in the minority among interviewees, these kinds of attitudes and approaches toward youth participants revealed gaps in quality of delivery. Dane and Schneider (1998) described the attitudes and enthusiasm among program leaders regarding the program and methodology as central features of quality of delivery in assessing implementation fidelity. The facilitators cannot effectively deliver the program’s methodology if their fundamental orientations toward young people do not reflect enthusiasm for or even accordance with the core features of the

methodology. As such, youths at some program centers would not have been exposed to the same level of implementation quality as youths in other centers, which could have diluted any intervention effects that might have existed otherwise.

Finally, this section considers the motivations behind facilitators' efforts in the QS NFE program as an aspect of quality of delivery. Any statement of motivation as to why facilitators participated in the program was coded. Two coding categories for facilitator motivation emerged through data analysis: service and financial. By far, facilitators predominately described their reasons for participating in the program as a desire to serve. Twenty-six references from eleven interviews articulated a desire to serve as justification for program participation compared to one reference that described financial reasons for participation. As one facilitator insisted, "I became a facilitator because I like voluntary work... Even if Questscope tells us they are going to stop giving us a salary, I am going to continue with the project." Another explained, "I wanted to fix this problem of youth dropping out of school. I wanted to be part of the solution and put them on the right track... It is a social and religious responsibility." Facilitators often cited religious convictions and desire to have a positive impact in the lives of vulnerable youths.

The facilitator that cited financial motivation also recognized service-based impetus for program participation: "Do not forget that, apart from being for the good of these students, there is also a financial gain that helps me as a teacher. I mean, you know the income of teachers is low, so it helps us manage. This is the truth with all honesty." Although other facilitators did not describe finances as a

motivation for participation, the majority described the financial remuneration as inadequate relative to the time and energy they exerted into the program. Nonetheless, nearly all facilitators articulated a passion for the cause and a commitment to the young people despite financial hardship, and this motivation exhibited a strength in the quality of delivery.

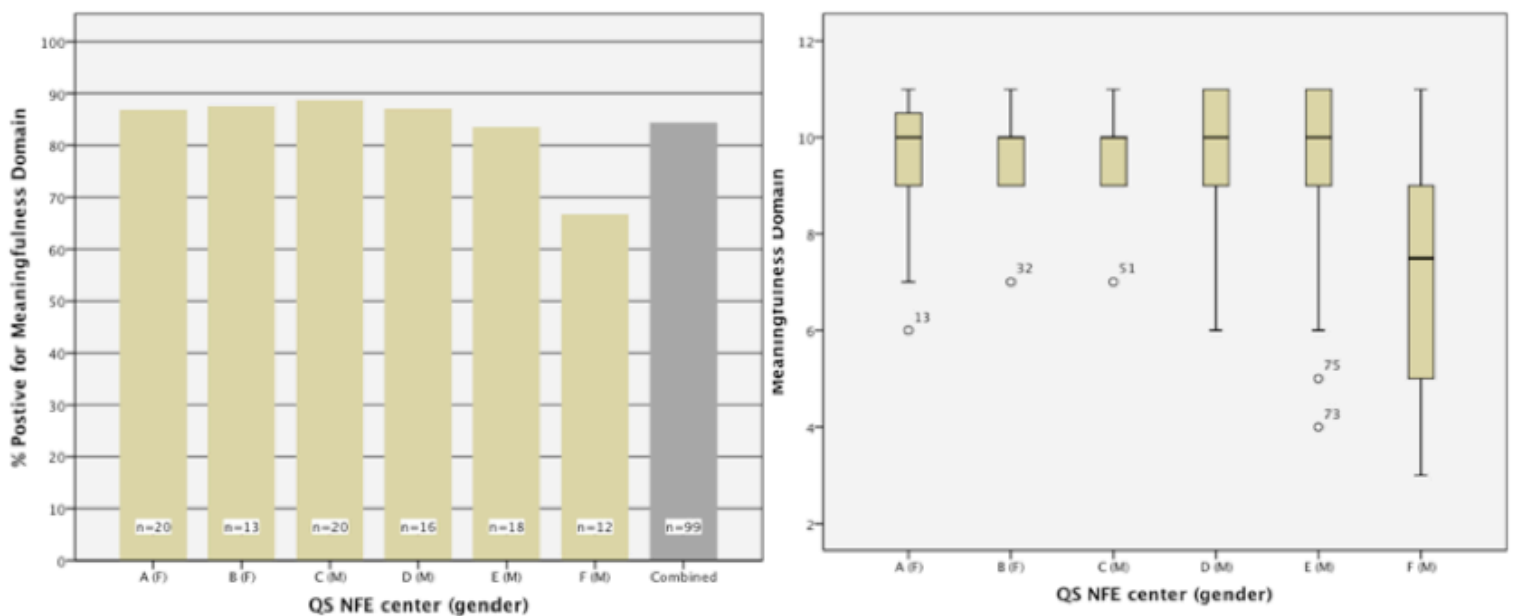
Participant responsiveness

Participant responsiveness indicates “the extent to which participants are engaged by and involved in the activities and content of the program” (Mihalic, 2009). As Bonell and colleagues (2006) contend, an intervention generally must be acceptable by participants in order to be effective. The most direct measure of participant responsiveness in this implementation study was the *meaningfulness* domain contained within the empowerment survey. Meaningfulness assesses “the value of a task in relation to one’s own beliefs, ideals, and standards” (Thomas & Velthouse, 1990). This subscale included items such as “I work hard for the group because I want to, not because I have to,” and, “The things I have to do for the group mean a lot to me personally.” The youths aggregately rated program meaningfulness higher (84.2% positive) than they rated the overall empowerment scale (80.3% positive). Relatively positive overall feedback on this subscale suggested that participants did feel that the program was relevant and valuable for them, which motivates higher engagement (Durlak & DuPre, 2008).

Some youths, such as the following, spoke of the particularly meaningful contribution the program made to their aspirations in the context of otherwise

discouraging life circumstances: “For me, this program is my main goal; it is the most important thing in my life.” As with the overall empowerment survey and other domains, however, feedback varied between centers (see figure 6.6). The average percentage of positive responses for the meaningfulness subscale among centers ranged from 66.6% to 88.6%. There were, in other words, clear differences in participant responsiveness, and these aligned with differences for other aspects of implementation quality and adherence assessed in this study.

Figure 6.6. Meaningfulness domain results (bar graph, left, and box plot, right)



Program participants often expressed engagement in the program due to the comfortable programming environment and recreational activities. When asked why one young person participated in the program, she summarized points that were prominent for many interviewees: “I was attracted to it. They taught us things, and there is entertainment, which is very nice. I mean, it’s an education and entertainment all at the same time. I felt it is better than the formal system; they

have nothing. You just follow the rules; you do not feel comfortable.” Another youth explained, “[W]e are not under pressure; we joke a little bit and we are serious a little bit; it is not all pressure.” The allusions to an engaging and comfortable program setting reinforced the concept of a youth-friendly, prosocial environment often advanced as a critical intervention component in youth empowerment literature (Nygreen et al., 2006; Scheve et al., 2006). Several participants described a range of recreational activities, including arts, sports, and games, as regular program features that excited them and encouraged youths to participate.

Conversely, in cases in which participants did not express a strong level of engagement in program activities, they tended to cite an uncomfortable program environment or a lack of recreational activities in the center. Two participants from the same center, for example, expressed dismay because participants were suggesting more artistic activities—especially in music—and these suggestions had yet to materialize into real activities. The example reinforced the link between the participatory methodology and feelings of program relevance and engagement. In cases where adherence to the empowerment-based methodology was lacking, youths also expressed disengagement with the prevalence or quality of activities.

Program differentiation

Program differentiation identifies the extent to which the critical components of an intervention are discernible from each other and from other programs with which it might be compared. Program differentiation is particularly important for a youth empowerment program (YEP) like QS NFE that must

negotiate inherent tensions between its goals as both a social intervention and an educational intervention. The program was designed first and foremost to foster positive youth development. Teachers are retrained to implement empowerment-based methodologies, and the physical setting and social processes are designed to differ markedly from formal schooling, reinforcing its purpose as a community program for out-of-school youth. Yet, even with these differences and substantial flexibility for facilitators and youths to operate within an empowerment framework, the program must still integrate a specialized educational curriculum to prepare participants for basic testing so that they can, in the end, obtain an alternative 10th grade certificate certified by the Ministry of Education. Moreover, the facilitators are still formal education teachers during the first part of the day, susceptible to reverting to formal education habits. Consequently, if the program wavers, the implementation could default into something that looks more like formal education than the program intends. This tension reinforces the importance of program differentiation.

Program differentiation identifies the critical components of a program that are essential for producing positive outcomes (i.e., 'component analysis') (Mihalic, 2009; Sanchez et al., 2007). The primary program components articulated by the basic theory of change given in chapter four include: (a) youth involvement in program decision-making, (b) asset-building activities, (c) a prosocial environment, and (d) trained adult support. Each of these components was clearly more distinguishable through qualitative investigation in centers with higher

implementation quality and adherence. The following paragraphs address the components in turn.

Component 1: youth involvement in decision-making - As described in the adherence section, all of the facilitator interviewees reported using participatory methodologies in the implementation of the program, typically involving youth in day-to-day decision-making processes concerning learning topics and activities. The frequency and quality with which facilitators implemented youth participation in program decision-making, and the extent to which they did it for *all* youths in the program and not just some youths, varied between centers, according to youth and facilitator feedback. Nevertheless, the participatory component—the most central to youth empowerment's theory of change—was generally articulated as a distinct aspect of the program methodology compared to formal education or other interventions that placed youths in passive learner roles.

Component 2: asset-building activities - Asset-building activities were frequently integrated as part of the day-to-day programming. In many cases, the participatory programming itself served as an opportunity for youths to contribute to decision-making, negotiate collective decisions, and engage in dialogue-based learning. Program centers often integrated team-based athletic activities, crafts, and educational games into the day-to-day operations. In some cases, however, data from youth interviews indicated minimal implementation of stimulating activities with intentional asset-building objectives. One young person, for example, appeared to find more stimulation in his formal education experience than with the QS NFE

activities: “When I used to go to school? There I was learning. Here, when we first come in, we sit at the computers for about half an hour; then they give us lessons.” Comments by this participant indicated that, unlike most centers, this center did not manage to utilize programming hours to facilitate the kind of structured, purpose-driven activities that could promote youth development.

Component 3: prosocial environment - A prosocial environment is perhaps the least tangible component to identify and assess in youth empowerment's theory of change. A prosocial environment implies a program setting that provides a physically and emotionally safe space for young people to express their ideas and interests freely and experience positive, encouraging group interactions. This kind of atmosphere establishes the requisite context for youths to engage effectively in the mastery experiences and collective decision-making processes that define youth empowerment. Qualitative feedback suggested that the intervention was largely successful in differentiating this aspect of the program. Most of the youths described their peer relationships as one of the most positive features of their program experience. Several youths explained how the positive social experience facilitated by the program contrasted with their previous experiences—e.g., “[Before,] I did not talk to anyone. When I came to the center, I met respectable people... Everyone in the program, no one is better than the other. We are all brothers.”

Some exceptions did arise in the qualitative interviews, particularly with participants' experiences in the female centers. At one center, two young females noted tensions between Jordanians and Iraqis, suggesting that that this center had

struggled to facilitate prosocial interactions across ethnic groups as most of the other centers appeared to have done. One of the girls described the dynamics as tensions between “gangs,” and these tensions were confirmed by interviews with the center’s facilitators as well. In general, however, the peer interactions were described favorably. Further to a youth-friendly environment, as mentioned above, several youths spoke of the comfortable, low-pressure program environment as an important aspect of the program experience and unique to their experiences with formal education. Some youths described having been hit or verbally castigated in the formal education system, and the QS NFE was described as a starkly different setting that provided a sense of physical and emotional safety.

Component 4: adult support - Trained adult support was one of the most explicit program components that differentiated QS NFE from formal education and alternative programming, and the other components were largely based on this component. Training constituted the basic premise for translating a participatory methodology, asset-building activities, and a prosocial environment into practice. Facilitators spoke universally favorably of the training and its uniqueness in the context of standard educational practice in Jordan. All facilitators interviewed participated in at least the first of three levels of the QS NFE training curriculum. Problematically, however, none of the twelve interviewed facilitators received the full three-level training curriculum.

A Questscope staff interviewee explained that the full, three-level training curriculum had not been implemented since 2006, and none of the facilitator

interviewees in this implementation study had started prior to 2007; indeed, most included QS NFE centers had not opened until 2007. As such, the training component, by consequence of partial adherence, was not differentiated as much as intended, and, therefore, the other components were also not as differentiated from standard practice as much as intended. Trained adults provided a substantial amount of support to many youth participants, as evidenced by close bonding and encouragement described in many youth and facilitator interviews, but a minority of youth interviews indicated that this supportive adult component was not universally experienced.

Discussion of implementation lessons

The implementation study provided valuable context for better understanding the findings from the impact study that were reported in chapter five. Only one outcome (SDQ conduct problems) showed statistically significant intervention effects at four-month posttest. Largely null findings left the possibilities that (a) the program model was largely ineffective; (b) the program was effective, but more time was required for measureable changes to materialize in the target youth; and/or (c) the program was inadequately implemented to achieve the impacts that the program model was capable of. As a multisite randomized controlled trial of a complex, setting-level intervention in a 'real world' context, this impact study was particularly susceptible to breaches in implementation fidelity. Indeed, heterogeneity in implementation fidelity was thematic across categories of fidelity that were assessed and reported in this chapter. This heterogeneity likely

contributed to diluted intervention effects and has important implications for this and similar YEPs moving forward.

Heterogeneity

Most of the QS NFE centers included in the research did appear to maintain a reasonably high fidelity to the program methodology and components and with strong staff commitment at the point of service. Strong intrinsic motivation toward the mission on the part of facilitators could contribute to a degree of implementation resilience when other aspects of implementation such as training and activity resources wane. However, in virtually all categories of implementation fidelity, at least one or two program centers performed relatively poorly, and all centers had some degree of difficulty with implementation fidelity.

Even under ideal circumstances, effect sizes in psychosocial interventions tend to be small (McCartney & Rosenthal, 2000), and achieving intervention effects grows increasingly difficult as children age into older adolescence (Heckman & Masterov, 2007). Even if just a substantive minority of participants do not experience the program with a high level of implementation fidelity, that could be sufficiently detracting from the results of those who did experience the program as it was intended to reduce effect sizes and preclude statistically significant findings.

Attendance

This problem is particularly compounded when program exposure is low, as was the case in this study with only 47.7% of the sample meeting the attendance standard set by the program (two days per week average). The low attendance for

many youths could reinforce implementation issues for youth empowerment. While involving young people in basic day-to-day decision-making processes, the program—at least at some sites—could have lacked adequate levels of empowerment in planning and implementation to facilitate the degree of program relevance and leadership experiences found in other cultural contexts to correlate with high program participation among adolescents (Altman, 1998; Deschenes et al., 2010). Exploratory analysis on attendance and outcomes reported in chapter five suggested that positive peer interactions may also play an important role in promoting retention. Given that youth empowerment's theory of change is highly process-oriented, program development and evaluation should carefully tend to the implementation of staff development and the experience of empowerment among program participants.

Empowerment

Several youth interviewees spoke positively of their relationship with facilitators as supporting their personal empowerment. In summarizing best practices based on case studies of YEPs, Zeldin and colleagues (2008) emphasized adult staff validating youth culture and encouraging creative forms of self-expression as a core value. Several youth interviewees from QS NFE described the program setting as more flexible and supportive of their youth culture and preferences than their experience with formal education. In some cases, however, youths and facilitators described more obedience-oriented program settings. These attitudes are perhaps unsurprising in the context of the Arab educational systems, which organizational researchers have described as highly authoritarian in nature

(Dwairy, 1998; Neuman et al., 1988), but they run contrary to the counter-cultural program methodology that aims to reformulate youth-adult interactions in terms of partnership and contribution rather than authority and compliance.

Increasing leadership role opportunities for more young people in the program may be a particularly important way to enhance empowerment implementation and outcomes. Several youths and facilitators described leadership roles as a valuable aspect of the participatory process. Further to the point, a survey of a representative sample of 1,822 11th grade students participating in a wide range of after-school activities in the USA found that adolescents with leadership roles reported substantially higher development experiences in the activities (accounting for 8% of variance) than students without leadership roles (Durlak & Weissberg, 2007). Even if youths are participating actively in collective decision-making, this participation may not provide the same extent of skill- and efficacy-building experiences provided through more demanding leadership opportunities possible through QS NFE, such as facilitating a project, coaching a team, mentoring younger youths, or leading a lesson or discussion.

Additionally, empowerment theory suggests that adults play a vital role in facilitating empowerment processes (O'Donoghue et al., 2002; Wong et al., 2010). Diminished levels of training reported in the implementation section may have inhibited implementation of the participatory methodology. The partial training (level one) and update training provided after 2006 may have been adequate to sustain a certain level of implementation adherence to the participatory

methodology, but the overall low adherence to the full training design marked an important constraint in implementation fidelity that likely has implications for more advanced levels of youth empowerment implementation.

Previous literature on professional development with teachers suggests that the most effective models engage teachers as active participants in the development process, are based in and connected to teachers' actual experience in a classroom context, are collaborative, and involve substantial follow-up and interaction (Pianta & Allen, 2008; Sandholtz, 2002). The lack of adherence to level two of the training curriculum (or alternative forms of on-site training), which was designed to provide the more contextualized and participatory professional development experience, forfeited many of the elements that educational theorists suggest are important for changing teacher behavior and sustaining changes.

Moreover, other youth empowerment models have integrated defined participatory training curricula to prepare young people early in the program with social and leadership skills to participate effectively and begin fostering positive youth-adult relationships (e.g., Berg et al., 2009; Suleiman et al., 2006b). Given the added challenges for youth participation in the Arab cultural context described in chapter three, it may be particularly important to facilitate similarly structured capacity-building curricula to strengthen out-of-school youths' competence and confidence for meaningful participation. Further financial and training investment in structured skills-training components could be a worthwhile direction.

Standardization-adaptation balance

The right standardization-adaptation balance can be especially difficult to attain when promoting implementation fidelity with highly process-oriented, complex interventions like YEPs. In fact, Hawe and colleagues (2004) have advocated for a looser perspective toward implementation fidelity with such interventions that emphasizes constancy of essential process ‘functions’ as core features while leaving substantial room for adaptation of the forms of program components at the site-level. This approach can be all the more important for empowerment-based programs for which activities and contents need to respond dynamically to the contributions of participants.

Yet, as Hawe and colleagues also clarified, this flexibility should not translate to an “anything goes” implementation paradigm. Furthermore, previous reviews of studies have found that most adaptations are not planned changes but rather result from a lack of staff understanding, time, or preparation to implement the program as intended (Hill et al., 2007; Mihalic, 2009). Indeed, considering the need for further manualization of the program’s theory of change and best practices, many facilitators’ expressions of feeling overburdened by the workload between formal education and QS NFE, and reduced quantity and regularity of facilitators’ professional development, all three of these issues have relevance for the present study.

Component analysis

The research design used in this project did not accommodate a rigorous component analysis. As such, this implementation study could adequately address differentiation in terms of which components were essential for achieving positive outcomes. Exploratory subgroup analysis presented in chapter five did indicate that the center with the most positive youth feedback on the empowerment survey produced more positive outcomes. This reinforced the hypothesis that youth empowerment is a pathway to psychosocial impacts, and it established a hypothesis for further testing—that the *level* of youth engagement may mediate the extent of positive intervention effects. Future research is needed with appropriate design and statistical power to determine how particular elements of youth empowerment's theory of change related to intended outcomes.

Implementation assessment

Developments in the setting-level youth intervention literature have introduced tools and approaches in order to mitigate implementation fidelity challenges as they scale-up, and these could be explored and perhaps adapted to foster stronger implementation with YEPs like QS NFE. The premise for these resources is that, especially as complex youth interventions expand, stakeholders at all program levels require clear, reliable, and standardized performance feedback at regular intervals to monitor and adjust implementation as they go.

Some tools might be adapted to gauge specific aspects of implementation fidelity as part of an ongoing system. Based on experience with scaling-up the Triple

P parenting intervention to a population-level, Prinz and Sanders (2007) described the need for creating an “evaluation culture” in which evaluation tools and data they produce are central to a program’s operation. For such information to be credible, however, Prinz and Sanders contended that measures need to be change-sensitive, brief, low-cost, and easy to use.

The Learner Empowerment Survey (Fymier et al., 1996) adapted and used in this research, as well as Lakin and Mahoney’s (2006) experimental evaluation of a YEP, could constitute one such resource. Shapiro and colleagues (1997) developed and validated the self-report Youth Client Satisfaction Questionnaire for use with adolescent participants in community mental health services ($\alpha = .87$). Items in this instrument are specific to therapeutic interventions, but a similar instrument for youth empowerment settings could be developed to assess program responsiveness. More specific to youth empowerment, Jones and Perkins (2005a) developed the self-report Involvement and Interaction Rating Scale to “assess the perception and experiences of youth and adult participants working together in community-based efforts” (subscales α ranged from .83 to .87; overall $\alpha = .94$). This kind of instrument could be used to help assess adherence and quality of delivery with respect to youth participation and youth-adult relationships, as well as differentiation on these constructs as program components.

Alternatively, the Youth Program Quality Assessment (PQA) developed by the High/Scope Educational Research Foundation offers a more multifaceted observational assessment tool used to “evaluate the quality of youth programs and

identify staff development needs” (Blazevski & Smith, 2007; Smith & Akiva, 2008). Subscales included in the instrument include safe environment, supportive environment, interaction, and engagement (intraclass correlations for paired-raters were .48, .69, .83, and .72, respectively; Blazevski & Smith, 2007). These domains align reasonably well with the main elements of YEPs outlined in the theory of change. The Youth PQA has the advantage of assisting setting-level youth programs with more holistic information, more defined feedback systems, and observational rather than self-report implementation assessment. Evaluators could also consider other instruments for measuring program quality for out-of-school programs compiled by The Forum for Youth Investment (Yohalem et al., 2009), but, so far, the Youth PQA is the most rigorously tested.

The implementation study reported in this chapter was limited by the absence of such standardized, observational metrics. Properly implementing this kind of tool, however, also requires additional resources, time, and training compared to simpler self-report surveys. Notably, all of the aforementioned tools were developed and tested in the United States. Especially given the cultural nuances of youth empowerment underscored by this chapter and chapter three, researchers need to ensure that any implementation measure systematically used to assess and guide practice in Arab countries is appropriately developed or adapted and tested for the context.

Process study results

The implementation study aimed to assess the intervention's fidelity to its intended components and practices. The process study discussed in this section draws from the same qualitative research (as well as added stakeholder interviews) but addresses a different question. The chapter turns now to the following query: what are the process factors that influence the extent to which youth empowerment is implemented and sustained in local programs? By integrating this investigation into process factors, the research aimed to augment knowledge usability and transferability. Implementation analysis alone highlights strengths and weaknesses in the implementation of youth empowerment, but process analysis uncovers underlying factors that can *explain* how those strengths and weaknesses materialize. Process analysis also provides an opportunity to extrapolate beyond a particular case of implementation and consider lessons in the context of broader social and organizational theory.

Organizational research can study how many different kinds of processes or innovations are or are not diffused through implementation. This study focused on diffusion specifically with regard to youth participation in program implementation—the signature quality of YEPs. A larger-scale organizational analysis of diffusion phenomena would likely have incorporated comparative case studies of multiple organizations, a larger data set, a longer observational timeframe, and extensive use of mixed methods (Hartley, 2004). This process study component could not provide the same level of comprehensiveness, but it could

provide valuable organizational insights into youth empowerment processes and formulate hypotheses for further investigation.

While a small number ($n = 8$) of the interviews were conducted with individuals representing external stakeholder groups (e.g., funders and policy-makers), the majority of interviews were conducted with those directly involved in program implementation at the site-level (facilitators and youth participants), and so this analysis primarily concentrated on process factors influencing youth participation at the program-level. A section is provided at the end reviewing potential external factors that were highlighted by the interviews, though these were not the primary focus of this investigation.

Process mechanisms for youth

The first level of qualitative data analysis explored program process factors that *directly* influenced young people's ability or propensity to engage in program decision-making or implementation. Twenty factor codes were initially established using NVivo 8™, and these were ultimately grouped into three overarching clusters: flexibility, support, and opportunity structure.

Flexibility. Flexibility encompassed the sensitivity and adaptability of the program to the youths' needs and preferences. Youth participation meant constant attentiveness to young people's ideas and desires. In order to accommodate these contributions, the program needed to be constantly responsive so as to adapt activities, lessons, and expectations to the group's expectations. Facilitators like the following described how flexibility enabled them to respond to youths'

contributions: “For example, today, we do not feel like studying; we want to play. Yes, we go out and play football matches, badminton matches... we listen to their suggestions.”

Without the flexibility among facilitators to be responsive, suggestions would go unrealized and youths would feel uncomfortable in the program environment. By contrast, formal education was often juxtaposed to Questscope non-formal education (QS NFE) as a more rigid environment that failed to accommodate youth culture and needs, much less their voice. This unresponsiveness was often cited as a reason for young people’s departure from formal education. A facilitator explained that a critical characteristic of the QS NFE model is that “there is more freedom for the facilitators and the students” compared with the formal education system.

Support. The second category for enablers of youth participation was *support*. In this case, support primarily encompassed positive, encouraging, and trust-based relations with peers and facilitators in the group. One of the youths, for instance, described her relationship with a facilitator: “Take the principal: we laugh and joke with her. I do not feel she is a facilitator; I feel she is my friend.” Supportive relationships established the kind of trust and reassurance many youths needed to participate. As one participant explained, “The first thing is, [the facilitators] understand you. They do not treat you as a school dropout or an ignorant person or things like that. On the contrary, they encourage you; they make you enjoy it.” Youths placed high worth on such horizontal relationships with caring adults.

Some youths expressed how encouragement and support from facilitators made them feel more valued when they tried to participate. “One of the best things,” a youth commented, “is that [the facilitator] makes me feel worthy in front of the other students.” This is consistent with social cognitive theory literature that posits that honest, positive feedback enhances self-efficacy and empowers young people to participate in increasingly challenging ways (Pajares, 2006, p. 350).

Support as a factor for promoting empowerment also included relations between youths and their peers. As one young person explained, “[Before,] I was just sitting at home, isolated from the world. When I enrolled, I met many girls, and I got attached to them and liked them, and I started participating... a lot has changed about my personality.” On the other hand, the implementation analysis above also highlighted how poor peer relations could foster an uncomfortable environment in which youths felt they competed among gangs and were unable to participate meaningfully.

Opportunity structure. Finally, *opportunity structure* was the third empowerment-enabling theme that emerged from data coding. The sociological term “opportunity structure” was originally used by Cloward and Ohlin (1960) as a theory for explaining pathways to success or to deviance. If pathways to success (positive opportunity structures) were blocked, the theory postulated, individuals would often pursue more delinquent opportunity structures through crime, social retreat, or conflict (Marshall, 1998). In this case, the term is used more specifically

and more literally to describe the structures that are provided for youths to participate in program decision-making and implementation.

The “structure” implies that the abstract ability of youths to contribute is insufficient; youths and facilitators highlighted the role of specific mechanisms for encouraging and acknowledging youth contributions in the empowerment process. In most cases, these opportunity structures were described in terms of facilitated “dialogue” or sessions for youths to contribute ideas and make choices about the program. Some interviewees noted that facilitators recorded the youths’ ideas (e.g., on a board), which visually reinforced a sense of recognition and responsiveness. “They must see the outcome of the decision to know that they made the right decision,” said one facilitator, “You implement their decision, which encourages their friends to make decisions.” Similarly, self-efficacy literature suggests that one way for youths to develop belief in their ability to successfully pursue challenging tasks is through vicarious experience of observing their peers do so (Pajares, 1996). Recording young people’s contributions and outcomes of those contributions can be a tangible method to stimulate vicarious experiences to promote young people’s proclivity to participate.

Others mentioned different kinds of opportunity structures for youths to contribute, such as leadership roles and elections. For the two male centers with the highest empowerment survey scores, leadership roles were seen as particularly important mechanisms for fostering youth participation, even through ephemeral roles with recreational activities, such as youths coaching sports activities or

refereeing. Older youths were sometimes given roles in supporting or mentoring younger youths. In general, the more opportunity structure available to the youths, the more engaged they felt. When opportunity structures were lacking, youth feedback reflected a sense of marginalization and defeat. A participant, for example, seemed to lose motivation to participate in the absence of opportunity structure: “Sometimes, I want to express an opinion or participate, but they do not allow me... I tried... they do not listen to all the youths; it’s normal.”

Even in cases where youths reported relatively high opportunity structure to contribute to decision-making through facilitated discussion and participatory sessions, facilitators suggested that more advanced opportunity structures would further strengthen youth participation. Facilitators at one center, for example, advocated for a site-level youth parliament involving regular elections, rotational leadership positions, and youth engagement in more long-term program planning. In centers in which youths felt empowered, young people were enabled and encouraged to assume leadership roles and contribute to decision-making through regular, structured opportunities for participation. In cases where such opportunity structures were described as less prevalent, facilitators were more inclined to describe youths’ behaviors as unmotivated or problematic, and youths tended to describe the program as stricter and their contributions as undervalued.

Process promoters for staff

The second level of qualitative data analysis explored factors that influenced *facilitators’* ability or propensity to implement the aforementioned program

mechanisms that influence young people's participation. Nineteen staff-level factors were initially coded, and these were clustered into three categories: cooperation, opportunity structures, and professional development.

Cooperation. Cooperation emerged as a means to strengthen facilitators' capacity to support youth in participatory processes. Cooperation included communication and coordination between facilitators and their peers, schools in which the QS NFE centers were located, the Ministry of Education, and Questscope.

Facilitators described the importance of easy working relations between facilitators in the center. The fact that each QS NFE center had four facilitators who shared common core training experience provided the opportunity for teamwork in facilitating programming and supporting youth participants. Interviews at two centers indicated tensions or lack of full cooperation between facilitators, and this left facilitators feeling overburdened.

Beyond the facilitator teams, interviewees frequently described the importance of open communications and positive, personal relationships with the institutions that affected the program's implementation—the school, the Ministry of Education, and the organization, Questscope. Cooperation was an important enabler for the facilitators to be responsive to youth participation. Cooperation with schools (and school principals in particular) was generally valued for easy use of facilities. If the school leadership was committed to the program and relationships with facilitators were good, then facilitators were able to use facilities more liberally in ways that supported the flexibility needed to encourage and respond to youth

participation. Cooperation with the Ministry and Questscope was viewed as essential for facilitators to respond with resources, permission, and appropriate skills to decisions and processes that arose from youth participation.

Too many bureaucratic layers that delayed action on programmatic ideas and stifled facilitators' ability to respond to youths' dynamic needs and preferences characterized instances of poor cooperation with the Ministry. Instances of poor cooperation with the organization, Questscope, were typically characterized by inadequate and impersonal communications. In several cases, facilitator interviewees felt that the "role of the facilitator was marginalized" because Questscope would send requirements to facilitators without soliciting and responding to feedback from facilitators. One facilitator summarized frustrations expressed by several others: "When the organization needs information from you they call: 'Hello, we need so and so...' Then, they do not call you, even to say, 'How are you? We do not need anything, we just want to make sure you're okay!'" This sense of lack of cooperation left some facilitator teams feeling isolated from the organization that was responsible for quality assurance.

Opportunity structure. Further to cooperation, one of the most significant process factors that emerged from the data was facilitators' perception of their ability to influence the decisions and implementation of the broader program. Just as the youths required *opportunity structure* to contribute to program decision-making, so too did adult staff. Facilitators frequently expressed positive feelings about the role of the organization in establishing the program and the methodology

and, yet, disgruntled feelings about the current state of active engagement of facilitator teams by the organization. “The facilitators’ participation,” suggested one facilitator, “I mean, they rarely allow us to participate in anything. We only get instructions.” Ultimately, lack of opportunity for facilitators to contribute led to a sense of marginalization and defeat among some facilitators—e.g., “I know you cannot change anything in the organization. They are just problems that you talk about; no more, no less.” It is unclear whether these kinds of sentiments differentiate from most other organizational settings, but, in an organization that aims to deliver empowerment-based programming, these feelings may be particularly threatening for implementation fidelity.

Of all of the process points made by facilitators, this theme of opportunity structures for facilitators to be heard and have an influence in the program had the most references coded under it. Although this dissertation focuses primarily on the empowerment of youth, the empowerment of adults directly responsible for engaging youth emerged as an important enabler in the process.

In most cases, facilitators stressed the value of any kind of structured opportunity for facilitators to contribute to broader program planning and implementation. To illustrate, ten program staff or facilitator interviewees spoke negatively about the lack of opportunity structures for themselves in the program, whereas only five spoke positively of available opportunity structures; reference codings for lacking opportunity structures were three times higher than those for positive opportunity structure references. A facilitator, for instance, suggested, “If

there was a training session or a workshop, and we propose a certain idea, and each person gave his opinion, we can say our voice is being heard. But where is the organization? And how is our voice supposed to be heard and we are not?" In other cases, facilitators, like youths, emphasized specific roles to contribute, such as leading training for other facilitators or organizing youth camps.

One area where facilitators did generally feel they had relatively high opportunity structure to lead and contribute was in the program center itself. As one facilitator stated, "I think that I have an influence in NFE in terms of implementing the program." The ability of facilitators to control much of the day-to-day programming was an important enabler for youth participation because it allowed facilitators the flexibility to respond to youth contributions in the day-to-day decision-making processes about group expectations, learning discussions, and activities. For this reason, another facilitator explained, "There is more freedom for the facilitators and for the students." This freedom provided facilitators with an opportunity structure that better enabled them to implement process mechanisms behind youth participation.

Professional development. Professional development was the third factor that promoted facilitators' ability and propensity to foster youth participation. Most facilitators recognized the Questscope training as the primary medium through which they adopted the participatory methodology. The following facilitator statement represented this trend: "I thought it was like being a regular teacher, but when we underwent the training, I felt differently... The way to treat these girls;

how to engage them and encourage them to come back to school, and be attracted to the center... They taught us all these things and we are applying them.” Substantive staff training in participatory methodologies was clearly an important enabler for implementing youth empowerment across program sites. In general, facilitators felt that the quality of training they received was high but that the amount and regularity of professional development over recent years was inadequate. This likely contributed to some implementation deficiencies as described in the implementation section of this chapter.

Professional development was also proposed by facilitators in terms of more regular opportunities for facilitators to meet with each other in structured forums to share knowledge and provide trainings for each other. These kinds of opportunities can help facilitators reinforce each other’s efforts in facilitating youth empowerment and exchange on best practices.

Individual and external factors

The qualitative research made clear that youth empowerment is not implemented in isolation. Previous sections focused on prominent process factors within the organization that appeared to promote or obstruct implementation fidelity to the methodology. Nevertheless, qualitative data did present numerous other factors at the level of individual facilitators and factors external to the organization that could have important implications for the implementation of youth participation. While this study was not designed to explore these factors and their relationships to youth empowerment implementation in depth, external

factors could play important roles in explaining implementation and therefore are addressed briefly here.

Facilitators' attitudes. Firstly, facilitators bring their own characteristics to the program that can moderate their propensity to implement empowerment-based programming. Perhaps most importantly, these characteristics include facilitators' basic biases and attitudes toward target youth. Such attitudes differed fairly substantially between facilitators. Some expressed strong beliefs in the participation of youths in program decision-making and were of the mind that the youths should have increasing opportunities to contribute. On the other end of the spectrum, some facilitators had relatively limited belief in the youths' ability or desire to contribute to program decision-making. Some facilitators were committed to youth participation but mostly for older youths or those who had been with the program for longer.

Youth empowerment programs also operate inside of complicated and dynamic broader social, institutional, and political environments that can significantly influence the extent to which programs successfully implement participatory methodology. External factors that emerged from the qualitative data as potential influences on the implementation of youth empowerment at the program level included: (a) cultural orientations toward target youth, (b) the availability of funding and characteristics of the funding source, (c) and the amenability of related policy to the program methodology.

Culture. Broader cultural orientations toward out-of-school youth in the Middle East can affect the biases that both facilitators and youths bring to program implementation and participation, and they can also influence the extent and quality of support for the program methodology from other actors, such as donors, policy-makers, and community stakeholders. As one organization staff member contended, in the view of broader society, the youth involved in QS NFE “are not important.” These views mirrored negative perceptions of out-of-school and street youth demonstrated elsewhere in the region, as reported in chapter three. Also as described in chapter three, research into Arab culture has revealed stronger orientations toward authority and vertical power relationships than other cultures (Dwairy, 1998; Neuman et al., 1988), which would likely influence the paradigms that facilitators and other stakeholders bring to implementing a YEP. Such context adds challenge to implementation fidelity.

Indeed, the cultural context for QS NFE in some respects may not necessarily reduce or increase youth participation but rather alter the way in which it is implemented. As predicted by Hart (2008) for youth empowerment in less individualistic societies, youths and facilitators were much more likely to describe youth participation as a collective process than an individual process (a group making program decisions rather than an individual). As for adult roles, even in centers showing indications of higher youth empowerment, adults maintained a degree of authority-based respect and distance in the relationship with youths, and youth interviews showed no signs of disillusionment with this arrangement so long as youths felt supported and able to contribute. Although the data affirmed the

relevance of youth participation for QS NFE's cultural context, it reinforced the importance of implementing youth participation in such a way that negotiates within local values and norms.

Funding. The nature of funding can also influence youth empowerment processes at the program level. As the organization and the QS NFE program expanded over recent years, the organization mainly relied on large grants from international aid agencies to fund the growth. Donor expectations for these grants were principally tied to outputs (e.g., 8,000 youth participants over two years), and this appeared to foster a tension between the program's historical emphasis on empowerment and changes in youth outcomes and taxing institutional burdens on reaching and documenting output targets for major funders. A Questscope staff interviewee recounted the dilemma:

Then the donors change their requirements. You know, 'we need this, we will do that, the target group should be like that...' So, then here we started to run after these things, and this keeps us away from doing other things, which are related to the quality. So our concerns started to be the quantity—the target group, the numbers. We have to meet this, we have to meet that... My impression... I think, somehow, in the last two years, we started to focus more on quantity rather than quality. Now there is a need to rebuild that quality and that path of the NFE as NFE.

Easterly (2002) and Edwards and Hulme's (1996) anticipations that this kind of funding relationship pushes non-governmental organizations (NGOs) toward more top-down organizational behavior seemed to materialize in comments by several facilitators that perceived the organization as increasingly distant from the program and the facilitators' roles as increasingly marginalized toward mechanistic

documentation of outputs rather than meaningful involvement in program development.

Funding may have particularly adverse effects on implementation fidelity when funder orientations toward the target youth contrast with those of the program methodology. In an interview with a representative from one of the program's main funders, the interviewee dismissed youth participation as a viable intervention strategy for the program's target youth until they reached the final program cycle: "You can't believe that [the youths] will do this high level of skills you are asking for... I mean that participation should be at the end of cycle three. They are not ready. You take them from the streets; you cannot tell them, 'you are responsible now.'" Instead, the interviewee contended that the program should be "giving information in a traditional [educational] way."

The view represented stark contrast with the intervention model, which assumes that youth participation is the *means* for engaging and impacting young people (especially early in the program) and that traditional educational approaches rebuffed rather than attracted target youth. The contrast between donor orientation and the program's theory of change highlighted how donor characteristics could compound funding-related factors that adversely influence a program's ability or propensity to implement empowerment-based methodologies with target youth.

Policy. Thirdly, YEPs operate within a political and policy context. In the case of QS NFE, which constituted a joint venture between a governmental and non-governmental organization, political and policy context was particularly pertinent.

In some respects, the political context appeared to promote youth empowerment at the program level. The Ministry's willingness to adopt the QS NFE program through a memorandum of understanding (MOU) and enthusiasm for the methodology expressed by Ministry of Education representatives were essential assets for diffusing youth participation. "The best thing about the methodology," stated a Ministry official, "is that it depends on learning through participation." This commitment to youth participation was reflected in the Ministry's cooperation with the organization in designing and certifying a uniquely flexible curriculum that enabled facilitators to build activities and learning topics around youths' ideas and preferences.

On the other hand, qualitative data also revealed instances of how the immediate policy context could indirectly stifle youth participation. As the Ministry became more involved, along with Questscope in overseeing the program, particularly centers hosted in schools, government staff and procedures became more involved in some day-to-day matters of program operations. Consequently, some facilitators reported frustrations with additional levels of bureaucracy within the government that they had to overcome when youths or facilitators wanted to introduce certain new activities or guest presenters to the program. "There is too much administrative work," explained one facilitator, "For example, the plan needs to be signed by three people." Ministry standards also limited the amount of remuneration that could be offered to facilitators, which could contribute to strain on facilitators and influence their quality of delivery.

Discussion of process lessons

The qualitative process research aimed to explore organizational factors that promote or obstruct the diffusion of youth empowerment at the program-level. The themes that emerged from the data were not meant to be exhaustive, and the research design was not comprehensive enough to fully saturate categories (a methodological aim of grounded theory research) and establish definitive conclusions or theories. Nevertheless, the investigation did present prominent themes from this particular program case, and these can be used to further explain implementation and impact study results and generate hypotheses for more intensive organizational investigation into the diffusion of youth empowerment.

As social constructionism implies, ideas are formed through participants' interactions with the world around them. As such, this process study tried to better understand how youth and facilitators interpret the program setting and the environment-level factors that appear to help explain those interpretations. The process study was not designed to make positivist, objective claims about relationships between variables, but it can set-up hypotheses to be tested through more deductive methods later on.

Process mechanisms for empowerment at the point of service

Themes that emerged from the qualitative data for supporting youth empowerment at the point of service are largely reflected in the extant literature of similar case studies. Cargo and colleagues (2003), for example, conducted qualitative research with youths and adults in the United States to investigate

process factors that support youth participation at the community level. The themes that arose as mechanisms for the point of service (“adult subprocess” factors) included a *welcoming social climate* (providing opportunities, believing, respecting, encouraging, caring) and *enabling youth* (facilitating, teaching, mentoring, and providing feedback). The present research has some parallels. Several descriptors under both of Cargo and colleagues’ themes (e.g., respecting, encouraging, mentoring, feedback) would be included under *support* in the present study, and this research suggests a somewhat more specific and prominent placing of the “providing opportunities” theme under the separate category, *opportunity structure*. *Flexibility* also emerged as an important theme in the present study. While not a key finding of Cargo and colleagues’, the finding was consistent with other youth empowerment case studies that described flexibility as an important enabler of meaningful participation (Nygreen et al., 2006; Ozer et al., 2008).

The feedback concerning meaningful roles for contribution that arose in the present study’s data may add additional relevance of role theory’s (Biddle, 1986) place in empowerment theory. Youths and facilitators described not only the value of opportunities to participate in program planning and implementation, but, in some cases, they emphasized the value of clear, structured leadership roles to do so. Similarly, Maton and Salem (1995) described the leadership development function of what they call “opportunity role structures,” which involves creating settings with many distinct roles in a setting for individuals to participate, learn, build deeper relationships, and develop confidence through experience (Speer, 2008). Whatever the form, opportunity structures emerged as important mechanisms for youth

empowerment. In the following excerpt, Pianta and Allen (2008, p. 25) emphasize the role of opportunities “carefully crafted” for youth to contribute in educational settings:

Any setting that intends to advance development and learning outcomes for youth must carefully craft the nature of experience it provides to give adolescent participants a sense of control, autonomy, choice, and mastery. Absent these considerations, or in settings that rely on top-down approaches, classrooms for youth are doomed to be places lacking in engagement and motivation. One of the most tragically avoidable errors that some secondary school teachers make is to assume that youth strivings for autonomy and self-expression represent negative forces to be countered rather than positive energies to be harnessed... This mismatch of classroom and development, driven by profound misunderstanding of teens, results in schools narrowing, rather than expanding, the “space” in which zones of proximal development can be created for youth.

Although QS NFE is at least as much of a social program as an educational one, Pianta and Allen highlighted empowerment themes that could be transferable across settings.

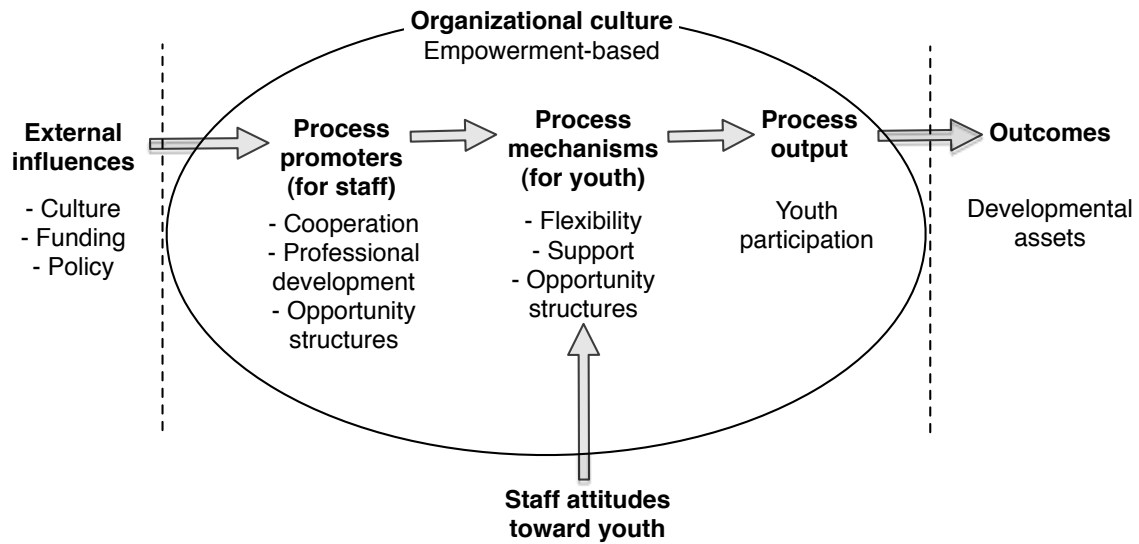
Process enablers of empowerment at the staff-level

Like most research into youth empowerment processes, the Cargo and colleagues study did not go on to assess the process factors that encourage or discourage *adults* from implementing mechanisms that facilitate youth empowerment. In lessons from a multiple case study analysis of model YEPs in the USA, Zeldin and colleagues stressed the importance of building an “organizational culture of partnership” (Zeldin et al., 2008, p. 17). A key trait for the examined programs was establishing youth-adult partnerships as a core organizational priority, not a side project or methodology narrowly defined to a particular piece of an organization’s portfolio.

As posited by Smith and Akiva, “We need to know if the organizations and systems in which the moment of adult-youth interaction is contextualized are helping or hindering the efforts of staff at the [point of service]” (2008, p. 208). Zimmerman reinforced the point: “Efforts to understand empowering processes and outcomes are not complete unless multiple levels of analysis are studied and integrated” (2000, p. 46). The diagram given in figure 6.7 suggests a basic conceptual model of how factors influencing youth empowerment (those that emerged from this study’s qualitative data, which are not intended to be exhaustive) can operate at multiple levels, both within the organizational system and externally to it to influence an organization’s culture of empowerment and ultimately its ability and propensity to diffuse youth empowerment at the point of service.

The first lesson that surfaced from the sum of the qualitative data was that empowerment (or disempowerment) functions and interacts at multiple levels. It appears difficult for staff to fully understand—much less fully implement—a process with which they themselves do not have direct experience. Much of the youth empowerment literature examines the extent to which adult staff empower youth, but very little of it examines the extent to which the adult staff feel empowered within the organizational or system context and the implications this might have for youth empowerment. As social constructionism would suggest, adults’ fundamental notions about empowerment are influenced by the nature of their own experiences (or lack thereof) in their social settings; these socially-shaped notions have implications for how facilitators implement the program methodology.

Figure 6.7. Factors influencing implementation of youth empowerment programming



Note: This conceptual model is not intended to be exhaustive but rather to illustrate themes that emerged from the qualitative data. It posits a series of hypotheses. Further research, using quantitative methods if possible, should test the hypotheses. It is likely that the relationships run in multiple directions, but this diagram presents a unidirectional pathway, as it focuses on process factors that influence youth participation.

Most empowerment research, as Zimmerman (2000) pointed out, has concentrated on the individual as the level of analysis. Given issues regarding facilitator empowerment raised from this case of a YEP, investigations of empowerment at the level of organizations or communities are merited to examine how feelings of empowerment interact across social systems. Schulz and colleagues described "empowering organizations" as those that are "cooperatively controlled by their members and work toward goals defined by those members within the parameters of external opportunities and constraints" (Schulz et al., 1995, pp. 3-4). By this criterion, the qualitative evidence made it difficult to classify QS NFE as nested within an empowering organization on the basis of facilitator feedback. While there is no technical requisite for a YEP to be implemented within an

empowering organization, this process study raises questions about the extent to which high implementation quality of youth empowerment is viable in the context of an organization in which staff perceive inadequate opportunity structures for influence in the organizational setting.

Concerned with the susceptibility of social workers to a sense of “professional malaise” resulting from “individual and professional powerlessness,” Frans (1993) developed a scale for measuring social worker empowerment. To this author’s knowledge, however, no research has attempted to employ such an instrument in youth programs in tandem with an instrument to measure youth participant’s perceptions of empowerment in order to investigate potential relationships between the two sets of data. The data generated in the present study indicate that this would be a useful endeavor for future organizational research.

Notably, some research has shown that, in the school context, higher teacher self-efficacy predicted higher student self-efficacy and academic achievement (Hoy & Davis, 2006, p. 120), and higher teacher enthusiasm predicted higher student intrinsic motivation (Patrick et al., 2000). These studies reinforce the notion of potential social contagion effects between adults and youths in program settings. If the organizational context fosters a higher sense of empowerment among facilitators in YEPs, it could cultivate an environment in which youths too feel more efficacious and engaged.

In terms of a second predominant process lesson, qualitative data also suggested that professional development was important for equipping facilitators to

implement empowerment with high fidelity. Consistent with empowerment theory, in fact, competency-building would be an essential aspect of facilitators' own empowerment (Fymier et al., 1996). In a diffusion study of youth-led participatory research in urban schools in the USA, Ozer and colleagues (2008) found that both general and innovation-specific capacity-building for school staff was important to the implementation of empowerment programming. Beyond discrete training schemes, however, QS NFE facilitators also described exchanges with other facilitators and partners as a valuable means for professional development in youth empowerment.

Others researching diffusion of innovation in organizational contexts have echoed similar recommendations for 'multi professional fora' or 'communities of practice' for individuals to engage in knowledge sharing and meaningful professional exchanges (Dopson, 2006; Storck, 2000). Similarly, Smith and Akiva (2008) posited the need for 'professional learning communities' (PLCs) to foster high implementation quality in youth development programs. Beyond improving the logistical ease to implementing youth empowerment, Ozer and colleagues (2008) suggested that frequent cooperation and communication among stakeholders could foster an organizational culture more supportive of youth empowerment.

Given the counter-cultural nature of youth empowerment in the context of an otherwise largely didactic and top-down educational system, facilitators who also serve as school teachers in the first part of the day can be highly susceptible to

returning to disempowering approaches that dominate social settings outside of the QS NFE program. As such, frequent refreshment training and exchanges with other facilitators and program staff may help sustain commitment to the methodology.

Context factors for empowerment implementation

Factors that extended beyond the organization were important contributors to staff members' ability and propensity to implement empowerment. Notably, different cultural contexts do not necessarily imply *more or less* empowerment; they could also influence the *approach* to empowerment. In challenging his own 'ladder of children's participation,' Hart (2008) suggested that, despite a common interpretation of youth empowerment in Western society to mean youth being "in-charge," in many contexts it may be more appropriate to consider participation as recognizing "the rights of others to have a voice and involv[ing] them." In other words, in contexts like Arab society, where respect and distinction of elders is pronounced, it may be particularly important to recognize the role of adult involvement in the empowerment process rather than to stress that "liberation from adults is the end goal" of youth empowerment efforts (Hart, 2008, p. 24).

One of the factors that emerged was the attitude that facilitators brought to the program. Similarly, in researching the implementation of youth participation in classrooms, Jones and colleagues (2008, p. 60) found that the "alignment of teachers' own values, beliefs, and perceptions of ability with the underlying pedagogy of a particular intervention is critical to their ability to understand, accept, and implement the intervention, and to the effectiveness of the intervention itself."

Mechanisms at the organizational level, such as professional development, cooperation, and opportunity structures, affect facilitators' implementation of youth empowerment, but facilitators' own beliefs influence how they approach youth and the extent to which organizational conditions can diffuse the methodology through the facilitators' implementation.

One study ($n = 108$) on adult attitudes toward youth found that adults in programs that promoted youth participation were significantly more positive toward youth involvement and youth-adult interaction than those in adult-led programs (Jones & Perkins, 2006). Because the study was cross-sectional, the direction of the relationship cannot be determined. It is conceivable that more empowerment-oriented adults were more likely to be selected or self-selected into programs that matched their values. On the other hand, it is also possible that, to some extent, programs that implement more youth empowerment-based methodologies could influence adult attitudes toward youth participation.

According to Allport's *intergroup contact theory*, bringing members of heterogeneous groups together has positive impacts on 'in-group' members' (i.e., those possessing power, privilege, and status) attitudes toward 'out-group' members (i.e., those outside of the in-group circle) when intergroup relationships include: (1) equal status among group members; (2) active, common goal-oriented endeavors; (3) intergroup cooperation that exposes members to personal qualities and skills of others in the group; and (4) a sense of shared values (Allport, 1954; Jones & Perkins, 2006). Accordingly, when YEPs are implemented with high fidelity

to the methodology, adults may be more likely to adopt strength-based attitudes toward youth that in turn reinforce stronger implementation of youth participation. In other words, while facilitator attitudes may affect implementation of youth empowerment, implementation of youth empowerment may also affect facilitators' attitudes toward youth. The relationship between the two variables should be investigated in future research.

Scaling-up effects on empowerment implementation

Another point that arose from the process study was that the scaling-up process could place significant strain on diffusing youth empowerment. This can be said for many types of psychosocial intervention implementation, but the highly process-oriented nature of youth empowerment, as well as its methodological novelty in the local social and educational context, makes implementation of youth empowerment all the more difficult in scaling-up.

Qualitative interviews with various stakeholder groups indicated that the scaling-up process with QS NFE had deleterious consequences for implementation fidelity. Having started in 2004 with a few community-based organizations, with the partnership of the MOE and attainment of large project grants, QS NFE has since expanded to include 57 sites (schools and community-based organizations) across Jordan and more than 7,000 participants. Rapid expansion can strain implementation fidelity, and the nature of international aid resourcing such expansion can deepen the problem by exerting disproportionate pressure on outputs and top-down organizational processes at the expense of program

emphases on impacts and the quality implementation processes believed to achieve them.

Much has been written on the nature of foreign aid documenting it as disproportionately focused on outputs (e.g., money spent or people served), as well as reinforcing top-down bureaucratic NGO operations in serving the poor (Easterly, 2002, 2006; Edwards & Hulme, 1996). As Easterly put it, “[A]id visionaries... map out top-down plans at the same time that they emphasize ‘local ownership’ and emphatically deny that they favor top-down planning... Unfortunately, decades of participation rhetoric have not changed the balance of power in foreign aid” (Easterly, 2006, pp. 196-197). The qualitative research gave indications of interactions consistent with these funding patterns.

Pfeffer and Salancik (1978) theoretically captured the tendency of organizations to alter their structures and goals to obtain resources to survive in the *resource-dependency model*; this kind of change tends to work adversely against empowerment. As Powell and Friedkin (1987) observed in a comparative case study analysis of the change processes for several large non-profit organizations in the USA, “The difficulty with external sources of financing... is that it can lead organizations to change in unanticipated ways.” For this reason, Bradach (2003) argued that programs that financially “struggle to stay afloat” are not good candidates for replication, even with good results. Bradach contended that programs without a reliable source of money are susceptible to decreased implementation fidelity as program leaders spend more time and energy on

fundraising than program development and programs more desperate for resources compromise their objectives or practices to satisfy funders. Unplanned change reflected in QS NFE's growth process has hindered empowerment of staff working at the point of service and reduced adherence to professional development.

The challenges for maintaining youth empowerment through program expansion could be identified in what Selznick (1957) recognized as the "the tragedy of organization" in nonprofit organizational behavior: "[T]he tendency of organizations to defeat the very purposes for which they were established" as operational goals and procedures supplant purposive goals (Powell & Friedkin, 1987). Similarly, Snook's (2000) organizational theory of practical drift concerned "the slow uncoupling of local practice from written procedure" (Snook, 2000, p. 225). Like Selznick's tragedy of organization theory, this implies a gradual departure of practice from initial goals and expectations for implementation. In Selznick's theory, however, practice drifts more toward bureaucratic distractions, and, in Snook's theory, practice drifts more toward local habits or biases on how to implement. Conceivably, both were to some extent true in the case of QS NFE, in which case expansion goals and donor expectations pressured Questscope into an increasingly operational focus on inputs and outputs, and this deemphasized empowerment-based methodologies leaving program staff more susceptible to local drift from written or trained practices for program methodology.

In order to minimize risks of deviation from strong implementation fidelity as social programs expand, Bradach (2003) emphasized the role of a clearly

articulated and disseminated theory of change, which standardizes as many elements as appropriate. Through the process of this research, Questscope has come to more clearly define and document its theory of change and key elements, which could help strengthen implementation fidelity as the program advances. The lack of a sufficiently documented and simplified theory of change supported by clear practice manuals going into this study, however, likely contributed to a level of heterogeneous language and implementation among facilitators.

Bradach further suggested that the role of the “national” entity—in this case Questscope—is to (1) ensure quality and protect the program brand, (2) facilitate learning throughout the network, and (3) provide centralized services for program sites. Yet, as both Bradach and this study’s stakeholder interviews suggest, funders tend to concentrate funding on direct program activities and are perennially reluctant to finance the role of the organization in supporting implementation fidelity. Organizations might be more successful in obtaining funds for their role in supporting implementation fidelity across multiple program sites by more clearly articulating their theory of change to prospective donors and the organization’s plan for monitoring and enhancing program quality and results. Donors, however, also need to better align their funding values and practices to match established frameworks for evidence-based practice.

Concluding remarks on implementation and process

This implementation and process study had limitations. Although procedures were set in place to gather a range of views into the process and implementation

factors, there was room for bias in sampling that could have influenced the data. For example, youths who were not attending the program regularly were less likely to participate in the qualitative interviews and the empowerment survey. As such, the views of this group of youth were likely underrepresented. Additionally, there was opportunity for interviewees such as facilitators to self-select into participation, which could have introduced bias through the characteristics of facilitators represented in the data.

The implementation study was constrained by the use of qualitative interviews and youth self-report measures. The investigation would have benefited from a culturally-validated program quality assessment framework including observational measurement. The development or adaptation of such an instrument, as well as further articulation of the standardized elements in QS NFE's theory of change to guide the assessment, would benefit future evaluation of implementation fidelity for the program. As mentioned before, the process study was less comprehensive than more intensive multi-method, longitudinal case studies of other examples of diffusion of innovation or other organizational behavior phenomena. The present study was not designed to enable the same level of breadth or thematic saturation.

Despite these limitations, the implementation and process research had the strengths of a mixed-methods investigation involving multiple stakeholder groups. The study revealed important context for better understanding the impact study results and informing program improvement. In doing so, the study contributed to

more useful, multidimensional intervention research that aligns with the MRC Framework for Complex Interventions' emphasis on conducting process research along with summative impact research. The research also makes contributions to the youth empowerment literature, particularly with respect to applying organizational lessons and theory to knowledge of the factors that influence the extent to which youth empowerment is or is not diffused into program implementation. To this end, the process study engendered directions for further research connecting organizational research theory and methods with the study of youth empowerment.

The QS NFE program had many implementation strengths, including a participatory program methodology that largely provided a vulnerable Arab youth population with opportunities to influence program processes and develop positive connections with peers and adults. The key elements of youth empowerment's theory of change were generally identifiable in the program and differentiated QS NFE from the youths' and facilitators' experiences with the formal education system. These strengths were reinforced by a strong personal commitment to the mission expressed by implementing staff.

On the other hand, program centers varied, in some cases substantially, in implementation fidelity to the essential practices that define the program and youth empowerment generally, and data from all centers revealed some degree of implementation difficulty. Not all centers or facilitators fostered adequate levels of youth participation or high-quality relationships with all participants.

Implementation gaps were compounded by, and perhaps connected to, very low attendance rates in some centers that translated to suboptimal exposure. These implementation heterogeneities likely contributed to low overall intervention effects on developmental assets as reported in chapter five.

The ability and propensity of program staff to implement youth empowerment are influenced by process factors operating both within and external to the organizational system. Researchers and stakeholders interested in the application of youth empowerment in programs must look beyond the narrow perspective of youth-adult interactions to understand contributors, especially cultural and organizational factors, to an empowering setting. This study highlighted how empowerment (and disempowerment) can operate at multiple interacting levels. As YEPs expand, the maintenance of an empowering organizational culture at multiple system levels can become increasingly difficult.

In order to maintain high fidelity to implementing youth empowerment through replication, programs like QS NFE (and their funders) may be able to increase the resilience and advancement of participatory methodology by integrating credible, ongoing assessment and feedback systems; regular skill-building and knowledge-sharing opportunities for both staff and youth; and high levels of communication and coordination between program centers and national entities to help staff at the point of service maintain the responsiveness and enthusiasm needed to effectively implement youth empowerment with vulnerable populations.

Chapter 7: Discussion

Introduction

Youth empowerment programs (YEPs) are increasingly promoted and implemented around the world as a means for developing adolescents' skills, attitudes, and behaviors, which are expected to increase resilience and positive outcomes across a range of important life areas. Drawing from previous literature, this dissertation defines YEPs as programs in which young people are regularly involved in program decision-making, supported by adults or older youth facilitators, participating in asset-building activities, and part of an environment that fosters prosocial interactions between peers and adults. The fundamental premise of empowerment theory and YEPs' theory of change is that the process of engaging program participants in planning and implementation with supportive facilitation, regardless of the particular content of the activities, is expected to build on individuals' strengths and develop positive assets, such as self-efficacy and social skills.

Despite the growing attention to youth empowerment, little is known about the extent to which YEPs actually affect outcomes of concern, which outcomes are affected and for which populations, and the components and process mechanisms that are responsible for intervention effects or lack thereof. This dissertation comprises multiple, progressive research components, guided by the MRC Framework for Complex Interventions, to contribute to early evidence concerning

these questions. The present chapter reviews findings from a systematic review, randomized controlled trial, and implementation and process study of youth empowerment programming conducted in the context of this dissertation. Youth empowerment's theory of change is revisited from chapter one in light of the evidence, and overall lessons are discussed followed by implications for research and practice.

Summary of findings

This section summarizes key findings from each of the dissertation's research components: the systematic review, the randomized controlled trial (RCT), the implementation study, and the process study.

Systematic review

Research question: What is the state of the evidence on the effects of YEPs on adolescents' self-efficacy, self-esteem, and other developmental outcomes?

1. Youth empowerment programs, although widely implemented, are largely untested and their effects are uncertain. Only two includable studies, excluding the Jordan RCT reported in this dissertation, were identified. One was very small ($n = 40$), and both had significant methodological and reporting limitations.
2. There were no statistically significant intervention effects among included studies on the review's primary outcomes (self-efficacy and self-esteem). A minority of secondary outcomes did show positive intervention effects in the

- individual trials but were not meta-analyzed due to heterogeneity of measures; these effects primarily included problem behavior outcomes.
3. Results were more encouraging from excluded school-based YEP studies ($k = 3$), but these too were very limited in terms of methodological weaknesses and could not be meta-analyzed.
 4. All of the discussed included ($k = 2$) and excluded studies ($k = 3$), excluding the Jordan RCT reported in this dissertation, were conducted in the USA.
- Impact studies of YEPs in other cultural contexts are thus especially lacking.

Randomized controlled trial

Research question: What is the impact of empowerment-based non-formal education on out-of-school Jordanian youths' social and emotional outcomes at 4 months?

1. There were no statistically significant overall intervention effects on the primary outcome (general self-efficacy) or on the majority (10 out of 11) of outcomes measured.
2. There was a significant intervention effect of medium effect size on SDQ conduct problems (95% $CI = 0.13$ to 1.48 ; $d = .57$).
3. Most of the positive changes in outcomes, including for conduct problems, were due to changes in the younger age group (13-15) rather than the older age group (16-21).
4. More positive changes, especially in SDQ prosocial behavior, were observed in the program center in which youth felt the most empowered. Participants with higher attendance also reported more positive change in friendship

supports. The results suggest mediating effects of empowerment implementation and attendance, but low power rendered these findings useful for hypothesis-building rather than conclusions.

Implementation study

Research question: How well was the Questscope non-formal education program implemented as intended?

1. There was overall positive youth feedback on program empowerment, but there was variability between centers that may have circumvented statistically significant intervention effects.
2. Facilitator training quality was perceived as high, but frequency and amount of training and professional development opportunities were perceived as low; this may have contributed to variations in implementation fidelity.
3. Exposure was particularly problematic, with only 48% of the intervention group meeting the basic attendance-level expectation (two days per week) set by the program.

Process study

Research question: What are the process factors that influence the extent to which youth empowerment is implemented and sustained in local programs?

1. Diffusion of empowerment into program implementation was likely influenced by factors at multiple levels of analysis.
2. Qualitative data showed program flexibility to youth's ideas and preferences, social supports, and opportunity structures for youth participation in

decision-making as prominently identified factors influencing youth empowerment at the point of service.

3. At the organization level, cooperation between facilitators and their colleagues and partner institutions, ongoing and dynamic professional development opportunities, and opportunity structures for facilitator participation in decision-making emerged as themes that promoted *facilitators'* ability and propensity to implement youth empowerment.
4. External influences, including cultural orientations toward target youth and participation, funding nature and availability, and policy context appeared to influence the *organization's* ability and propensity to foster high-quality implementation of youth empowerment.
5. The degree of youth empowerment in a program setting may be connected to the degree of facilitator (adult staff) empowerment in the organizational system; this should be empirically studied.

Discussion of findings

Systematic review

Despite the considerable amount of literature and institutions promoting the believed impacts of YEPs on positive attitudes and behaviors, this review concluded that there is thus far insufficient empirical evidence to adequately substantiate the claim that YEPs are effective in changing youth outcomes. None of the three included studies independently demonstrated significant intervention effects on this review's primary outcomes, self-efficacy and self-esteem, nor was there a significant effect in the limited, two-study meta-analysis of self-efficacy that was possible. Many

outcomes had null effects, and, for all three included studies, these outnumbered the outcomes that did show significant intervention effects. All three included studies assessed self-efficacy, but none showed significant intervention effects compared to control groups at posttest. On the other hand, no study showed evidence of harm according to the outcomes measured, nor were there any studies that failed to produce at least one statistically significant, positive intervention effect for the review's secondary outcomes—including social skills (team skills), coping skills (proactive coping), and problem behaviors (conduct problems, marijuana use, and number of sex partners).

A number of factors may help explain the insubstantial impact results of YEPs to-date. Firstly, the mere size and quantity of studies renders any findings—positive, null, or negative—inherently inconclusive. Given the complex nature of studying YEPs and the dearth of experimental or quasi-experimental studies available, the trials to-date could be viewed as much as pilot studies to prime the field for further rigorous evaluation as they could be viewed as contributors to the evidence base. Additional and larger studies in the future, building and improving from these early design examples, could identify more compelling intervention effects if and where they exist. The most important point that emerged from this systematic review was that the existing evidence base is thin, and more research using experimental methods when appropriate is needed to rigorously assess the effects of YEPs.

Secondly, study design and intervention implementation limitations among included studies could have diluted or precluded intervention effects. In terms of study design, one trial (Berg et al., 2009) used a quasi-experimental design and had differences between groups at baseline that could have confounded results at posttest. One of the RCTs (Olson-Merichko, 2006) comprised a very small sample size ($n = 40$) for this kind of intervention without any statistical justification (e.g., a power calculation) given, and the unpublished study was weakly reported in terms of its methods and implementation by CONSORT standards. Both of the included RCTs (including the Jordan RCT presented in this dissertation) measured change at four-month follow-ups. Although some studies of youth development programs have detected intervention effects on intermediate attitudes and behaviors within a similar timeframe (Grolnick et al., 2007; Lakin & Mahoney, 2006; Shelton, 2009b; Winkleby et al., 2004), the designs nevertheless prevented the studies' ability to detect any effects that might have materialized over a longer period of time.

Implementation limitations might also have affected results. Youth empowerment programs are complex interventions involving a high degree of social process and adaptability. The opportunities for diminished implementation fidelity in this context are many. Two of the included studies (Berg et al., 2009; Olson-Merichko, 2006) reported qualitative feedback from program participants, but systematic implementation fidelity assessments were lacking. As such, it is unclear as to the extent to which factors such as low attendance or low quality delivery of participatory methodologies might have existed and influenced results.

As the implementation study of the Jordan YEP discussed later showed, implementation and exposure can in practice fall below standards in some cases, and this could help explain null results. As such, it is vital that future impact studies of YEPs are accompanied by theory-driven implementation studies, preferably integrating quantifiable process measures that can be examined in statistical analyses with outcomes for interaction effects. Likewise, the literature on systematic reviews of complex community interventions increasingly emphasizes the importance of integrating implementation data with reviews of intervention effects (Durlak & DuPre, 2008; Roen et al., 2006); this should continue to be a priority for updates of this dissertation's systematic review as well as other reviews of youth development interventions.

Collecting and synthesizing implementation data could have a number of useful implications for the field. For example, if YEPs chronically struggle with low attendance, particularly in out-of-school programs where attendance may be more variable than in formal education-based programs, sample size calculations for future trials of YEPs should consider the impact that lower levels of attendance will have on statistical power and adjust the sample size targets accordingly.

Randomized controlled trial

Following from the systematic review, this pilot study involved the first completed randomized controlled trial (RCT) of a YEP in the Middle East and one of the few globally. In terms of evidence-based practice, the undertaking demonstrated

that rigorous impact study designs can be integrated into youth development efforts in the region in order to build a stronger evidence base for what works and provide practical data that can direct stakeholders toward areas for improvement.

Significant intervention effects were not detected by four-month posttest for ten out of eleven observed outcomes, including the primary outcome, general self-efficacy. Overall results also did not show evidence of adverse program effects on youths, though exploratory analysis indicated some possibility of harm in intervention sites where implementation quality was low. Results did show a medium-sized intervention effect on reducing and protecting against SDQ conduct problems over the course of four months. Previous research suggests that improved conduct can serve as an important proximal indicator for resilience in other aspects of adjustment into adolescence and early adulthood (Burke et al., 2007; Moffitt et al., 1996). Interestingly, the effect on conduct problems may reinforce a valuable role for strength-based intervention models in promoting resilience even before or without improving developmental assets.

Exploratory analysis provided informative insights into aspects of implementation and possible relationships with program effects. The program site with the highest youth ratings of empowerment outperformed the site with the lowest empowerment ratings on the majority of measures and appears to have been especially more effective in strengthening participants' prosocial behavior. The subgroup analysis was exploratory and insufficient to claim causality, but the finding might plausibly indicate that the degree of program empowerment

influences developmental outcomes. The more young people can contribute meaningfully to planning and implementation within a supportive environment, the more opportunity they might have to augment self-efficacy beliefs, connections with adults, and skills exercised through empowerment experiences.

Exploratory analysis indicated that the intervention had a stronger effect on younger youths (13 to 15) than older youths (16 to 21). More challenging empowerment experiences could be particularly important for older adolescents, for whom this study's analysis showed very little effect. The age range of participants, 13 to 21, captures more than one developmental period (Cohen et al., 1993). Previous longitudinal research has found that young people's desire to participate in decision-making processes increases as they transition into older adolescence (Midgley & Feldlaufer, 1987), implying that participatory opportunities may need to be more intensive for older age groups. Unlike QS NFE, some YEPs, like the two presented as examples in chapter one (RIATT, 2010; Wilson et al., 2007), include distinct leadership roles for older youth participants. This could be a promising strategy for building more developmentally appropriate levels of mastery experiences into the empowerment process.

Implementation study

As a multisite RCT of a complex, setting-level intervention in 'real world' context, the aforementioned impact study was particularly susceptible to breaches in implementation fidelity (Zvoch, 2009). Most of the QS NFE centers included in the research did appear to maintain a reasonably high fidelity to the program

methodology and components and with strong staff commitment at the point of service. However, in virtually all categories of implementation fidelity (adherence, quality of delivery, exposure, participant responsiveness, and program differentiation) at least one or two program centers performed relatively poorly, and all centers had some degree of difficulty with implementation fidelity.

This heterogeneity likely contributed to diluted intervention effects and has important implications for this and similar YEPs moving forward. Notably, chapter five reported the results of exploratory analysis for separated outcomes of the male QS NFE center with the highest implementation fidelity (as measured by youth feedback on the Learner Empowerment Survey) and attendance. Even with low statistical power, it was apparent that the findings were more encouraging for improvements in youths' developmental assets compared with the center in which youth perceived relatively low levels of empowerment.

Problems in implementation fidelity were particularly apparent with respect to exposure, as measured by attendance, with only 47.7% of the sample having met the attendance standard set by the program (two days per week average). The low attendance for many participants further diminished the trial's power to detect intervention effects if and where they exist. The attendance results could also reinforce broader implementation challenges for YEPs in challenging community contexts with marginalized youths.

Several youths and facilitators described practical barriers, such as family constraints, work demands, and lack of transportation assistance as impediments

for attendance. Some participants overcame these challenges, but youth and facilitator interviewees often expressed that a high-quality program environment (e.g., high levels of youth participation and adult support) and frequency of engaging activities (e.g., trips and recreational activities) were important draws to encourage out-of-school youths to overcome practical obstacles in order to participate regularly. Exploratory analysis on attendance and outcomes reported in chapter five suggested that positive peer interactions may also play an important role in promoting retention, though the direction of this association is unclear.

This implementation study was limited by lack of a validated and culturally appropriate measure (or collection of measures) to assess multiple dimensions of implementation fidelity according to the program's theory of change. Standardized instruments, largely using observational methods, have emerged in recent years to assess setting-level program processes with youth development interventions (Yohalem et al., 2009). Youth empowerment programs could benefit from adapting these toward the youth empowerment theory of change. Instruments such as the Youth Program Quality Assessment (PQA), which has shown generally favorable reliability and validity results in its early testing (Blazevski & Smith, 2007; Smith & Akiva, 2008), specifically address youth engagement, among other constructs relevant to the youth empowerment process, and therefore may be useful for assessing implementation of YEPs with relatively little adaptation outside of that needed to meet the respective cultural context of the study.

Process study

The qualitative process research aimed to explore organizational factors that promote or obstruct the diffusion of youth empowerment at the program-level. The implementation study sought to assess implementation fidelity to program activities and methodologies, but that does not offer explanations as to how or why a program is implemented in a certain way. The process study explored how and why the program was implemented with high or low fidelity, particularly with respect to the empowerment methodology, and the organizational lessons on diffusing empowerment that might be learned from this.

Two main process lessons surfaced from the sum of the qualitative data for diffusing youth empowerment in the context of organizational behavior. The first lesson was that empowerment (or disempowerment) functions and interacts at multiple levels. As Zimmerman has pointed out, “[a]lthough empowerment theory has consistently included multiple levels of analysis, the preponderance of research has been on psychological (individual) empowerment” (2000, p. 58). Findings from chapter six suggest that this limited examination of empowerment in programs constitutes a handicap for understanding empowerment theory.

For example, it appears difficult for staff to fully understand—much less fully implement—a process with which they themselves do not have direct experience. Much of the youth empowerment literature examines the extent to which adult staff empower youth, but very little of it examines the extent to which the adult staff feel empowered within the organizational system or broader social context and the

implications this might have for youth empowerment. While qualitative and exploratory, the process study data from QS NFE indicated that facilitators were more likely to implement the empowerment-based methodology for youth when provided themselves with empowering opportunities to develop professionally, receive open and transparent communications as partners in the program, and participate in program planning and decision-making.

The second lesson that arose from the process study was that the scaling-up process can place significant strain on youth empowerment. Qualitative interviews with various stakeholder groups indicated that the scaling-up process with QS NFE has had deleterious consequences for implementation fidelity. Having started in 2004 with a few community-based organizations, QS NFE has since expanded to include 57 sites (schools and community-based organizations) across Jordan and more than 7,000 participants with the partnership of the MOE and attainment of large project grants. Rapid expansion can strain implementation fidelity, and the nature of international aid resourcing such expansion can deepen the problem by exerting disproportionate pressure on outputs and top-down organizational processes at the expense of program emphases on impacts and the quality implementation processes believed to achieve them.

Considering results in the context of extant literature

This dissertation's title is *measuring the impacts of youth empowerment*. The primary task as such was to investigate the impacts of YEPs, specifically on adolescents' social and emotional developmental outcomes. Preceding sections

summarized and discussed the results of this dissertation's secondary and primary research; this section considers findings in the context of extant literature—including previous evidence and theoretical arguments.

Discussion of findings concerning impacts has to be explicit and realistic about the strength of any causality indicated in the evidence produced by this work or others. Bradford-Hill (1965) presented a prominent set of criteria used to determine the extent to which any case for a causal relationship could be justified in scientific research. In this dissertation, the foremost question is whether the experience of empowerment processes (via youths' supported participation in program decision-making and asset-building activities) causes improved social and emotional outcomes.

Bradford-Hill's nine criteria for distinguishing causation from association include: *strength* (i.e., effect size or clinical significance), *consistency* (i.e., repeatedly observed in different persons, places, and circumstances), *specificity* (i.e., the association detected is limited to the independent variable of concern and not something else), *temporality* (i.e., the 'cause' takes place before the 'effect'), *biological gradient* (i.e., dose-response curve), *plausibility* (i.e., the relationship is scientifically plausible—though findings can sometimes be both improbable and true), *coherence* (i.e., the cause-and-effect interpretation does not seriously conflict with generally known scientific facts), *experiment* (i.e., experimental or quasi-experimental studies show that introduction of a variable affects changes in another), and *analogy* (i.e., cause-and-effect between one set of variables can be

inferred from evidence that a theoretically similar set of variables have a similar causal relationship).

Bradford-Hill (1965) warned that none of his nine "viewpoints" present "hard-and-fast rules" for securing indisputable evidence for or against cause-and-effect relationships. He did, however, suggest that associations should be considered in light of each of the criterion when considering whether associations qualify as causation. To the contrary, much of the existing literature describing empowerment as causally related to wellbeing has been largely based on *plausibility*—theoretical arguments that persuasively describe participatory engagement of marginalized individuals and communities as a means to improving their life outcomes, substantiated by what is understood about human behavior through extant psychosocial studies and theories.

While plausibility is important, Bradford-Hill's criteria posit a more systematic framework through which claims of causality should be appraised. In particular, Bradford-Hill (1965) conveyed the *experiment* as the "strongest support for the causation hypothesis." This dissertation was designed to contribute to evidence in this criterion. In so doing, it also addressed several other criteria to varying degrees—e.g., effect sizes from the RCT and systematic review conveyed *strength*, the systematic review evaluated *consistency* between studies, subgroup analysis according to attendance and empowerment surveys explored possible *biological gradients*, and *specificity* and *temporality* were accounted for with the RCT design. As such, this study gives some empirical basis with which to examine the

strength (or lack thereof) of causal arguments concerning youth empowerment for changing developmental outcomes from multiple viewpoints—perhaps, more than any study of youth empowerment has enabled in the past.

Youth empowerment and self-efficacy

Unlike health interventions that aim to treat a particular disease and/or its symptoms, broad youth development programs like YEPs often target a range of attitudinal and behavioral changes rather than a single dominant outcome (Roth, 2004). As such, the concept of a “primary outcome” is less exacting in the context of a youth development program’s theory of change than it would be in many medical interventions’ theories of change. Nonetheless, development of self-efficacy is a frequent aim, among others, with empowerment theory and YEPs in particular (Holden et al., 2005; Jennings, 2006; Wong et al., 2010), and this dissertation held self-efficacy—general self-efficacy (GSE)—as the trial’s primary outcome due to its prevalence in the theoretical literature and relatedness to the goals of the intervention.

As reported, no intervention effects were detected for this outcome in the Jordan RCT. Consequently, the study failed to demonstrate causality between empowerment and self-efficacy by way of Bradford-Hill’s experimental criterion. The systematic review demonstrated that there is very little with which to compare the present RCT in terms of other experimental studies of YEPs. As such, the RCT acts as a contributor to a nascent evidence base rather than a test of the repeatability or transportability (or ‘consistency,’ using Bradford-Hill’s terminology)

of results from a previously efficacious intervention. Still, the two other included trials of YEPs showed similarly null results on self-efficacy (as well as self-esteem, the review's other primary outcome). These trials involved very small sample size or quasi-experimental design with significant selection bias vulnerabilities, but it is nevertheless interesting that none of the three studies showed statistically significant results on self-efficacy, an intermediate outcome that is frequently promoted in the youth empowerment literature. The null finding, in other words, is thus far consistent. This finding merits further examination of whether YEPs are appropriate for changing self-efficacy beliefs and whether current evaluation methods and measures are appropriate for detecting such changes.

Theoretical literature on self-efficacy suggests that individuals develop self-efficacy beliefs through mastery experience, vicarious experience, social persuasions, and physiological reactions (Pajares, 2006). Mastery experience is believed to be the most important source of self-efficacy beliefs but others play important roles as well (Pajares, 2006). Even though YEPs theoretically provide mastery experiences for youth through participation, in practice these experiences may not be as frequent as needed to change attitudes, especially about general self-efficacy beliefs. In YEPs, some youths may participate at very high levels—either by their own accord or by preferential opportunities facilitated by adults or other youth leaders—and this could limit the extent to which the majority of youths in programs can gain mastery experiences that might mediate self-efficacy development.

Further process evaluation of YEPs should assess the degree to which participatory opportunities are facilitated for all youths in the group; these evaluations should monitor program processes over time, rather than at one static point, in order to assess whether they are participating in progressively challenging mastery experiences over a duration of their involvement in the program. A better understanding of the type and frequency of mastery experiences facilitated by YEPs, and how these experiences interact with self-efficacy and other outcomes, would add a lot to theoretical development of how youth empowerment situates (or does not) in self-efficacy theory, and its related theories, social learning theory and social cognitive theory.

Furthermore, the lack of intervention effects on self-efficacy raises questions about measurement. For example, the findings renew the ongoing tension between those who promote GSE measures and those who prefer task specific self-efficacy (SSE) measures. Although no statistically significant intervention effects were either measured or detected among the included ($k = 3$) or excluded school-based YEP studies ($k = 3$) for GSE in the systematic review, the largest trial ($n = 813$; Winkleby et al., 2004), an excluded cluster RCT of a school-based tobacco prevention advocacy intervention, did show a significant positive intervention effect on self-efficacy beliefs specific to tobacco prevention advocacy.

On one hand, GSE measures can be more useful given that jobs and roles are becoming increasingly broad, complex, and demanding (Chen et al., 2001), and its global nature could be more amenable to YEPs' general theory of change, which is

broadly focused on adolescent development, rather than on any particular content task. General self-efficacy measures are also more easily combined in a meta-analysis to compare effects across studies than SSE measures that tend to be more heterogeneous. On the other hand, general self-efficacy has been described as a more stable and trait-like belief. Although this may reinforce its value as an outcome in youth empowerment, it also implies greater difficulty in changing GSE through a short-term intervention compared with SSE, which has been described as more malleable, assuming interventions target the particular task on which the SSE measure is centered (Chen et al., 2004).

The tension highlights strengths and trade-offs inherent in either type of measure. In the case of QS NFE, it is not clear that any particular SSE measure would have fit with the program's priorities or theory of change. Consequently, the local stakeholders on the research coordinating team expressed preference for the GSE outcome in the planning stage of the RCT. Nonetheless, especially given that GSE has been found to predict and influence SSE measures (Chen et al., 2001; Shelton, 1990; Sherer & Adams, 1983), where SSE measures are appropriate for the intervention, future impact studies of YEPs should follow changes in both GSE and relevant SSE outcomes and take care to analyze potential interactions between GSE and SSE over time if multiple follow-ups are feasible.

Revisiting theory

In chapter one, the dissertation presented a theoretical context in which youth empowerment's theory of change is situated. Most prominently,

empowerment theory influences the beliefs about change processes embedded in youth empowerment, but a number of other constituent psychosocial theories help to explain mechanisms by which empowerment might affect attitudinal and behavioral outcomes. This section reconsiders the applications of related theory for YEPs in light of the evidence presented by the components of this dissertation.

Empowerment theory's basic premise is that individuals' wellbeing (measured proximally in this dissertation by social and emotional outcomes) is connected to their engagement in the processes that affect their lives and surroundings (Zimmerman, 2000). As Holden and colleagues (2005) postulated with respect to youth, "Empowerment theory asserts that positive youth development emerges through promotion of skills development with greater participation and involvement by youth in the public affairs of their community." Youth empowerment programs facilitate these opportunities for participation within the program's decision-making, and sometimes more broadly in the community as well through youth-led service or advocacy projects.

In general, this dissertation's systematic review demonstrated that, to-date, there is insufficient evidence from experimental or quasi-experimental studies to confirm or reject the hypothesis that youth empowerment decisively produces positive adolescent outcomes. Bradford-Hill's (1965) experimental criterion for establishing causality remains largely unfulfilled. This dearth limits the dissertation's scientific capacity to confirm or deny empowerment theory's relevance for improving youth development outcomes. However, although the

Jordan RCT did not yield overall significant intervention effects on developmental outcomes, closer scrutiny of the results and empowerment theory may lend useful insights.

In the Jordan RCT, the intervention showed more positive results across various social and emotional outcomes from the center with the highest youth ratings of program empowerment compared with the center with the lowest empowerment ratings. Given low statistical power, findings from these subgroup analyses have to be treated with caution. Nonetheless, even with low power, the center with high empowerment showed an intervention effect on prosocial behavior that trended toward statistical significance ($p = .06$, $d = .65$).

Similarly, the Teen Outreach Program (TOP) that was evaluated by one of the school-based excluded studies in the systematic review showed a positive relationship between participants' perceptions of program empowerment (measured by autonomy—having input into the nature of the work performed) and their outcomes (Allen et al., 1994). The TOP study was limited in its theoretical contribution since only problem behaviors (teen pregnancy, academic failure, and suspension) were measured, and not strength-based outcomes more closely associated with empowerment's theory of change. Nonetheless, Allen and colleagues ventured that the finding “suggests that developing interventions for teenagers in which they can experience autonomy within a program is clearly *not* a contradiction in terms but may actually be a means of facilitating a critical developmental task” (Allen et al., 1994, p. 635).

The subgroup findings from QS NFE and TOP were limited and only exploratory, but they suggest an association between higher degrees of youth empowerment and positive outcomes. The association provides very modest evidence of a possible ‘biological gradient’ (more ‘dosage’ of empowerment renders more positive outcomes response) in Bradford-Hill’s causality criteria, but this and other criteria, such as *experimental* and *strength* are clearly unsatisfied for making any authoritative arguments about causality. Still, the finding that an association appears to exist between levels of empowerment and outcomes provokes the speculative explanation that the QS NFE RCT failed to produce overall intervention effects on developmental assets not because empowerment theory failed, but because the mechanisms of empowerment theory were not adequately implemented across all sites.

The one outcome for which the Jordan RCT did show statistically significant intervention effects was SDQ conduct problems. Similarly, four of the other five included or excluded school-based YEP studies in the systematic review also reported intervention effects on reducing various antisocial behavior outcomes (e.g., substance abuse, teenage pregnancy, and academic suspension). More research is needed, but this finding shows a modest level of *consistency* in the context of Bradford-Hill’s causality criteria. If future studies replicate this finding of YEPs showing intervention effects on problem behaviors without, or prior to, changing developmental outcomes, this could have implications for youth empowerment’s theory of change.

Firstly, *social control theory* could offer a framework for explaining impacts on problem behaviors. Social control theory posits that strengthening young people's bonds to positive individuals and institutions increases their resilience to deviant behavior (Hirschi, 1969). Empowerment theory suggests that programs that are based on participatory methodology are more likely to increase the meaningfulness of programs and young people's relationships with adults that act as facilitators and partners in the programs (Fymier et al., 1996; Wong et al., 2010). Considering the two in tandem, YEPs might be particularly effective in establishing positive bonds that protect youths from antisocial behavior even before other developmental outcomes materialize. This theory calls to question, then, why the Jordan RCT showed effects on conduct problems and not social supports as well. It could be that the social supports measures were not specific enough to relationships with adults and youths involved in the intervention, and therefore links between particular social bonds facilitated by YEPs and problem behaviors could have been missed. Notably, qualitative data did largely indicate strong bonds between participants and their peers and facilitators in the QS NFE program.

Secondly, the findings respond to a chief concern of harm reflected in the literature. Dishion and colleagues (2005; 1999; 2004; 2011) have demonstrated that well-intended programs serving vulnerable adolescents can have adverse impacts on problem behaviors by way of peer contagion effects that arise when programs aggregate deviant youth. Not only was there no evidence of such harm from the overall intervention effects of the YEP studies discussed in this dissertation's systematic review, but, in several cases, studies reported positive effects on

antisocial behaviors. In addition to highlighting social control theory, these results could reinforce YEPs' prosocial element of the theory of change. In other words, YEPs appear to manage to inculcate positive values and norms in the empowerment process that avoid or override the potentially harmful influences of peer contagion effects within group settings.

Thirdly, the finding that YEPs affect problem behavior outcomes without or prior to affecting strength-based outcomes could prompt reconsideration of how the variables interact. It may be that reducing problem behaviors is an important intermediate outcome of YEPs before strengths can be developed. This could alter the paradigm in positive youth development that programs strengthen developmental assets, which protect youths from negative behaviors (Kia-Keating et al., 2011). Future time series data of the impacts of YEPs could show that empowering programs facilitate processes that first alter antisocial behavioral tendencies among at-risk groups of youths, and this change makes subsequent development of assets possible. The notion is speculative, but further research is merited.

Implications for practice

Is youth empowerment investment-worthy?

In chapter one, the author presented counterarguments to youth empowerment programming, including economic interpretations by the likes of Heckman and colleagues (2006; 2007) that suggested that investments in interventions for vulnerable adolescents in general could be ill-advised in the face of

limited economic resources and more encouraging results from experimental studies of interventions aimed at early childhood. Although some modest promising findings emerged from this dissertation's systematic review and RCT, the evidence base for YEPs is clearly not developed enough to enable persuasive cost-benefit analyses and put these arguments at bay.

Still, the notion that there are diminishing returns to targeting investments in child development at later age groups should come as little surprise. The argument that investments should be made in early childhood development and *not* adolescent development, however, represents a more disputable position. Not all adolescents have had the advantage of early childhood intervention or adequate supports, and even those who have may still require intervention at the adolescent period to sustain and build on any positive gains made during early years (a point that Heckman (2006) endorsed with his "skill begets skill" concept). Youth empowerment programs may offer a promising strategy for fostering positive adolescent development and transitions to adulthood, but funders and policy-makers need more quantity and quality of scientific evidence to inform the levels and specifications of investments to be made in this area.

For many aligned with youth empowerment movements, the logic is obvious and irrefutable—so much so that it is worth advancing even without rigorous experimental evidence to support it. After reviewing evaluations of youth development programs broadly, Roth and colleagues (1998) concluded, "the current mismatch between the enthusiasm for these programmatic efforts and the empirical

evidence calls into question the effectiveness of such efforts.” Thirteen years later, this conclusion largely applies to YEPs. The hard question, “does youth empowerment work?” is virtually absent and arguably taboo in the topical literature. Rather, the majority of the literature offers case studies of variations of YEPs, discrepancies over the extent to which youth and adults should be involved, resource guides, and conceptual models. In other words, the literature generally assumes effectiveness despite evidence, and it proceeds from that assumption.

Much of youth empowerment support has worked to discredit traditional, problem-focused views of youths and corresponding ineffective strategies such as detention centers (Wooldredge, 1988), boot camps (Meade & Steiner, 2010), and scared straight programs (Petrosino et al., 2003) and present YEPs as the natural alternative. Yet, the fact that interventions with a different ideology have failed to produce positive youth outcomes does not demonstrate that YEPs will succeed. This premise has lulled the field of youth empowerment into hazardous lethargies.

In the absence of serious research evidence, youth empowerment is characterized by muddled and even contradictory guidance for practitioners and policy-makers. Trends among YEPs are largely driven by views from individual or institutional authorities in the field rather than by credible data. Despite broad rhetoric promoting youth empowerment around the globe, the lack of rigorous evidence for YEPs’ effectiveness presents an additional challenge for the translation of rhetoric to funding for YEPs in difficult economic times when scarcer resources may be more concentrated on proven interventions.

Although rigorous evidence is rarely sufficient to guarantee funding in a world of complicated influences and biases, or withdrawal of funding in the case of programs that are fiscally resilient despite ineffectiveness (Black & Donald, 2001), it does appear to play an increasingly important role. The UK government, for example, is dramatically reducing public spending on youth participation groups at the same time as focusing new spending on evidence-based interventions like the Family Nurse Partnership (FNP) program (Ashton, 2010; Cameron, 2010; Hillier, 2011). Across the Atlantic, the Obama Administration has directed The White House's Office of Management and Budget (OMB) to review the federal budget and make cuts and priorities based on rigorous evidence (Haskins & Baron, 2011; Orszag, 2009), and the Bill & Melinda Gates Foundation, the world's largest private philanthropic organization, is committed to making grants to projects based on results (Bill & Melinda Gates Foundation, 1999-2011). Program areas like youth empowerment that have grown accustomed to sustaining themselves in the absence of rigorous evaluation may find it increasingly difficult to do so in this kind of political and economic climate.

Addressing the empowerment-implementation problem

Although some evidence indicates that high levels of empowerment process may yield positive outcomes, the largely null findings among included studies of YEPs may highlight the difficulties in implementing empowerment sufficiently well to achieve such outcomes.

Gallagher and colleagues (2005) concluded from their process study of thirteen community-based youth development programs (many of which integrated empowerment-based methodologies) related to teenage pregnancy prevention in the USA that “the diffuse nature of their efforts, the global nature of their expected outcomes, and the diversity in their implementation” presented challenges for evaluating impact. Likewise, the implementation study reported in this dissertation demonstrated disparities in implementation of empowerment practices between centers under the *same* program, much less the implementation differences we could expect between different programs, as Gallagher and colleagues exhibited. Smith and Akiva (2008) also reported findings from a study that collected program quality data using a standardized assessment instrument with 1,635 youths across 164 programs nested within 51 organizations, which showed just as much variance in quality scores between staff *within* the same organization as there was quality variance across organizations.

The challenges to implementing high-quality youth empowerment could be better understood when empowerment is viewed as an *innovation* in youth services (Ozer et al., 2008). The difficulties to diffusing empowerment as an intervention are compounded by the nature of the innovation. Youth empowerment is highly process-oriented and culturally intertwined. Empowerment, by definition, is not formulaic in terms of the activities it generates nor does it lend itself easily to highly prescriptive manualization. Zimmerman argued that “[b]oth empowerment processes and outcomes vary in their outward form because no single standard can fully capture its meaning for all people in all contexts” (Zimmerman, 2000, p. 45).

Given that specificity has been shown to increase adherence in other intervention areas, this abstraction presents problems for empowerment interventions (Carroll et al., 2007).

Literature on diffusion of innovation has identified, among other things, *trialability* (the degree to which innovation appears easy to implement), *observability* (the degree to which an innovation can be identified, measured, and noticed), and *communicability* (the degree to which an innovation is easy to describe) as important dimensions of innovation qualities that influence whether an innovation is adopted, implemented, and routinized into social practice (Mayer & Davidson, 2000). A highly process-oriented complex social intervention like youth empowerment would take on these qualities much less than many medical or structured psychotherapeutic interventions would.

Even some complex interventions such as the Triple P parenting program or the Nurse-Family Partnership (NFP) program have lent reasonably well to fairly prescriptive manualization and implementation guidelines, which may have contributed to their ability to achieve and maintain demonstrated impact shown through rigorous experimental study and to scale-up to multiple locations with high implementation fidelity (e.g., the NFP program includes approximately 1,300 pages of instruction and facilitator manuals; Pursley, 2011). Considering the intrinsic difficulties of diffusing especially process-oriented, pliable, and complex interventions like YEPs, dissemination prospects of the intervention into high-quality, sustained implementation fidelity confront added adversity.

Given difficulties of defining and standardizing YEPs, should YEPs be exempted from standards of implementation fidelity? Should policy-makers and funders abandon YEPs for interventions that are more readily standardized and calculable? Zimmerman argued that the difficulty in standardizing empowerment “has led some to dismiss its usefulness, but that does not diminish its validity as a vital concept for the field” (2000, p. 57). Zimmerman’s point may be true, but there is also more opportunity to apply implementation fidelity standards to empowerment programs than is often projected. Chapter six discussed developments in the implementation fidelity literature for complex interventions, which suggest that key functions and processes of complex social interventions can be standardized and assessed even while particular activities and forms are left to local adaptation (Hawe et al., 2004).

This is important for YEPs because functions such as training of facilitators, skills training for youths to foster their preparedness for meaningful participation, and processes ensuring regular youth involvement in decision-making could be manualized and standardized to a greater extent than they are. These functions need not be overly prescriptive, but they could be documented and staff could be trained in them enough to establish clear standards for implementing stakeholders to work from and to facilitate consistency in the key processes embedded in YEPs’ theory of change: youth participation, asset-building activities, supportive relationships with adults, and a prosocial program environment.

Recent developments in instrumentation, such as the Youth Program Quality Assessment (PQA) (Blazevski & Smith, 2007), have emerged from the assumption that key *setting-level processes* within a youth program are important across various settings, can be standardized, and are measureable, even if the specific forms or activities through which the processes materialize vary from one setting to the next. The Youth PQA validation study showed high program quality scores, especially at the empowerment levels of program quality, to significantly predict positive self-reported youth outcomes, including safety, interest, sense of growth, and skill building (Smith & Akiva, 2008; Smith & Hohmann, 2005). The developments suggest that implementation fidelity is relevant and important for YEPs, and implementation resources are emerging to help programs diffuse and monitor participatory methodologies at the point of service.

Cross-cultural applications of youth empowerment

All of the studies of YEPs prior to this dissertation that were identified in the systematic review were from the United States. Moreover, the vast majority of the topical literature on youth empowerment, and, indeed, on child development in general, that is used on a global scale to inform research and practice has come from Western societies (Hart, 2008). Cultural psychologists have increasingly pointed to the precariousness of this position. Among them, Henrich and colleagues (2010b) emphasized the disproportionate research focus on what they call Western, educated, industrialized, rich, and democratic (WEIRD) societies.

A review by Henrich and colleagues (2010a) showed substantial cross-cultural variability in experimental results across a range of psychological and behavioral domains that alter fundamental assumptions underlying behavioral psychology and corresponding interventions. More unsettlingly, they concluded that research subjects from WEIRD societies, including children, “are among the least representative populations one could find for generalizing about humans.” Yet, a 2008 survey of leading psychology journals found that 96% of research participants were from Western industrialized countries, which comprise 12% of the world’s population (Arnett, 2008). The same paper reported that even within the American samples, 77% of them were majority European American ethnicity in a 2007 review of leading journals.

This background frames the importance of studying youth empowerment in different cultural contexts, as this dissertation has with primary research in a Middle East setting. As discussion of the results has already conveyed, the pilot RCT largely did not demonstrate that youth empowerment programming was effective in changing social or emotional outcomes for out-of-school Arab youth. The trial did, however, show intervention effects on conduct problems and subgroup analysis indicated more promising results for other outcomes such as prosocial behavior in centers where youths reported higher feelings of empowerment. Additionally, qualitative process data indicated that youth were responsive to participatory methodologies and wanted to see more of them. All of these points run contrary to the position that youth empowerment is not culturally applicable to Arab youth.

More experimental evaluation supported by culturally appropriate measures is warranted.

That said, chapter three's literature review also highlighted potential barriers to youth empowerment programming rooted in modern Arab society. These included hierarchical organizational systems, deficit-based views toward out-of-school youth, and disempowering tendencies in the educational system. Indeed, qualitative interviews reported in chapter six revealed evidence of each of these phenomena, and these presented additional challenges for diffusing empowerment-based methodologies into local practice. To this end, future research and efforts on disseminating youth empowerment in the region might benefit from considering culturally appropriate ways to target capacity building for staff and organizations responsible for implementing youth empowerment programming.

Implications for research

Youth empowerment and the evidence-based framework

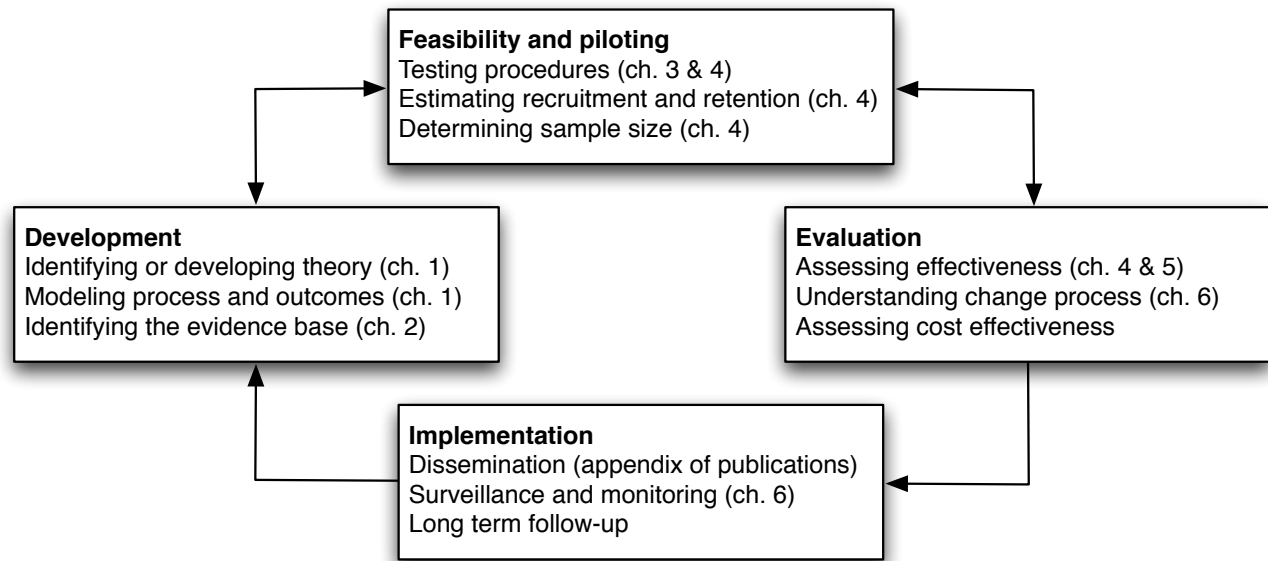
This dissertation set out to assess the impacts of youth empowerment. The result is the first project to have systematically investigated the impacts of YEPs guided by an evidence-based social intervention framework. A systematic, evidence-based approach to interventions like YEPs is best encapsulated by the MRC's Framework for Complex Interventions (MRC, 2008), which guided much of this research. While recognizing that the process of developing and evaluating complex interventions is rarely linear, the MRC Framework posits phases through which interventions should be studied insofar as is possible (see figure 7.1).

All four elements of the framework's development and evaluation process are represented in this dissertation:

Development – 'Identifying or developing theory' was a primary focus on the background development and the literature review it comprised (chapter 1), and 'identifying the evidence base' was achieved through the first systematic review on the effects of YEPs (chapter 2). The review revealed a meager evidence-base from experimental and quasi-experimental trials, indicating the need for further primary evaluation.

Feasibility and piloting – Because the YEP in Jordan was already operating prior to the study, a formal feasibility study to establish the early intervention design was not applicable. An in-depth review of the local context helped examine feasibility issues for implementing youth empowerment methodology in the Jordanian context and identify key assumptions in the theory of change (chapter 3). A collaborative research team and preliminary stakeholder discussions (chapter 4) also helped to assess and develop evaluation goals, indicators, and design. Outcome and process measures were translated, adapted, and field piloted before use in evaluation (chapter 4).

Figure 7.1. MRC Framework: Key elements of the development and evaluation process



Adapted from: Craig et al., 2008

Evaluation – A randomized controlled trial, the optimal method for impact evaluation recommended by the MRC Framework if appropriate, was conducted to study the effectiveness of a YEP implemented in ‘real world’ context (chapters 4 and 5). This trial also overlaps with the feasibility and piloting stage, as it was intended as a pilot evaluation to inform further program development and evaluation efforts, given its precedence as the first rigorous impact study of a youth development intervention in the Middle East. A mixed-methods ‘process evaluation’ (described in this dissertation as an implementation study) was conducted with the trial to assess fidelity and quality of implementation and how the intervention could be optimized (chapter 6).

Implementation and beyond – The qualitative process study investigated factors that support or hinder implementation of youth empowerment and revealed issues for further research and program development to this end (chapter 6). The

discussion of the process and implementation study also considered tools and strategies for assessing and maintaining high implementation fidelity as the program continues and potentially expands. The present chapter aims to summarize key lessons for further research and practice moving forward. Though not discussed in this dissertation, multiple efforts, including publications and stakeholder forums, have occurred or are planned in order to promote dissemination and knowledge transfer based on the research findings (see the list of publications and presentations on p. 6), which are intended to circle back to inform YEP development, as the MRC Framework intends.

Although guided by the MRC Framework, this dissertation was not exhaustive of it. Studies, for example, are warranted to better model and develop outcomes for youth empowerment at the development stage. At the evaluation phase, larger experimental evaluations with longer follow-up periods are needed to properly assess causal mechanisms, moderator and mediator variables, and potential time lags in intervention effects using appropriate design and statistical power to do so. At the implementation phase, monitoring and feedback systems and setting-level instruments designed to help YEPs assess and maintain high implementation quality over extended time periods and various cultural contexts are needed. Given these remaining gaps, more work clearly remains.

Incubation periods for research and development

The conditions of the interventions studied could make a difference in results. This dissertation included an impact evaluation of a YEP that was designed and implemented in ‘real world’ context by practitioners rather than researchers; this decision involved strengths and weaknesses. On one hand, the program was already operating, expanding, and gaining substantial interest among decision-makers. As such, there was an immediate practice value and applicability of findings produced from the research to inform policy and program development. Additionally, most YEPs operate in this way—organized by community institutions and implemented in dynamic, real world situations. Consequently, studying a YEP in its natural conditions provided insights into impacts and implementation issues that better represent the realities of the intervention, and these would not be reflected under highly researcher-controlled circumstances.

Efficacy trials were initially designed to test new medical drugs or procedures (hence the synonym, ‘clinical trials’), which, unlike YEPs, are often not yet available to and commonplace in the public at the efficacy testing phase. Furthermore, conducting impact evaluation with an existing YEP can be more feasible than designing and implementing a new intervention for the purpose of research—a substantial undertaking that can be economically and logistically prohibitive. Despite these benefits associated with studying the impacts of a ‘real world’ YEP, however, the decision also involved drawbacks.

As discussed in the previous section, quantitative and qualitative data suggested that problems in delivery may constitute a primary explanation for largely null findings from the RCT of QS NFE. A more tightly controlled YEP intervention designed and implemented by researchers might have ensured higher implementation fidelity to the components and processes expressed in youth empowerment's theory of change. Higher fidelity, in turn, enables more precise testing of the theoretical model of an intervention since implementation fissures can confound results. Consequently, some, like Flay and colleagues (2005), advocate a more regimented progression from 'efficacy' trials (impact study under researcher-controlled intervention conditions) to 'effectiveness' trials (impact study under real world intervention conditions).

This dissertation has aligned more with the updated MRC Framework on Complex Interventions, which recognizes that reality often stipulates a less linear approach to evidence-based social intervention. Nevertheless, the efficacy-to-effectiveness continuum logic has some merits for YEPs that could inform future research and practice. Youth empowerment programs may benefit from what the business sector refers to as a more intentional 'incubation period.' This would comprise a period of time (e.g., a couple of years) in which a YEP is carefully designed, manualized, and implemented (whether by practitioners and/or researchers) with acute focus on evaluation and the theory of change.

An incubation phase would concentrate resources on quality and adherence at both the staff development level and the point of service level, and it would

minimize the scale and distractions of the program to reinforce this focus. During this period, the YEP could be rigorously evaluated so as to contribute to better evidence regarding ‘active ingredients’ in and overall *efficacy* of youth empowerment (as much as an ‘efficacy’ trial is feasible with an intervention like YEPs). From a practice perspective, this period could also give the program leaders an opportunity to fortify and adjust the intervention model and develop stakeholder competencies for more successful scaling-up along the guidelines recommended by Bradach (2003) and discussed in chapter six.

Although much of the institutional rhetoric globally has promoted the generic dissemination of youth empowerment methodologies, this may in fact not be the most sensible position given the state of the evidence. Rather, donors and decision-makers ought to channel more support and attention toward a more modest and concerted strategy of better incubating YEPs with a lot of emphasis on manualizing, implementing, evaluating, monitoring, and refining YEPs for optimal impact. The field needs better models, supported by rigorous evidence, to justify and inform wider expansion of youth empowerment.

Measurement issues for empowerment as an outcome

Expanding on the opportunity for further work in modeling outcomes according to the MRC Framework’s development phase components, more consensus and development on theory-driven outcomes and measures for youth empowerment would aid future evaluation. A wide range of outcomes is frequently cited in association with youth empowerment, and yet these associations often lack

articulation of strong theory or evidence. Various YEPs may have outcomes that are unique to their particular goals. A YEP, for example, with a content focus on HIV/AIDS prevention might evaluate related health changes or knowledge. This, however, should not preclude content-specific YEPs from measuring outcomes that reflect a general youth empowerment theory of change and are relevant across YEPs due to common process commitments to youth participation, adult support, asset-building activities, and prosocial environments, regardless of their content foci.

Importantly, theorists have defined empowerment itself as both a process and an outcome (Zimmerman, 1995, 2000). This dissertation has primarily concentrated on empowerment as a process; its lack of conceptualization and measurement of empowerment as an outcome could be considered a limitation. Many of the outcomes included in this dissertation's systematic review and RCT, such as self-efficacy, self-esteem, and social skills, have been described as components of psychological empowerment (Zimmerman, 1995). Consequently, the utilized outcomes arguably had value as theoretically supported proxies of significant aspects if not the whole of psychological empowerment. Yet, future impact studies of YEPs might assess changes according to a clearly conceptualized and validated multi-component scale of outcome empowerment, as Holden and colleagues (2005) have proposed for tobacco advocacy YEPs. This pursuit, however, requires substantially more development and testing and therefor may not always be the most feasible approach for evaluators of YEPs when "proxy" outcomes are available.

Zimmerman's (1995) conceptualization of psychological empowerment includes three components each with their own subcomponents: intrapersonal (e.g., domain specific perceived control, domain-specific self-efficacy, perceived competence, perceived control), interactional (e.g., critical awareness, understanding of causal agents, skill development, resource mobilization), and behavioral (e.g., community involvement, organizational participation, and coping behaviors). Constructing an outcome measure consistent with Zimmerman's conceptualization of psychological empowerment, however, presents numerous practical and methodological difficulties.

Zimmerman suggested that a universal or global measure of psychological empowerment might not be suitable because the nature of empowerment is theoretically context dependent; consequently, a construct may be more useful if it is developed for specific situations or interventions. Zimmerman recognized that this approach necessarily limits the generalizability of psychological empowerment measures as a trade-off for securing internal validity. This recognition, however, does not alter the fact that systematic development and psychometric testing of measures are highly resource, time, and expertise intensive. Developing a new scale to capture the construct of psychological empowerment for every situation might exceed realistic expectations for youth empowerment research at a time when even experimental studies using existing measures of empowerment-related outcomes are rarely ventured.

Nevertheless, the youth empowerment field needs clearer, theoretically driven consensus on outcomes for impact study. Whether the corresponding measures are generalizable across YEPs or specific to particular YEPs, if even they sit within a common, consensual framework, they could support further impact evaluation and better comparison of impact evaluation results on YEPs. To this end, Zimmerman's framework could contribute to improved measurement of empowerment as an outcome to assess impacts of empowerment as an intervention, and this consistency between process and outcome could also strengthen future theory of change modeling for YEPs.

Component analysis

In addition to impact evaluation, further research is needed into how process factors influence outcomes and how organizational factors (internal and external) affect youth empowerment implementation. Process, key themes, and associations were examined using mixed methods in this dissertation's chapter six, but more authoritative methods are needed to test hypotheses. No experimental study of youth empowerment has yet been conducted with adequate power and design to include strong component analysis. This is considered a valuable aspect of assessing program differentiation in the context of implementation fidelity (Mihalic, 2009) and corresponds with the MRC Framework's evaluation stage.

If a YEP is shown to be effective, does its participatory component contribute to outcomes independently of other aspects of the program, such as a caring adult presence or the types of activities? Research by Smith and Akiva (2008), which

showed that high levels of participation and engagement are related to youth outcomes above and beyond social supports and activity types, lends some support to the hypothesis that empowerment has a mediating role, but experimental studies testing the effects of different components are lacking. Better attending to these questions is vital for assessing the validity of key notions espoused by theories like empowerment theory and self-determination theory in youth programming.

Dissemination research

In terms of research into the organizational factors influencing empowerment implementation, both deductive and inductive methods could play important roles for further investigation. Deductively, for example, research could quantitatively test the hypothesis that youth and adult empowerment are connected in organizational systems, and it could test relationships with other variables of interest. This might be ventured by surveys conducted at the program level of analysis across multiple program sites over multiple time points. Subsequent analysis could examine interactions between youth and adult scores on strategic process variables, including sense of empowerment, for potential moderators or mediators. Analysis using data from subjective measures could be supplemented by using objective measures (e.g., standardized, third-party observational ratings) of organizational and/or implementation processes.

Inductively, more thorough qualitative and case study investigation into processes and their circumstances could help develop understanding of factors promoting or hindering sustained youth empowerment. Zimmerman suggested that

“[i]n-depth study of the research setting and population would not only add to our understanding of [empowerment], but would also add insight into the organizational and community settings in which it develops” (2000, p. 51). An inductive mixed-methods approach would help to identify hypotheses and aspects of the phenomena that might not be currently understood or considered. The small-scale qualitative process study included in this dissertation added some useful insights and contributed to hypothesis building. However, it simply could not provide the detailed and intimate level of understanding into the complex social and cultural processes through which implementation of youth empowerment pass that would be possible in a more long-term, ethnographic investigation in which the researcher is committed to a closer observational co-existence with the research participants.

Concluding remarks

This closing begins with a brief illustration from another context of social research: neighborhood effects. Particularly since Wilson (1987) authored *The Truly Disadvantaged*, the social science community has taken keen interest in the role of neighborhoods as potential influents of child and youth outcomes. Yet, despite burgeoning research into the topic, the precise ways in which neighborhood contexts did or did not contribute to individuals’ wellbeing remained elusive. A multitude of observational studies helped to generate hypotheses and associations, but these methods could not overcome the problem of selection bias leaving studies persistently incapable of deciphering causalities.

Despite the difficulties involved, researchers advocated the need for experimental designs to overcome selection bias problems and make more authoritative inferences about the effects of neighborhood conditions on individuals' wellbeing (Sampson & Morenoff, 2000). Eventually, one such experiment did take place: Moving to Opportunity (MTO), a sophisticated, multi-city housing relocation randomized controlled trial in the United States (Leventhal & Brooks-Gunn, 2003b, 2004). Yet, although an enormously important contribution to the field, the MTO trial prompted numerous critiques of its methodology and the extent to which its design was informed by sound ecological theory, and its mixed results left social scientists still speculating about the role of neighborhood effects (Sampson, 2008; Sobel, 2006).

In light of these critiques and mixed results, Sampson, in defending the overall value of the project, ventured a more sobering view of the study's contribution: "MTO has been important in debunking simple-minded hypotheses: no simple conclusion can or should be drawn about neighborhood effects in the abstract" (Sampson, 2008, pp. 224-225). Indeed, the early applications of rigorous evaluative methods to the challenging topic of youth empowerment and the result of, at best, mixed effects have yielded much the same reality.

As the task demands, theory, methods, and data abound in this dissertation. Perhaps, the most important subtext behind it all is simply that significantly and quantifiably changing the lives of vulnerable adolescents is not easy. Adolescents, and especially marginalized and older adolescents, are harder to attract to and retain in programs (Deschenes et al., 2010), harder to change outcomes with (Weisz

et al., 1987), and, some would argue, harder to justify public investments in (Heckman & Masterov, 2007). In fact, youth empowerment has largely emerged as a programmatic field because many previous interventions and strategies failed to change youth outcomes; many even did more harm than good. In response, YEPs not only offer an alternative intervention, but they embody a radically different paradigm about how best to change youth outcomes, as well as what kinds of outcomes should be changed.

In the early-1960s, Friedenbergr (1962) observed that “the teenager seems to have replaced the Communist as the appropriate target for public controversy and foreboding.” The statement captures an enduring sentiment that society largely conceptualizes adolescents as problems, and policies and programs react with deficit-based orientations. In contrast, youth empowerment is rooted in a strengths-based orientation to adolescents that views them as capable of sharing power in the processes that affect their lives and surroundings and more likely to experience positive development when given that experience. Unfortunately, the evidence is still too meager to scientifically substantiate this expectation, but modest developments suggest it may have some merit. This study is only a beginning, but it has endeavored to contribute to the theoretical, methodological, and evidence basis from which the youth empowerment field can continue to advance through a more evidence-based course of research and action.

As this dissertation highlights, the task of measurably impacting the lives of vulnerable adolescents is a great challenge, both in terms of designing and

implementing programs to achieve impacts and in terms of designing and implementing evaluations to measure them. Nonetheless, the need is significant with the problems many adolescents face. Youth empowerment may offer opportunities for young people to not only avoid or overcome these problems, but to concurrently contribute to advancement of selves, programs, communities, and societies. To this end, the task is well justified, however arduous.

References

- Aboudan, F., & Murray, H. (2009). *Scoping Study of Education Contexts in the Middle East: Egypt, Israel, Jordan, Lebanon, Palestine and Syria*. London: Open Society Institute.
- Abu-Saad, I., & Hendrix, V. L. (1995). Organizational climate and teachers' job satisfaction in a multi-cultural milieu: The case of the Bedouin Arab schools in Israel. *Int. J. Educational Development*, 15(2), 141-153.
- African Union. (2006). *African Youth Charter*. Retrieved from http://www.africa-union.org/root/UA/Conferences/Mai/HRST/Charter_english.pdf.
- Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect Size and Power in Assessing Moderating Effects of Categorical Variables Using Multiple Regression: A 30-Year Review. *Journal of Applied Psychology*, 90(1), 94-107.
- Ahmad, M., & Coombs, P. H. (Eds.). (1975). *Education for rural development: Case studies for planners*. New York: Praeger Publications.
- Akers, R. L., Krohn, M., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behavior: A specific test of a general theory. *American Sociological Review*, 44, 636-655.
- Al-Faleh, M. (1993). Cultural Influences on Arab Management Development: A Case Study of Jordan. *Journal of Management Development*, 6(3), 19-33.
- Al-Khalidi, M. D. (2006). Educational Attainment of Jordan's Population and Labor Force. *Journal of Social Sciences*, 2(1), 1-6.
- Al-Khouli, T., Al-Zo'bi, H., Mihyar, H., & Rhodes, C. (2006). Assessing the Impact of the Mentoring Program on Increasing At-Risk Children and Youth's Efficiency in Integrating into Society. *Questscope for Social Development in the Middle East*.
- Al-Kubaisy, A. (1985). Namothoj Al-tinmeha Al-edaiya Fi Dual Al-khalej Al-arabi. *The Arab Gulf*, 17(2), 29-48.
- Al-Sa'd, A. (2007). *Evaluation of students' attitudes toward vocational education in Jordan*. Malmö University, Malmö, Sweden.
- Ali, A. J. (1990). Management Theory in a Transitional Society: the Arab's Experience. [Article]. *International Studies of Management & Organization*, 20(3), 7-35.
- Allen, J. P., Kuperminc, G., Philliber, S., & Herre, K. (1994). Programmatic prevention of adolescent problem behaviors: The role of autonomy, relatedness, and volunteer service in the teen outreach program. *American Journal of Community Psychology*, 22(5), 617-638.
- Allen, J. P., Philliber, S., Herrling, S., & Kuperminc, G. P. (1997). Preventing teen Pregnancy and Academic Failure: Experimental Evaluation of a Developmentally Based Approach. *Child Development*, 68(4), 729-742.
- Allport, G. W. (1954). *The nature of prejudice*. New York: Doubleday.
- Altman, D., Feighery, E., Robinson, T., Haydel, K., Strausberg, L., Lorig, K., Killen, J. (1998). Psychosocial factors associated with youth involvement in

- community activities promoting heart health. *Health Education & Behavior*, 25.
- Altman, D. G. (2009). Missing outcomes in randomized trials: addressing the dilemma. *Open Medicine*, 3(2), 31-50.
- Altman, D. G., & Feighery, E. C. (2004). Future Directions for Youth Empowerment: Commentary on Application of Youth Empowerment Theory to Tobacco Control. *Health Educ Behav*, 31, 641-647.
- Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, 332(7549), 1080.
- Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., et al. (2001). The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration. *Ann Intern Med*, 134(8), 663-694.
- Ambasht, N. K. (1992). Non-formal Education. In V. S. Upadhyay & V. P. Sharma (Eds.), *Contemporary Indian Society* (pp. 58-72). Anmol Publications: New Delhi.
- Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children; prevalence in a large sample from the general population. *Archives of General Psychiatry*, 44, 69-76.
- Anderson, K. S., & Sandmann, L. (2009). Toward a Model of Empowering Practices in Youth-Adult Partnerships. *Journal of Extension*, 47(2).
- Anderson, R. (2008). New MRC guidance on evaluating complex interventions. *BMJ*, 337.
- Aptekar, L. (1994). Street Children in the Developing World: A Review of Their Condition. *Cross-Cultural Research*, 28, 195.
- Arnett, J. J. (1999). Adolescent storm and stress, reconsidered. *American Psychologist*, 54(5), 317-326.
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469-480.
- Arnett, J. J. (2006). Emerging Adulthood: Understanding the New Way of Coming of Age. In J. J. Arnett & J. L. Tanner (Eds.), *Emerging adults in America: Coming of age in the 21st century*. (pp. 3-19): Washington, DC, US: American Psychological Association.
- Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, 63(7), 602-614.
- Arnold, H. J., & Feldman, D. C. (1981). Social Desirability Response Bias in Self-Report Choice Situations. *Academy of Management Journal*, 24(2), 377-385.
- Arnold, M. E., & Hughes, J. N. (1999). First Do No Harm: Adverse Effects of Grouping Deviant Youth for Skills Training. *Journal of School Psychology*, 37(1), 99-115.
- Aseltine, R. H. (1995). A reconsideration of parental and peer influences on adolescent deviance. *Journal of Health and Social Behavior*, 36(2), 103-121.
- Ashton, C. (2010). Family nurses offer young mothers a bright future. Retrieved May 30, 2011, from <http://www.bbc.co.uk/news/education-11062614>
- Aspy, C., Oman, R., Vesely, S., McLeroy, K., Rodine, S., & Marshall, L. (2004). Adolescent Violence: The Protective Effects of Youth Assets. *Journal of Counseling & Development*, 82(3).

- Assaad, R., & Barsoum, G. (2009). Rising Expectations and Diminishing Opportunities for Egypt's Young. In N. Dhillon & T. Yousef (Eds.), *Generation in Waiting: The Unfulfilled Promise of Young People in the Middle East* (pp. 67-94). Washington, D.C.: Brookings Institution Press.
- ASSDA'A Burson-Marsteller. (2011). *3rd Annual ASDA'A Burson-Marsteller Arab Youth Survey*. Dubai: ASSDA'A Burson-Marsteller.
- Azer, A. (1993). *Situation Analysis of Disadvantaged Children in Egypt*. Cairo: UNICEF.
- Baldwin, A. L., Baldwin, C., & Cole, R. E. (1990). Stress-resistant families and stress-resistant children. In J. Rolf, A. S. Masten, D. Chicchetti, K. Nuechterlein & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 257-280). Cambridge: Cambridge University Press.
- Baldwin, S. A., & Hoffman, J. P. (2002). The Dynamics of Self-Esteem: A Growth-Curve Analysis. *Journal of Youth and Adolescence*, 31(2), 101-113.
- Bandura, A. (1986). *Social Foundations of Thought and Action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, 28(2), 117-148.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (2006a). Adolescent Development from an Agentic Perspective. In F. Pajares & T. Urdan (Eds.), *Adolescence and Education: Self-Efficacy Beliefs of Adolescents* (Vol. V, pp. 1-43). Greenwich: Information Age Publishing.
- Bandura, A. (2006b). Guide for Constructing Self-Efficacy Scales. In F. Pajares & T. Urdan (Eds.), *Adolescence and Education: Self-Efficacy Beliefs of Adolescents* (Vol. V, pp. 307-337). Greenwich: Information Age Publishing.
- Bandura, A., Pastorellia, C., Barbaranelli, C., and Caprara, G. . (1999). Self-efficacy pathways to childhood depression. *Journal of Personality and Social Psychology*, 76(2).
- Barber, T. (2007). Who is Youth Work For? Distortions and Possibilities. *Scottish Youth Issues Journal*(9), 77-88.
- Bardone, A. M., Moffitt, T. E., Caspi, A., Dickson, N., & Silva, P. A. (1996). Adult mental health and social outcomes of adolescent girls with depression and conduct disorder. *Development and Psychopathology*, 8(04), 811-829
- Barnes, G. M., & Farrell, M. P. (1992). Parental Support and Control as Predictors of Adolescent Drinking, Delinquency, and Related Problem Behaviors. *Journal of Marriage and Family*, 54(4), 763-776.
- Barrera, M., & Garrison-Jones, C. (1992). Family and peer social support as specific correlates of adolescent depressive symptoms. *Journal of Abnormal Child Psychology*, 20(1), 1-16.
- BASW. (2011). Code of Ethics. Retrieved June 30, 2011, from <http://www.basw.co.uk/about/code-of-ethics/>
- Bateson, M. C. (2003). *Arabic Language Handbook*. Washington, D.C.: Georgetown University Press.
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, 103(1), 5-33.

- Baumrind, D. (1972). An Exploratory Study of Socialization Effects on Black Children: Some Black-White Comparisons. *Child Development*, 43(1), 261-267.
- Baumrind, D. (1991). The Influence of Parenting Style on Adolescent Competence and Substance Use. *The Journal of Early Adolescence*, 11(1), 56-95.
- BBC. (2011). Clashes break out at Jordan anti-government protest. Retrieved May 2, 2011, from <http://www.bbc.co.uk/news/world-middle-east-12857360>
- Berg, M., Coman, E., & Schensul, J. (2009). Youth Action Research for Prevention: A Multi-level Intervention Designed to Increase Efficacy and Empowerment Among Urban Youth. *American Journal of Community Psychology*, 43(3), 345-359.
- Berndt, T. J., & Keefe, K. (1995). Friends' Influence on Adolescents' Adjustment to School. *Child Development*, 66(5), 1312-1329.
- Bibars, I. (1998). Street children in Egypt: from the home to the street to inappropriate corrective institutions *Environment and Urbanization*, 10, 201-216.
- Biddle, B. J. (1986). Recent Development in Role Theory. *Annual Review of Sociology*, 12, 67-92.
- Bill & Melinda Gates Foundation. (1999-2011). Guiding Principles. Retrieved May 30, 2011, from <http://www.gatesfoundation.org/about/Pages/guiding-principles.aspx>
- Billig, S., Rot, S., & Jesse, D. (2005). The relationship between the quality indicators of service-learning and student outcomes: Testing professional wisdom. *Improving service-learning practice: Research on models to enhance impacts* (pp. 97-115). Greenwich, CT: Information Age Publishers.
- Birman, B. F., Desimone, L., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership*, 57(8), 1-8.
- Bishop, J., Bishop, M., Bishop, M., Gelbwasser, L., Green, S., Peterson, E., Rubinsztaj, and Zuckerman, A. (2004). Why we harass nerds and freaks: a formal theory of student culture and norms. *Journal of School Health*, 74(7).
- Black, N., & Donald, A. (2001). / Evidence based policy: proceed with care / Commentary: research must be taken seriously. *BMJ*, 323(7307), 275-279.
- Blanchard, K., Carlos, J., & Randolph, A. (1996). *Empowerment takes more than a minute*. San Francisco: Berret-Koehler.
- Blazevski, J., & Smith, C. (2007). *High/Scope Youth PQA Technical Report: Inter-rater Reliability on the Youth Program Quality Assessment*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Block, P. (1987). *The empowered manager: Positive political skills at work*. Sa Francisco: Joessey-Bass.
- Blum, R. (2003). Positive youth development: A strategy for improving adolescent health. In F. Jacobs, D. Wertlieb & R. M. Lerner (Eds.), *Enhancing the Life Chances of Youth and Families: Contributions of Programs, Policies, and Service Sytems* (Vol. 2). Thousand Oaks, CA: Sage Publications, Inc.
- Blum, R., Reinehart, P. (1997). *Reducing the risk: Connections that make a difference in the lives of youth*. Rockville: National Institute on Alcohol Abuse and Alcoholism.

- Blum, R. W., & Ireland, M. (2004). Reducing risk, increasing protective factors: Findings from the Caribbean Youth Health Survey. *Journal of Adolescent Health, 35*(6), 493-500.
- Boden, J. M., Fergusson, D. M., & Horwood, L. J. (2008). Does adolescent self-esteem predict later life outcomes? A test of the causal role of self-esteem. *Development and Psychopathology, 20*(01), 319-339.
- Boeck, T. (2009). Social Capital and Young People. In J. Wood & J. Hine (Eds.), *Work with Young People: Theory and Policy for Practice* (pp. 88-103). Thousand Oaks, CA: SAGE.
- Bogenschneider, K. (1996). An Ecological Risk/Protective Theory for Building Prevention Programs, Policies, and Community Capacity to Support Youth. *Family Relations, 45*(2), 127-138.
- Bolognini, M., Plancherel, B., Bettschart, W., & Halfon, O. (1996). Self-esteem and mental health in early adolescence: development and gender differences. *Journal of Adolescence, 19*(3), 233-245.
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V., & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ, 333*(7563), 346-349.
- Booth, M. (2002). Arab Adolescents Facing the Future: Enduring Ideals and Pressures for Change. In B. Brown, R. Larson & T. S. Saraswati (Eds.), *The World's Youth: Adolescence in Eight Regions of the Globe* (pp. 207-242). Cambridge: Cambridge University Press.
- Borenstein, M., Hedges, L. V., & Rothstein, H. (2007). *Meta-Analysis: Fixed effect vs. random effects*: Meta-analysis.com.
- Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., Ravaud, P., & for the, C. G. (2008). Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. *Annals of Internal Medicine, 148*(4), 295-309.
- Box, G., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B, 26*(2), 211-252.
- Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review, 26*(3), 291-345.
- Bradach, J. (2003). Going to Scale: The challenge of replicating social programs. *Stanford Social Innovation Review*(Spring).
- Bradford-Hill, A. (1965). The environment and disease: Association or causation? . *Proceedings of the Royal Society of Medicine, 58*, 295-300.
- Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*, 164-173.
- Brieger, W. R., Delano, G. E., Lane, C. G., Oladepo, O., & Oyediran, K. A. (2001). West African youth initiative: outcome of a reproductive health education program. *Journal of Adolescent Health, 29*(6), 436-446.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.
- Brooks-Gunn, J., Duncan, G., and Aber, L. (Ed.). (1997). *Neighborhood Poverty* (Vol. I). New York, New York: Russell Sage Foundation.

- Brown, L. (2009). *HEY Guide: Youth Empowerment Model A Step-By-Step Guide to Empower Youth Leaders*. San Francisco, CA: Honoring Emancipated Youth (HEY).
- Bryant, A., & Charmaz, K. (2007). Grounded Theory Research: Methods and Practices. In A. Bryant & K. Charmaz (Eds.), *The SAGE Handbook of Grounded Theory* (pp. 1). London: SAGE.
- Burke, J., Loeber, R., & Lahey, B. (2007). Adolescent conduct disorder and interpersonal callousness as predictors of psychopathy in young adults. *Journal of Clinical Child Psychology*, 36(3), 334 - 346.
- Burns, A. A., Ruland, C. D., & Finger, W. (2004). *Reaching out-of-school youth with reproductive health and HIV/AIDS information services*. Arlington, VA: Family Health International.
- Call, K. T., Riedel, A. A., Hein, K., McLoyd, V., Petersen, A., & Kipke, M. (2002). Adolescent health and well-being in the twenty-first century: A global perspective. *Journal of Research on Adolescence*, 12(1), 69-98.
- Cameron, D. (2010). Speech: David Cameron: Supporting parents. Retrieved May 30, 2011, from http://www.conservatives.com/News/Speeches/2010/01/David_Cameron_Supporting_parents.aspx
- Camino, L. A. (2000). Youth-Adult Partnerships: Entering New Territory in Community Work and Research. *Applied Developmental Science*, 4, 11-20.
- Capaldi, D. M., & Stoolmiller, M. (1999). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: III. Prediction to young-adult adjustment. *Development and Psychopathology*, 11(01), 59-84
- Cargo, M., Grams, G., Ottoson, J., Ward, P., & Green, L. (2003). Empowerment as fostering Positive Youth Development and citizenship. *American Journal of Health Behavior*, 27(1), S66-S79.
- Carnegie. (2008). *CYPI Final Report*. London: The Carnegie UK Trust.
- Carneiro, P., Crawford, C., & Goodman, A. (2007). *The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes*. London: Centre for the Economics of Education, London School of Economics.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation science* : IS, 2, 40.
- Castelloe, P., & Watson, T. (1999). Participatory Education as a Community Practice Method. *Journal of Community Practice*, 6(1), 71-89.
- Catalano, R., Berglund, M., Ryan, J., Lonczak, H., & Hawkins, J. (2004). Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Program. *The ANNALS of the American Academy of Political and Social Sciences*, 591(1), 98-124.
- Cauffman, E. (2004). The Adolescent Brain: Excuse versus Explanation: Comments on Part IV. *Annals of the New York Academy of Sciences*, 1021(1), 160-161.
- Chaaban, J. (2008). *The Costs of Youth Exclusion in the Middle East*. Washington, DC: Brookings Institution.
- Chaaban, J. (2009). Youth and Development in the Arab Countries: The Need for a Different Approach. *Middle Eastern Studies*, 45(1), 33 - 55.

- Chambers, R. (1992). *Rural appraisal: rapid, relaxed and participatory*. Brighton, UK: Institute for Development Studies, University of Sussex.
- Chambers, R. (1994). The origins and practice of participatory rural appraisal. *World Development*, 22(7), 953-969.
- Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4(1), 62-83.
- Chen, G., Gully, S. M., & Eden, D. (2004). General self-efficacy and self-esteem: toward theoretical and empirical distinction between correlated self-evaluations. *Journal of Organizational Behavior*, 25(3), 375-395.
- Chinman, M., & Linnery, J. (1998). Toward a model of adolescent empowerment: Theoretical and empirical evidence. *Journal of Primary Prevention*, 18, 393-413.
- Church, A. T. (2010). Measurement Issues in Cross-cultural Research. In G. Walford, E. Tucker & M. Viswanathan (Eds.), *The Sage Handbook of Measurement* (pp. 152-175). London: SAGE.
- Cloward, R. A., & Ohlin, L. (1960). *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press.
- Cochran, W. G., & Cox, G. M. (1957). *Experimental Designs* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7(3), 249-253.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cohen, J. M. (1977). Sources of Peer Group Homogeneity. *Sociology of Education*, 50(4), 227-241.
- Cohen, P., Cohen, J., Kasen, S., Velez, C. N., Hartmark, C., Johnson, J., et al. (1993). An Epidemiological Study of Disorders in Late Childhood and Adolescence—I. Age- and Gender-Specific Prevalence. *Journal of Child Psychology and Psychiatry*, 34(6), 851-867.
- Connell, J., & Halpern-Felsher, B. (1997). How neighborhoods affect educational outcomes in middle childhood and adolescence: conceptual issues and an empirical example. In J. Brooks-Gunn, Duncan, G., and Aber, L. (Ed.), *Neighborhood Poverty* (Vol. I). New York: Russell Sage Foundation.
- Connell, J. P., Aber, J. L., & Walker, G. (1995). How do urban communities affect youth? Using social science research to inform the design and evaluation of comprehensive community initiatives. In J. P. Connell, A. Kubisch, L. B. Schorr & C. Weiss (Eds.), *New Approaches To Evaluating Community Initiatives: Concepts, Methods, And Contexts* (Vol. 1, pp. 93). Washington, D.C.: Aspen Institute.
- Connell, J. P., & Kubisch, A. C. (1998). *Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives: Progress, Prospects, and Problems*: The Aspen Institute.
- Cook, D., Gebiski, V. J., & Keech, A. C. (2004). Subgroup analysis in clinical trials. *The Medical Journal of Australia*, 180(6), 289-291.
- Cooke, B., & Kothari, U. (2001). *Participation: the New Tyranny?* London: Zed Books.
- Cortina, J. M. (1993). What is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 78(1), 98-104.

- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, 337.
- Crowley, A., & Skeels, A. (2010). Getting the measure of children and young people's participation. In B. Percy-Smith & N. Thomas (Eds.), *A Handbook of Children and Young People's Participation: Perspectives from theory and practice* (pp. 184-192). Abingdon, UK: Routledge.
- Cunha, F., Heckman, J., Lochner, L., & Masterov, D. (2005). *Interpreting the evidence on life cycle skill formation* (No. 1675).
- Cunha, F., & Heckman, J. J. (2006). *Investing in our Young People* (Working Paper). Chicago: University of Chicago.
- Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3), 883-931.
- D'Aogostino, R. B. (1986). Tests for the Normal Distribution. In R. B. D'Aogostino & M. A. Stephens (Eds.), *Goodness-of-Fit Techniques* (pp. 36). New York: Marcel Dekker, Inc.
- Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. keynote address. *Annals of the New York Academy of Sciences*, 1021(1), 1-22.
- Dane, A. V., & Schneider, B. H. (1998). Program Integrity In Primary and Early Secondary Prevention: Are Implementation Effects Out of Control? *Clinical Psychology Review*, 18(1), 23-45.
- de Boer, H., Bosker, R. J., & van der Werf, M. P. C. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology*, 102(1), 168-179.
- de Vries, H., Dijkstra, M., & Kuhlman, P. (1988). Self-efficacy: the third factor besides attitude and subjective norm as a predictor of behavioural intentions. *Health Educ. Res.*, 3(3), 273-282.
- Deschenes, S. N., Arbreton, A., Little, P. M., Herrera, C., Baldwin Grossman, J., Weiss, H. B., et al. (2010). *Engaging older youth: Program and city-level strategies to support sustained participation in out-of-school time*. Cambridge, MA: Harvard Family Research Project and Public/Private Ventures.
- Dhillon, N. (2008, January 7). *From Oil Boom to Youth Boon: Tapping the Middle East Demographic Gift*, Washington, D.C.
- Dickerson, F. B. (1998). Strategies that foster empowerment. *Cognitive and Behavioral Practice*, 5(2), 255-275.
- Dishion, T. J. (2005). Peer contagion in interventions for children and adolescents: Moving toward an understanding of the ecology and dynamics of change. *Journal of Abnormal Child Psychology*, 33(3).
- Dishion, T. J., McCord, J., & Poulin, F. o. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, 54(9), 755-764.
- Dishion, T. J., Nelson, S. E., & Bullock, B. M. (2004). Premature adolescent autonomy: parent disengagement and deviant peer process in the amplification of problem behaviour. *Journal of Adolescence*, 27(5), 515-530.
- Dishion, T. J., & Tipsord, J. M. (2011). Peer Contagion in Child and Adolescent Social and Emotional Development. *Annual Review of Psychology*, 62, 189-214.

- Dixon-Woods, M., Bonas, S., Booth, A., Jones, D. R., Miller, T., Sutton, A. J., et al. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, 6(1), 27-44.
- Donner, A., & Klar, N. (2002). Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine*, 21, 2971-2980.
- Dopson, S. (2006). Debate: Why Does Knowledge Stick? What We Can Learn From the Case of Evidence-Based Health Care. *Public Money & Management*, 26(2), 85-86.
- DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of Mentoring Programs for Youth: A Meta-Analytic Review. *American Journal of Community Psychology*, 30(2), 157-197.
- Dubow, E. F., & Tisak, J. (1989). The Relation between Stressful Life Events and Adjustment in Elementary School Children: The Role of Social Support and Social Problem-Solving Skills. *Child Development*, 60(6), 1412-1423.
- Dubow, E. F., & Ullman, D. G. (1989). Assessing Social Support in Elementary School Children: The Survey of Children's Social Support. *Journal of Clinical Child Psychology*, 18(1), 52 - 64.
- Duke, D. L. (1978). The Etiology of Student Misbehavior and the Depersonalization of Blame. *Review of Educational Research*, 48(3), 415-437.
- Dumont, M., & Provost, M. A. (1999). Resilience in Adolescents: Protective Role of Social Support, Coping Strategies, Self-Esteem, and Social Activities on Experience of Stress and Depression. *Journal of Youth and Adolescence*, 28(3), 343-363.
- Durlak, J., & DuPre, E. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology*, 41(3), 327-350.
- Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning (CASEL).
- Duyan, V. (2005). Relationships between the Sociodemographic and Family Characteristics, Street Life Experiences and the Hopelessness of Street Children. *Childhood*, 12(4), 445-459.
- Dwairy, M. (1998). *Cross cultural psychotherapy: The Arab Palestinian case*. New York: Haworth Press.
- Dwairy, M. (2004). Parenting styles and mental health of Palestinian-Arab adolescents in Israel. *Transcultural Psychiatry*, 41(2), 233-252.
- Dwairy, M., Achoui, M., Abouserie, R., & Farah, A. (2006). Parenting Styles, Individuation, and Mental Health of Arab Adolescents. *Journal of Cross-Cultural Psychology*, 37(3), 262-272.
- Dwairy, M., Achoui, M., Abouserie, R., Farah, A., Sakhleh, A., Fayad, M., et al. (2006). Parenting styles in Arab societies. *Journal of Cross-Cultural Psychology*, 37(3), 230-247.
- Dyer, C. (2007). Working children and educational inclusion in Yemen. *International Journal of Educational Development*, 27(5), 512-524.
- Easterly, W. (2002). The cartel of good intentions: The problem of bureaucracy in foreign aid. *The Journal of Policy Reform*, 5(4), 223 - 250.

- Easterly, W. (2006). *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York, NY: Penguin Press.
- Eccles, J., & Appleton-Gootman, J. (2001). *Community-based programs for youth: Lessons learned from general developmental research and from experimental and quasi-experimental evaluations*. Paper presented at the Urban Seminar Series.
- Eccles, J., Early, D., Fraser, K., Belansky, E., & McCarthy, K. (1997). The relation of connection, regulation, and support for autonomy to adolescents' functioning. *Journal of Adolescent Research*, 12.
- Eccles, J. S., & Appleton Gootman, J. (Eds.). (2002). *Community Programs to Promote Youth Development*. Washington, D.C.: National Academy Press.
- Eccles, J. S., & Roeser, R. W. (2011). Schools as Developmental Contexts During Adolescence. *Journal of Research on Adolescence*, 21(1), 225-241.
- Eden, D. (1988). Pygmalion, goal setting, and expectancy: Compatible ways to raise productivity. *Academy of Management Review*, 13(639-652).
- Edwards, M., & Hulme, D. (1996). Too close for comfort? The impact of official aid on nongovernmental organizations. *World Development*, 24, 961-973.
- EES. (2007). *EES Statement: The importance of methodologically diverse approach to impact evaluation -- specifically with respect to development aid and development interventions*. Nijkerk, The Netherlands: European Evaluation Society.
- El-Said, H., & Harrigan, J. (2009). "You Reap What You Plant": Social Networks in the Arab World--The Hashemite Kingdom of Jordan. *World Development*, 37(7), 1235-1249.
- Enzmann, D., Marshall, I. H., Killias, M., Junger-Tas, J., Steketee, M., & Gruszczynska, B. (2010). Self-reported youth delinquency in Europe and beyond: First results of the Second International Self-Report Delinquency Study in the context of police and victimization data. *European Journal of Criminology*, 7(2), 159-183.
- Ernst, M., Pine, D. S., & Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. *Psychological Medicine*, 36(03), 299-312.
- EU. (1999). *Resolution on youth participation*. Retrieved from http://europa.eu/legislation_summaries/other/c11604_en.htm.
- Euling, S. Y., Herman-Giddens, M. E., Lee, P. A., Selevan, S. G., Juul, A., Sorensen, T. I. A., et al. (2008). Examination of US Puberty-Timing Data from 1940 to 1994 for Secular Trends: Panel Findings. *Pediatrics*, 121(Supplement_3), S172-191.
- Evans, W. D., Ulasevich, A., & Blahut, S. (2004). Adult and Group Influences on Participation in Youth Empowerment Programs. *Health Educ Behav*, 31(5), 564-576.
- Eyango, V. M., & AlShaer, M. (2003). *Educational Assessment for Jordan 2004-2009 Country Strategy*. Amman: USAID/Jordan.
- Fagan, A., Van Horn, M., Hawkins, J., & Arthur, M. (2007). Gender Similarities and Differences in the Association Between Risk and Protective Factors and Self-Reported Serious Delinquency. *Prevention Science*, 8(2), 115-124.

- Farah, A. M. (1992). Guidance and counseling in the Hashemite kingdom of Jordan: Some observations. *International Journal for the Advancement of Counselling*, 15(1), 17-26.
- Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *BMC Medical Research Methodology*, 2, 8.
- Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). London: SAGE.
- Flannery-Schroeder, E. C., & Kendall, P. C. (2000). Group and Individual Cognitive-Behavioral Treatments for Youth with Anxiety Disorders: A Randomized Clinical Trial. *Cognitive Therapy and Research*, 24(3), 251-278.
- Flay, B., Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., et al. (2005). Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*, 6(3), 151-175.
- Frans, D. J. (1993). A Scale for Measuring Social Worker Empowerment. *Research on Social Work Practice*, 3(3), 312-328.
- Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing Moderator and Mediator Effects in Counseling Psychology Research. *Journal of Counseling Psychology*, 51(1), 115-134.
- Freeman, M. (2005). Beyond rhetoric: Extending rights to young people. In M. Barry (Ed.), *Youth Policy and Social Inclusion: Critical debates with young people* (pp. 55-68). London: Routledge.
- Freire, P. (1972). *Pedagogy of the oppressed*. Harmondsworth: Penguin.
- Freire, P. (2005). *Education for critical consciousness*. New York: Continuum.
- Friedenberg, E. Z. (1962). *The Vanishing Adolescent*. New York, NY: Dell Publishing.
- Fuller, G. E. (2003). *The Youth Factor: The New Demographics of the Middle East and the Implications for U.S. Policy*. Washington, D.C.: The Saban Center for the Middle East Policy at the Brookings Institution.
- Furman, W., & Buhrmester, D. (1992). Age and Sex Differences in Perceptions of Networks of Personal Relationships. *Child Development*, 63(1), 103-115.
- Fymier, A. B., Shulman, G. M., & Houser, M. (1996). The development of the learner empowerment measure. *Communication Education*, 43(3), 181-199.
- Gallagher, K. M., Stanley, A., Shearer, D., & Mosca, C. (2005). Implementation of youth development programs: promise and challenges. *Journal of Adolescent Health*, 37(3, Supplement 1), S61-S68.
- Gardner, F., Burton, J., & Klimes, I. (2006). Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: outcomes and mechanisms of change. *Journal of Child Psychology and Psychiatry*, 47(11), 1123-1132.
- Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who Benefits and How Does It Work? Moderators and Mediators of Outcome in an Effectiveness Trial of a Parenting Intervention. *Journal of Clinical Child & Adolescent Psychology*, 39(4), 568-580.
- Gardner, M., & Steinberg, L. (2005). Peer influence on risk-taking, risk preference, and risky decision-making in adolescence and adulthood: An experimental study. *Developmental Psychology*, 41, 625-635.

- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What Makes Professional Development Effective? Results From a National Sample of Teachers. *American Educational Research Journal*, 38(4), 915-945.
- Garmezy, N. (1985). Stress-resistant children: The search for protective factors. In J. E. Stevenson (Ed.), *Recent Research in Developmental Psychopathology* (pp. 213-233). Oxford, UK: Pergamon.
- Gartlehner, G., Hansen, R. A., Nissman, D., Lohr, K. N., & Carey, T. S. (2006). A simple and valid tool distinguished efficacy from effectiveness studies. *Journal of Clinical Epidemiology*, 59(10), 1040-1048.
- Gates, S., McCambridge, J., Smith, L., & Foxcroft, D. (2006). Interventions for prevention of drug use by young people delivered in non-school settings. *Cochrane Database of Systematic Reviews*(1).
- Gavin, L. A., & Furman, W. (1989). Age Differences in Adolescents' Perceptions of Their Peer Groups. *Developmental Psychology*, 25(5), 827-834.
- George, A. (2005). *Jordan: Living in the Crossfire*. London: Zed Books Ltd.
- Gilad, C., Stanley, M. G., & Dov, E. (2004). General self-efficacy and self-esteem: toward theoretical and empirical distinction between correlated self-evaluations. *Journal of Organizational Behavior*, 25(3), 375-395.
- Gittell, R., & Vidal, A. (1998). *Community organizing: Building social capital as a development strategy*. Thousand Oaks, CA: Sage.
- Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why Don't We See More Translation of Health Promotion Research to Practice? Rethinking the Efficacy-to-Effectiveness Transition. *American Journal of Public Health*, 93(8), 1261-1267.
- Glasser, W. (1990). *The quality school: Managing students without coercion*. New York: Harper & Row.
- Glasser, W. (1999). *Choice Theory in the Classroom*. New York, NY: Harper Collins.
- Glazer, S., Levy, D. M., & Myers, D. (2003). Nonexperimental Versus Experimental Estimates of Earnings Impacts. *The ANNALS of the American Academy of Political and Social Science*, 589(1), 63-93.
- Glaziou, P. (2003-2005). Sampsiz. Retrieved July 3, 2009, from <http://sampsiz.sourceforge.net/iface/s2.html> - means
- Godinot, X., & Quentin, W. (Eds.). (2006). *Participatory Approaches to Attacking Extreme Poverty*. Washington, D.C.: World Bank.
- Goffman, E. (1991). *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. Harmondsworth: Penguin.
- Golding, W. (1954). *Lord of the Flies*. Boston: Faber & Faber.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38(5), 581-586.
- Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of Amer Academy of Child & Adolescent Psychiatry*, 40(11), 1337-1345.
- Goodman, R., Meltzer, H., & Bailey, V. (1998). The strengths and difficulties questionnaire: A pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry*, 7(3), 125-130.

- Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is Small Beautiful? *Journal of Abnormal Child Psychology*, 27(1), 17-24.
- Gottfredson, D., Gerstenblith, S., Soule, D., Womer, S., & Lu, S. (2004). Do after school programs reduce delinquency? *Prevention Science*, 5(4), 253-266.
- Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1, 201-233.
- Gray, A., & Hayes, C. (2008). *Understanding the state of knowledge on youth engagement financing and sustainability*. Washington, DC: The Finance Project.
- Grolnick, W. S., Farkas, M. S., Sohmer, R., Michaels, S., & Valsiner, J. (2007). Facilitating motivation in young adolescents: Effects of an after-school program. *Journal of Applied Developmental Psychology*, 28(4), 332-344.
- Grossman, J. B., & Tierney, J. P. (1998). Does Mentoring Work? An Impact Study of the Big Brothers Big Sisters Program. *Evaluation Review*, 22(3), 403-426.
- Group, C. S. (2010). Resources for Review Authors. Retrieved September 15, 2010, from <http://webcache.googleusercontent.com/search?q=cache:-m3mkhLGqnkJ:szg.cochrane.org/resources-review-authors+cochrane+locf&cd=7&hl=en&ct=clnk&gl=uk&client=safari>
- Gundlach, M. J., Martinko, M. J., & Douglas, S. C. (2003). Emotional Intelligence, Causal Reasoning, and the Self-Efficacy Development Process. *International Journal of Organizational Analysis* (2003), 11(3), 229-246.
- Haase, C. M., & Silbereisen, R. K. (2011). Effects of positive affect on risk perceptions in adolescence and young adulthood. *Journal of Adolescence*, 34(1), 29-37.
- Hahn, A., Leavitt, T., & Aaron, P. (1994). *Evaluation of the Quantum Opportunities Program: Did the program work?* Waltham, MA: Brandeis University, Heller Graduate School.
- Hall, G., & Gruber, D. (2007). *Back to the future: engaging older youth. Issue Brief: The Massachusetts Special Commission on After School and Out of School Time*.
- Hall, G. S. (1904). *Adolescence*. New York: Appleton.
- Haller, G., Haller, D. M., Courvoisier, D. S., & Lovis, C. (2009). Research Paper: Handheld vs. Laptop Computers for Electronic Data Collection in Clinical Research: A Crossover Randomized Trial. *JAMIA*, 16, 651-659.
- Halpern, D. (2005). *Social Capital*. Cambridge: Polity.
- Hamer, R. M., & Simpson, P. M. (2009). Last Observation Carried Forward Versus Mixed Models in the Analysis of Psychiatric Clinical Trials. *Am J Psychiatry*, 166(6), 639-641.
- Hamid, S. (2011a). *The struggle for Middle East democracy*. Washington, D.C.: Brookings Institute.
- Hamid, S. (2011b, March 8, 2011). Why Middle East monarchies might hold on. *The Atlantic*.
- Hammand, S. H. (1999). The CRC: 'Words on paper' or a reality for children? A case study of Jordan. *The International Journal of Children's Rights*, 7, 215-237.
- Hannam, D. (2001). *A pilot study to evaluate the impact of student participation aspects of the citizenship order on standards of education in secondary schools*. Cambridge: Cambridge University.

- Harris, K., Ahluwalia, J., Okuyemi, K., Turner, J., & Woods, M. (2001). Addressing cultural sensitivity in a smoking cessation intervention: Development of the kick it at swope project. *Journal of Community Psychology*, 29(2).
- Harris, K. J., Ahluwalia, J. S., Okuyemi, K. S., Turner, J. R., Woods, M. N., Backinger, C. L., et al. (2001). Addressing cultural sensitivity in a smoking cessation intervention: Development of the Kick It at Swope project. *Journal of Community Psychology*, 29(4), 447-458.
- Hart, R. (1992). *Children's participation from tokenism to citizenship*. Florence: UNICEF Innocenti Research Centre.
- Hart, R. (2008). Stepping Back from 'The Ladder': Reflections on a Model of Participatory Work with Children. In A. Reid, B. Jensesn, J. Nikel & V. Simovska (Eds.), *Participation and Learning*. New York: Springer Netherlands.
- Hartley, J. (2004). Case Study Research. In C. Cassell & G. Symon (Eds.), *Essential Guide to Qualitative Methods in Organizational Research* (pp. 323-333). London: SAGE.
- Haskins, R., & Baron, J. (2011). *The Obama Administration's evidence-based social policy initiatives: An overview*. Washington, D.C.: Brookings Institution.
- Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ*, 328(7455), 1561-1563.
- HCY. (2004). *National Youth Strategy for Jordan: 2005-2009*. Amman: Higher Council for Youth.
- Healy, K. (2005). *Social work theories in context: Creating frameworks for practice*. New York, NY: Palgrave Macmillan.
- Heckman, J., Strixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411-482.
- Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54(1), 3-56.
- Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 312(5782), 1900-1902.
- Heckman, J. J., & Masterov, D. V. (2007). The Productivity Argument for Investing in Young Children. *Applied Economic Perspectives and Policy*, 29(3), 446-493.
- Hemmati, T., Mills, J. F., & Kroner, D. G. (2004). The validity of the Bar-On emotional intelligence quotient in an offender population. *Personality and Individual Differences*, 37(4), 695-706.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Beyond WEIRD: Toward a broad-based behavioral science. *Behavioral and Brain Sciences*, 33(2-3), 111-135.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. [10.1038/466029a]. *Nature*, 466(7302), 29-29.
- Henry, D. (2008). Changing classroom social settings through attention to norms. In M. Shinn & H. Yoshikawa (Eds.), *Toward Positive Youth Development: Transforming Schools and Community Programs*. New York: Oxford University Press.

- HFRP. (2006). Learning From Small-Scale Experimental Evaluations of After School Programs. Snapshot Number 8. *Harvard Family Research Project, Harvard University*.
- Higgins, J. P. T., & Green, S. (Eds.). (2009). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2: The Cochrane Collaboration*.
- Hill, L., Maucione, K., & K. Hood, B. (2007). A Focused Approach to Assessing Program Fidelity. *Prevention Science*, 8(1), 25-34.
- Hill, R. F., & Dennis Fortenberry, J. (1992). Adolescence as a culture-bound syndrome. *Social Science & Medicine*, 35(1), 73-80.
- Hillier, A. (2011). Uncertain future for youth participation groups. February 20. Retrieved May 30, 2011, from <http://www.cypnow.co.uk/news/1055871/Uncertain-future-youth-participation-groups/>
- Hirsch, B. J., & DuBois, D. L. (1992). The relation of peer social support and psychological symptomatology during the transition to junior high school: A two-year longitudinal analysis. *American Journal of Community Psychology*, 20(3), 333-347.
- Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.
- HM Treasury. (2010). *Spending Review 2010*. Retrieved from Office for Budgetary Responsibility.
- Hoffman, M. A., Ushpiz, V., & Levy-Shiff, R. (1988). Social support and self-esteem in adolescence. *Journal of Youth and Adolescence*, 17(4), 307-316.
- Hogarty, G. E., Flesher, S., Ulrich, R., Carter, M., Greenwald, D., Pogue-Geile, M., et al. (2004). Cognitive Enhancement Therapy for Schizophrenia: Effects of a 2-Year Randomized Trial on Cognition and Behavior. *Arch Gen Psychiatry*, 61(9), 866-876.
- Holden, D. J., Evans, W. D., Hinnant, L. W., & Messeri, P. (2005). Modeling Psychological Empowerment Among Youth Involved in Local Tobacco Control Efforts. *Health Education & Behavior*, 32(2), 264-278.
- Holden, D. J., Messeri, P., Evans, W. D., Crankshaw, E., & Ben-Davies, M. (2004). Conceptualizing Youth Empowerment within Tobacco Control. *Health Education & Behavior*, 31(5), 548-563.
- Howard, G. S., & Dailey, P. R. (1979). Response-shift bias: A source of contamination of self-report measures. *Journal of Applied Psychology*, 64(2), 144-150.
- Howick, J. (2009). Questioning the Methodologic Superiority of 'Placebo' Over 'Active' Controlled Trials *The American Journal of Bioethics*, 9(9), 34 - 48.
- Hoy, A. W., & Davis, H. A. (2006). Teacher self-efficacy and its influence on the achievement of adolescents. In F. Pajares & T. Urdan (Eds.), *Adolescence and Education: Self-Efficacy Beliefs of Adolescents* (Vol. V, pp. 117). Greenwich: Information Age Publishing.
- Huebner, A. J., & Betts, S. C. (2002). Exploring the Utility of Social Control Theory for Youth Development. *Youth & Society*, 34(2), 123-145.
- Hunter, J. P., & Csikszentmihalyi, M. (2003). The Positive Psychology of Interested Adolescents *Journal of Youth and Adolescence*, 32(1), 27-35.
- Hurrelmann, K. (1990). Parents, peers, teachers and other significant partners in adolescence. *International Journal of Adolescence and Youth*, 2(3), 211-236.

- Hutchings, J., Bywater, T., Daley, D., Gardner, F., Whitaker, C., Jones, K., et al. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *BMJ*, 334(7595), 678.
- Imas, L. G. M., & Rist, R. C. (2009). *The Road to Results: Designing and Conducting Effective Development Evaluations*. Washington, D.C.: The World Bank.
- IMF. (2011). World Economic Outlook Database. Retrieved May 1, 2011, 2011, from <http://www.imf.org/external/pubs/ft/weo/2011/01/weodata/weorept.aspx?sy=2008&ey=2011&scsm=1&ssd=1&sort=country&ds=.&br=1&c=439&s=NGDPD%2CNGDPDPC%2CPPPGDP%2CPPPPC%2CLP&grp=0&a=&pr.x=38&pr.y=14>
- IRIN. (2008, July 8, 2008). Yemen: New study highlights plight of street children *IRIN*. Retrieved from <http://www.irinnews.org/Report.aspx?ReportId=79145>
- Jadad, A. R., & Enkin, M. W. (2007). *Randomized Controlled Trials: Questions, Answers and Musings* (2nd ed.). Malden, MA: Blackwell Publishing.
- Jansen, W. (2006). Gender and the expansion of university education in Jordan. *Gender and Education*, 18(5), 473 - 490.
- Jennings, L., Parra-Medina, D., Messias, D., McLoughlin, K. (2006). Toward a critical social theory of youth empowerment. *Journal of Community Practice*, 14(1/2).
- Jones, K. (2004). Mission Drift in Qualitative Research, or Moving Toward a Systematic Review of Qualitative Studies, Moving Back to a More Systematic Narrative Review. *The Qualitative Report*, 9(1), 95-112.
- Jones, K. R., & Perkins, D. (2005a). Determining the quality of youth-adult relationships within community-based youth programs. *Journal of Extension*, 45(3).
- Jones, K. R., & Perkins, D. F. (2005b). Youth-adult partnerships. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental science: An encyclopedia of research, policies, and programs*. Thousand Oaks, CA: Sage Publications.
- Jones, K. R., & Perkins, D. F. (2006). Youth and Adult Perceptions of Their Relationships Within Community-Based Youth Programs. *Youth & Society*, 38(1), 90-109.
- Jones, L., Koegel, P., & Wells, K. B. (2008). Bringing Experimental Design to Community Partnered Participatory Research. In M. Minkler & N. Wallerstein (Eds.), *Community-Based Participatory Research for Health: From process to outcomes* (2nd ed.). San Francisco: Jossey-Bass.
- Jones, S., Brown, J., & Aber, L. (2008). Classroom settings as targets of intervention and research. In M. Shinn & H. Yoshikawa (Eds.), *Toward Positive Youth Development: Transforming Schools and Community Programs*. New York: Oxford University Press.
- Jordan Times. (2008). King expresses full support for projects proposed by Madaba youth. *Jordan Times*. Retrieved from <http://www.jordanembassyus.org/new/newsarchive/2008/05222008002.htm>

- Jose, P. (2008). ModGraph-I: A programme to compute cell means for the graphical display of moderational analyses: The internet version, Version 2.0. Retrieved June 21, 2011, from <http://www.victoria.ac.nz/psyc/paul-jose-files/modgraph/modgraph.php>
- Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits--self-esteem, generalized self-efficacy, locus of control, and emotional stability--with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*(1), 80-92.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology, 83*(3), 693-710.
- Kalichman, S. C., & Nachimson, D. (1999). Self-efficacy and disclosure of HIV-positive serostatus to sex partners. *Health Psychology, 18*(3), 281-287.
- Kaltiala-Heino, R., Marttunen, M., Rantanen, P., & Rimpela, M. (2003). Early puberty is associated with mental health problems in middle adolescence. *Social Science & Medicine, 57*(6), 1055-1064.
- Kanaan, T., Al-Salamat, M. N., & Hanania, M. (2009). *Higher Education in Jordan*. Amman: Jordan Center for Policy Research and Dialogue (JCPP).
- Kanaan, T., & Hanania, M. (2009). The disconnect between education, job growth, and employment in Jordan. In N. Dhillon & T. Yousef (Eds.), *Generation in Waiting: The Unfulfilled Promise of Young People in the Middle East* (pp. 142-165). Washington, D.C.: Brooking Institution Press.
- Kaplowitz, P. B., Slora, E. J., Wasserman, R. C., Pedlow, S. E., & Herman-Giddens, M. E. (2001). Earlier Onset of Puberty in Girls: Relation to Increased Body Mass Index and Race. *Pediatrics, 108*(2), 347-353.
- Kemmler, G., Hummer, M., Widschwendter, C., & Fleischhacker, W. W. (2005). Dropout Rates in Placebo-Controlled and Active-Control Clinical Trials of Antipsychotic Drugs: A Meta-analysis. *Arch Gen Psychiatry, 62*(12), 1305-1312.
- Kia-Keating, M., Dowdy, E., Morgan, M. L., & Noam, G. G. (2011). Protecting and promoting: An integrative conceptual model for healthy development of adolescents. *Journal of Adolescent Health, 48*(3), 220-228.
- Kim, S., Crutchfield, C., Williams, C., & Hepler, N. (1998). Toward a new paradigm in substance abuse and other problem behavior prevention for youth: Youth development and empowerment approach. *Journal of Drug Education, 28*(1), 1-17.
- Kirby, P. (2004). *A Guide to Actively Involving Young People in Research: For researchers, research commissioners, and managers*. London: INVOLVE Support Unit.
- Kirby, P., & Bryson, S. (2002). *Measuring the Magic? Evaluating and researching young people's participation in public decision making*. London: Carnegie Young People Initiative.
- Konrath, S. H., O'Brien, E. H., & Hsing, C. (2010). Changes in Dispositional Empathy in American College Students Over Time: A Meta-Analysis. *Personality and Social Psychology Review, XX*(X), 1-19.

- Kraemer, H. C., Frank, E., & Kupfer, D. J. (2006). Moderators of Treatment Outcomes. *JAMA: The Journal of the American Medical Association*, 296(10), 1286-1289.
- Kraemer, H. C., & Robinson, T. N. (2005). Are certain multicenter randomized clinical trial structures misleading clinical and policy decisions? *Contemporary Clinical Trials*, 26(5), 518-529.
- La Cava, G., Clert, C., & Lytle, P. (2004). *Investing in Youth Empowerment and Inclusion: A Social Development Approach*. Washington, DC: World Bank.
- Lahey, B. B., Schwab-Stone, M., Goodman, S. H., Waldman, I. D., Canino, G., Rathouz, P. J., et al. (2000). Age and gender differences in oppositional behavior and conduct problems: A cross-sectional household study of middle childhood and adolescence. *Journal of Abnormal Psychology*, 109(3), 488-503.
- Lakin, R., & Mahoney, A. (2006). Empowering youth to change their world: Identifying key components of a community service program to promote positive development. *Journal of School Psychology*, 44(6), 513-531.
- Lane, P. (2008). Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharmaceutical Statistics*, 7(2), 93-106.
- Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, 55(1), 170-183.
- Larson, R. W., & Richards, M. H. (1991). Boredom in the Middle School Years: Blaming Schools versus Blaming Students. *American Journal of Education*, 99(4), 418-443.
- Leigh, K. (2011, March 29). As Protests Turn Violent, Jordan Ponders Reforms. *TIME*.
- Leung, K., Lau, S., & Lam, W. L. (1998). Parenting styles and achievement: A cross-cultural study. *Merrill-Palmer Quarterly*, 44, 157-172.
- Leventhal, T., & Brooks-Gunn, J. (2003a). Children and youth in neighborhood contexts. *Current Directions in Psychological Science*, 12(1), 27-31.
- Leventhal, T., & Brooks-Gunn, J. (2003b). Moving to Opportunity: an Experimental Study of Neighborhood Effects on Mental Health. *Am J Public Health*, 93(9), 1576-1582.
- Leventhal, T., & Brooks-Gunn, J. (2004). A Randomized Study of Neighborhood Effects on Low-Income Children's Educational Outcomes. *Developmental Psychology*, 40(4), 488-507.
- Light, D., Method, F., Rockman, C., Cressman, G. M., & Daly, J. (2008). *Synthesis Report: Overview and Recommendation to the Jordan Education Initiative*. Amman: USAID/Jordan.
- Liu, M., Wei, L., & Zhang, J. (2006). Review of guidelines and literature for handling missing data in longitudinal clinical trials with a case study. *Pharmaceutical Statistics*, 5, 7-18.
- Lofquist, W. (1989). *The technology of prevention*. Tucson, AZ: Associates for Youth Development.
- Lorente, F. O., Souville, M., Griffet, J., & GrÈlot, L. (2004). Participation in sports and alcohol consumption among French adolescents. *Addictive Behaviors*, 29(5), 941-946.
- LoSciuto, L., Freeman, M. A., Harrington, E., Altman, B., & Lanphear, A. (1997). An Outcome Evaluation of the Woodrock Youth Development Project. *The Journal of Early Adolescence*, 17(1), 51-66.

- Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, 23(1), 151-169.
- Lusk, M. W., Peralta, F., & Vest, G. W. (1989). Street children of Juarez: A field study. *International social work*, 32(4), 289.
- Luszczynska, A., Scholz, U., & Schwarzer, R. (2005). The General Self-Efficacy Scale: Multicultural validation studies. *The Journal of Psychology*, 139(5), 439-457.
- Makhoul, J., Nakkash, R., El Hajj, T., Abdulrahim, S., Kanj, M., Mahfoud, Z., et al. (2010). Development and Validation of the Arab Youth Mental Health Scale. *Community Mental Health Journal*, 1-10.
- Marsh, H. (1989). Age and Sex Effects in Multiple Dimensions of Self-Concept: Preadolescence to Early Adulthood. *Journal of Educational Psychology*, 81(3), 417-430.
- Marshall, G. (1998). Opportunity structure. *A Dictionary of Sociology* Retrieved April 12, 2011, from <http://www.encyclopedia.com/doc/1088-opportunitystructure.html>
- Martin, J. J., & Gill, D. L. (1991). The relationships among competitive orientation, sport-confidence, self-efficacy, anxiety, and performance. *Journal of Sport & Exercise Psychology*, 23, 149-159.
- Masri, M., Jemni, M., A-Ghassani, A. M., & Badawi, A. A. (2010). *Entrepreneurship Education in the Arab States*. Beirut: UNESCO.
- Masten, A. S., & Garmezy, N. (1985). Risk, vulnerability, and protective factors in developmental psychopathology. *Advances in Clinical Child Psychology*, 8, 1-52.
- Maton, K. I., & Salem, D. A. (1995). Organizational characteristics of empowering community settings: A multiple case study approach. *American Journal of Community Psychology*, 23, 631-656.
- Matthews, H. (2001). Citizenship, Youth Councils and Young People's Participation. *Journal of Youth Studies*, 4(3), 299-318.
- Mayer, J. P., & Davidson, W. S. (2000). Dissemination of Innovation as Social Change. In J. Rappaport & E. Seidman (Eds.), *Handbook of Community Psychology* (pp. 421-438). New York: Kluwer Academic/Plenum Publishers.
- Mayo-Wilson, E. (2007). Reporting Implementation in Randomized Trials: Proposed Additions to the Consolidated Standards of Reporting Trials Statement. *Am J Public Health*, 97(4), 630-633.
- Mays, N., & Pope, C. (1995). Qualitative Research: Rigour and qualitative research. *BMJ*, 311(6997), 109-112.
- McCartney, K., & Rosenthal, R. (2000). Effect Size, Practical Importance, and Social Policy for Children. *Child Development*, 71(1), 173-180.
- McCord, J. (2003). Cures That Harm: Unanticipated outcomes of crime prevention programs. *ANNALS*, 587(May).
- McCord, J., & McCord, W. (1959). A Follow-up Report on the Cambridge-Somerville Youth Study. *The ANNALS of the American Academy of Political and Social Science*, 322(1), 89-96.
- MDRC. (2011). YouthBuild Evaluation. Retrieved May 13, 2011, from http://www.mdrc.org/project_32_103.html

- Meade, B., & Steiner, B. (2010). The total effects of boot camps that house juveniles: A systematic review of the evidence. *Journal of Criminal Justice*, 38(5), 841-853.
- Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2003). Mental health of children and adolescents in Great Britain. *International Review of Psychiatry*, 15, 185-187.
- Midgley, C., & Feldlaufer, H. (1987). Students' and Teachers' Decision-Making Fit Before and after the Transition to Junior High School. *The Journal of Early Adolescence*, 7(2), 225-241.
- Mihalic, S. (2009). *Implementation Fidelity*. Boulder, CO: Blueprints Initiative, Center for the Study and Prevention of Violence, University of Colorado.
- Miles, R. (2002). Employment and Unemployment in Jordan: The Importance of the Gender System. *World Development*, 30(3), 413-427.
- Miller, C. L., Druss, B. G., & Rohrbaugh, R. M. (2003). Using Qualitative Methods to Distill the Active Ingredients of a Multifaceted Intervention. *Psychiatr Serv*, 54(4), 568-571.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 40-48.
- Mills, J., Bonner, A., & Francis, K. (2006). The Development of Constructivist Grounded Theory. *International Journal of Qualitative Methods*, 5(1), 25-35.
- Mitra, D. L. (2005). Adults Advising Youth: Leading While Getting Out of the Way. *Educational Administration Quarterly*, 41(3), 520-553.
- MOE. (2008). *The Development of Education National Report of the Hashemite Kingdom of Jordan*. Amman: The Ministry of Education.
- Moffitt, T., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males. *Development and Psychopathology*, 8, 399-424.
- Mohajer, N., & Earnest, J. (2009). Youth empowerment for the most vulnerable: A model based on the pedagogy of Freire and experiences in the field. *Health Education*, 109(5), 424-438.
- Montgomery, P., Gardner, F., Operario, D., Mayo-Wilson, E., Tamayo, S., & Underhill, K. (2007). *The Oxford Implementation Reporting Index: Development of an indicator of treatment fidelity in systematic reviews of psychosocial interventions*.
- Morton, M. (2009a). *Applicability of Impact Evaluation to Cohesion Policy*. Brussels: European Commission.
- Morton, M. (2009b). Questscope Non-Formal Education Impact Study (QSNFEIS). Retrieved September 10, 2010, from <http://clinicaltrials.gov/ct2/show/NCT01001286>
- Morton, M. (2011). *Youth Empowerment with Out-of-School Adolescents in Jordan: A Multi-Method Study of Process and Impact*. University of Oxford, Oxford.
- Morton, M., & Montgomery, P. (2011). Empowerment-based non-formal education for out-of-school youth in Jordan: A pilot randomized controlled trial. *In submission*.

- Moulton, J. (1997). *Formal and Nonformal Education and Empowered Behavior: A Review of the Research Literature*. Washington, D.C.: Support for Analysis and Research in Africa (SARA), Academy for Educational Development.
- Mounts, N. & Steinberg, L. (1995). An ecological analysis of peer influence on adolescent grade point average and drug use. *American Psychological Association*, 31(6).
- MRC. (2008). *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council.
- Mukoma, W., Mathews, C., Flisher, A. J., Seebregts, C., Jansen, S., Ahmed, N., et al. (2004). *Use of electronic questionnaires on handheld devices to evaluate the effects of a school-based HIV prevention programme on adolescent sexual behavior*. Paper presented at the International Conference on AIDS.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30-38.
- Mumford, M. D., Zaccaro, S. J., Harding, F. D., Jacobs, T. O., & Fleishman, E. A. (2000). Leadership skills for a changing world: Solving complex social problems. *The Leadership Quarterly*, 11(1), 11-35.
- NASW. (2011). Code of Ethics of the National Association of Social Workers. Retrieved June 30, 2011, from <http://www.socialworkers.org/pubs/code/code.asp>
- National Research Council. (2002). *Community programs to promote youth development*. Washington, D.C.: National Academy Press.
- Natvig, G. K., Albrektsen, G., & Qvarnstrom, U. (2003). Associations between psychosocial factors and happiness among school adolescents. *International Journal of Nursing Practice*, 9(3), 166-175.
- NCFA. (2004). *Jordan country study of disadvantaged children*. Amman: The National Council for Family Affairs.
- NCFY. (2007). *Putting positive youth development into practice: A resource guide*. Washington, DC: National Clearinghouse on Families & Youth, US Department of Health & Human Services.
- Neuman, Y., Reiche, A., & Abu Saad, I. (1988). Organizational climate and work satisfaction: The case of Bedouin elementary schools in Israel. *The Journal of Educational Administration*, 26, 82-96.
- Newbegin, I., & Owens, A. (1996). Self-esteem and anxiety in secondary school achievement. *Journal of Social Behavior & Personality*, 11(3), 521-530.
- Nimeh, Z., & Bauchmuller, R. (2010). "Out of School: A look at deprived children in Jordan". Paper presented at the Poverty Reduction, Equity and Growth Network (PEGNet) annual conference.
- Norman, G. R. (1989). Issues in the use of change scores in randomized trials. *Journal of Clinical Epidemiology*, 42(11), 1097-1105.
- Nutbeam, D. (1999). Oakley's case for using randomised controlled trials is misleading. *BMJ*, 318(7188), 944b-.
- Nygreen, K., Kown, S., & Sanchez, P. (2006). Urban youth building community: Social change and participatory research in schools, homes, and community-based organizations. *Journal of Community Practice*, 14(1/2).

- Nylen, W. (2003). *Participatory Democracy Versus Elitist Democracy: Lessons from Brazil*. New York, NY: Palgrave Macmillan.
- O'Donoghue, J., Kirshner, B., & McLaughlin, M. (2002). Introduction: Moving youth participation forward. *New Directions for Youth Development*, 2002(96), 15-26.
- Oakley, A. (2001). Making Evidence-Based Practice Educational: A Rejoinder to John Elliott. *British Educational Research Journal*, 27(5), 575-576.
- Oakley, A., Strange, V., Bonell, C., Allen, E., Stephenson, J., & Team, R. S. (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ*, 332(7538), 413-416.
- OECD. (2005). *Definition and Selection of Key Competencies - Executive Summary*: OECD (Organisation for Economic Co-Operation and Development).
- Office of Management and Budget. (2011). *Budget of the U.S. Government: Fiscal Year 2010*. Retrieved from <http://www.whitehouse.gov/omb/budget/Overview/>.
- Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., et al. (1987). Ontario Child Health Study: II. Six-Month Prevalence of Disorder and Rates of Service Utilization. *Arch Gen Psychiatry*, 44(9), 832-836.
- Oliver, K. G., Collin, P., Burns, J., & Nicholas, J. (2006). Building resilience in young people through meaningful participation. *Australian e-Journal for the Advancement of Mental Health*, 5(1), 1-7.
- Olson-Merichko, J. A. (2005). *Youth leadership program evaluation: Implications for implementation*. Unpublished Dissertation, Indiana University of Pennsylvania, Indiana, PA.
- Olson-Merichko, J. A. (2006). *Youth leadership program evaluation: Implications for implementation*. Unpublished Dissertation, Indiana University of Pennsylvania, Indiana, PA.
- Onyx, J., & Bullen, P. (2000). Measuring Social Capital in Five Communities. *The Journal of Applied Behavioral Science*, 36(1), 23-42.
- Open Source Center. (2009). Jordan: Opinion Poll Finds 74% of People Fear Criticizing Government Openly. *Al-Ghad*.
- Orszag, P. (2009). Building Rigorous Evidence to Drive Policy. Retrieved May 30, 2011, from <http://www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidenceToDrivePolicy/>
- Overholser, J. C., Adams, D. M., Lehnert, K. L., & Brinkman, D. C. (1995). Self-Esteem Deficits and Suicidal Tendencies among Adolescents. *Journal of American Academy of Child & Adolescent Psychiatry*, 34(7), 919-928.
- Ozer, E., Cantor, J., Cruz, G., Fox, B., Hubbard, E., & Moret, L. (2008). The diffusion of youth-led participatory research in urban schools: The role of the prevention support system in implementation and sustainability. *American Journal of Community Psychology*, 41(3-4).
- Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research*, 66(4), 543-578.
- Pajares, F. (2006). Self-efficacy during childhood and adolescence: Implications for teachers and parents. In F. Pajares & T. Urdan (Eds.), *Adolescence and*

- education: Self-efficacy beliefs of adolescents* (Vol. V, pp. 339-366). Greenwich: Information Age Publishing.
- Pajares, F., & Urdan, T. (Eds.). (2006). *Adolescence and Education: Self-Efficacy Beliefs of Adolescents* (Vol. V). Greenwich: Information Age Publishing.
- Palen, L.-A., Graham, J. W., Smith, E. A., Caldwell, L. L., Mathews, C., & Flisher, A. J. (2008). Rates of Missing Responses in Personal Digital Assistant (PDA) Versus Paper Assessments. *Evaluation Review*, 32(3), 257-272.
- Palinkas, L., Aarons, G., Horwitz, S., Chamberlain, P., Hurlburt, M., & Landsverk, J. (2011). Mixed Method Designs in Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 44-53.
- Park, N. (2004). Character Strengths and Positive Youth Development. *The ANNALS of the American Academy of Political and Social Science*, 591(1), 40-54.
- Parker, J. D. A., Creque, R. E., Barnhart, D. L., Harris, J. I., Majeski, S. A., Wood, L. M., et al. (2004). Academic achievement in high school: does emotional intelligence matter? *Personality and Individual Differences*, 37(7), 1321-1330.
- Patall, E. A., Cooper, H., & Wynn, S. R. (2010). The effectiveness and relative importance of choice in the classroom. *Journal of Educational Psychology*, 102(4), 896-915.
- Patrick, B. C., Hisley, J., & Kempler, T. (2000). "What's Everybody So Excited About?": The Effects of Teacher Enthusiasm on Student Intrinsic Motivation and Vitality. *The Journal of Experimental Education*, 68(3), 217-236.
- Patterson, G. R., Degarmo, D., & Knutson, N. (2000). Hyperactive and antisocial behaviors: Comorbid or two points in the same process? *Development and Psychopathology*, 12(01), 91-106.
- Patterson, G. R., Dishion, T. J., & Yoerger, K. (2000). Adolescent Growth in New Forms of Problem Behavior: Macro- and Micro-Peer Dynamics. *Prevention Science*, 1(1), 3-13.
- Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods* (3 ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Perkins, D. (2009). Community Youth Development. In J. Wood & J. Hine (Eds.), *Work with Young People: Theory and Policy for Practice* (pp. 104-113). Thousand Oaks, CA: SAGE.
- Petersen, A. C. (1988). Adolescent Development. *Annual Review of Psychology*, 39(1), 583-607.
- Petrides, K. V., Frederickson, N., & Furnham, A. (2004). The role of trait emotional intelligence in academic performance and deviant behavior at school. *Personality and Individual Differences*, 36, 277-293.
- Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2003). Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency: A Systematic Review of the Randomized Experimental Evidence. *The ANNALS of the American Academy of Political and Social Science*, 589(1), 41-62.
- Pfeffer, J., & Salancik, G. (1978). *The External Control of Organizations*. New York: Harper & Row.
- Pianta, R., & Allen, J. (2008). Building capacity for positive youth development in secondary school classrooms: Changing teachers' interactions with students. In M. Shinn & H. Yoshikawa (Eds.), *Toward Positive Youth Development:*

- Transforming Schools and Community Programs*. New York: Oxford University Press.
- Pittman, K. (1999). The Power of Engagement. *Youth Today*. Retrieved from <http://www.forumforyouthinvestment.org/node/500>
- Pittman, K., & Irby, M. (2000). Unfinished business: Further reflections on a decade of promoting youth development. In P. P. Ventures (Ed.), *Youth development: Issues, challenges, and directions*. Philadelphia: Public/Private Ventures.
- Plato. (2009). *The Republic* (B. Jowett, Trans.): BN Publishing.
- Population Council. (2009). *The adolescent experience in-depth: Using data to identify and reach the most vulnerable young people: Jordan 2007*. New York: The Population Council, Inc.
- Possel, P., Baldus, C., Horn, A. B., Groen, G., & Hautzinger, M. (2005). Influence of general self-efficacy on the effects of a school-based universal primary prevention program of depressive symptoms in adolescents: a randomized and controlled follow-up study. *Journal of Child Psychology and Psychiatry*, 46(9), 982-994.
- Potter, R. B., Darmame, K., Barham, N., & Nortcliff, S. (2007). *An Introduction to Urban Geography of Amman, Jordan*: University of Reading.
- Potter, R. B., Darmame, K., Barham, N., & Nortcliff, S. (2009). "Ever-growing Amman", Jordan: Urban expansion, social polarisation and contemporary urban planning issues. *Habitat International*, 33(1), 81-92.
- Powell, W. W., & Friedkin, R. (1987). Organizational Change in Nonprofit Organizations. In W. W. Powell (Ed.), *The Nonprofit Sector: A Research Handbook*. New Haven, CT: Yale University Press.
- Preston, S. H. (1984). Children and the Elderly: Divergent Paths for America's Dependents. *Demography*, 21(4), 435-457.
- Prinz, R. J., & Sanders, M. R. (2007). Adopting a population-level approach to parenting and family support interventions. *Clinical Psychology Review*, 27(6), 739-749.
- Prowse, M. (2007). *Aid effectiveness: the role of qualitative research in impact evaluation*. London: Overseas Development Institute.
- Pursley, S. (2011) NFP manual. email correspondence. June 7, 2011. Denver, CO. Nurse-Family Partnership National Service Office
- Putnam, R. (1995). Tuning in, tuning out: the strange disappearance of social capital in America. *Political Science and Politics*, 28, 1-20.
- Putnam, R. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
- Puzzanchera, C. (2009, December). Juvenile Arrests 2008. *Juvenile Justice Bulletin*.
- Questscope. (2010). Closing the Gap: 2009 Annual Report. Questscope for Social Development in the Middle East.
- Rappaport, J. (1981). In praise of paradox: A social policy of empowerment over prevention. *American Journal of Community Psychology*, 9(1), 1-23.
- Raudebush, S. W., Martinez, A., Bloom, H., Zhu, P., & Lin, F. (2008). *An Eight-Step Paradigm for Studying the Reliability of Group-Level Measures*. New York, NY: W.T. Grant Foundation.

- Reynolds, B. (1951). *Social Work and Social Living: Exploration in Philosophy and Practice*. New York: Citadel Press.
- Rhodes, J., Lowe, S. R., Litchfield, L., & Walsh-Samp, K. (2008). The role of gender in youth mentoring relationship formation and duration. *Journal of Vocational Behavior*, 72(2), 183-192.
- RIATT. (2010). *Child and youth participation in East and Southern Africa: Taking stock and moving forward*: Regional Inter Agency Task Team on Children & AIDS.
- Riger, S. (1993). What's wrong with empowerment. *American Journal of Community Psychology*, 21(3), 279-292.
- Rodrik, D. (2009). The New Development Economics: We Shall Experiment, but How Shall We Learn? In J. Cohen & W. Easterly (Eds.), *What Works in Development? Thinking Big and Thinking Small*. Washington, D.C.: Brookings Institution Press.
- Roen, K., Arai, L., Roberts, H., & Popay, J. (2006). Extending systematic reviews to include evidence on implementation: Methodological work on a review of community-based initiatives to prevent injuries. *Social Science & Medicine*, 63(4), 1060-1071.
- Romer, D., Duckworth, A., Sznitman, S., & Park, S. (2010). Can Adolescents Learn Self-control? Delay of Gratification in the Development of Control over Risk Taking. *Prevention Science*, 11(3), 319-330.
- Rosen, M., & Maureen, S. (2001). *Changing the face of giving: An assessment of youth philanthropy*. San Francisco, CA: James G. Irvine Foundation.
- Rosenberg, M. (1989). *Society and the Adolescent Self-Image*. Middletown, CT: Wesleyan University Press.
- Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global Self-Esteem and Specific Self-Esteem: Different Concepts, Different Outcomes. *American Sociological Review*, 60(1), 141-156.
- Rosenthal, D., Moore, S., & Flynn, I. (1991). Adolescent self-efficacy, self-esteem and sexual risk-taking. *Journal of Community & Applied Social Psychology*, 1(2), 77-88.
- Roth, J., & Brooks-Gunn, J. (2003a). Youth development programs: Risk, prevention and policy. *Journal of Adolescent Health*, 32(3), 170-182.
- Roth, J., Brooks-Gunn, J., Murray, L., & Foster, W. (1998). Promoting healthy adolescence: Synthesis of youth development program evaluations. *Journal of Research on Adolescence*, 8(4), 423-459.
- Roth, J. L. (2004). Youth Development Programs. *Prevention Researcher*, 11(2), 3-7.
- Roth, J. L., & Brooks-Gunn, J. (2003b). What Exactly Is a Youth Development Program? Answers From Research and Practice. *Applied Developmental Science*, 7(2), 94-111.
- Roth, J. L., & Brooks-Gunn, J. (2003c). What is a Youth Development Program? Identification of Defining Principles. In F. Jacobs, D. Wertlieb & R. M. Lerner (Eds.), *Enhancing the Life Chances of Youth and Families: Contributions of Programs, Policies, and Service Systems* (Vol. 2, pp. 197). Thousand Oaks, CA: Sage Publications, Inc.

- Rothwell, P. M. (2005). Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *The Lancet*, 365(9454), 176-186.
- Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry*, 57, 316-331.
- Rutter, M. (2001). Psychosocial adversity: Risk, resilience and poverty. In J. M. Richman & M. W. Fraser (Eds.), *The context of youth violence: Resilience, risk, and protection*. Westport, CT: Praeger.
- Rutter, M., & Smith, D. J. (1995). Towards causal explanations of time trends in psychosocial disorders of youth. In M. Rutter & D. J. Smith (Eds.), *Psychosocial disorders in young people: Time trends and their causes*. Chichester: Wiley.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- Sabaratna, P., & Klein, J. D. (2006). The READY Tool: A Youth Development Outcomes Measure: University of Rochester, Division of Adolescent Medicine.
- Sackett, D. L., & Wennberg, J. E. (1997). Choosing the best research design for each question. *BMJ*, 315(7123), 1636-.
- Saleebey, D. (1996). The strengths perspective in social work practice: Extensions and cautions. *Social Work Research*, 41, 296-305.
- Salehi-Isfahani, D., & Dhillon, N. (2008). *Stalled Youth Transitions in the Middle East: A Framework for Policy Reform*. Washington, D.C.: Wolfensohn Center for Development at the Brookings Institution.
- Samad, L., Hollis, C., Prince, M., & Goodman, R. (2005). Child and adolescent psychopathology in a developing country: testing the validity of the Strengths and Difficulties Questionnaire (Urdu version). *International Journal of Methods in Psychiatric Research*, 14(3), 158-166.
- Sampson, R. J. (2008). Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure. *American Journal of Sociology*, 114(1), 189-231.
- Sampson, R. J., & Morenoff, J. D. (2000). Ecological perspectives on the neighborhood context of urban poverty: past and present. In J. Brooks-Gunn, G. J. Duncan & L. J. Aber (Eds.), *Neighborhood Poverty: Policy Implications in Studying Neighborhoods* (Vol. II): Russell Sage Foundation.
- Sanchez, V., Steckler, A., Nitirat, P., Hallfors, D., Cho, H., & Brodish, P. (2007). Fidelity of implementation in a treatment effectiveness trial of Reconnecting Youth. *Health Education Research*, 22(1), 95-107.
- Sandholtz, J. H. (2002). Inservice training or professional development: Contrasting opportunities in a school/university partnership. *Teaching and Teacher Education*, 18(7), 815-830.
- Scherbaum, C. A., Cohen-Charash, Y., & Kern, M. J. (2006). Measuring General Self-Efficacy. *Educational and Psychological Measurement*, 66(6), 1047-1063.
- Scheve, J., Perkins, D., & Mincemoyer, C. (2006). Collaborative teams for youth engagement. *Journal of Community Practice*, 14(1/2).
- Schirm, A., Rodriguez-Planas, N., Maxfield, M., & Tuttle, C. (2003). *The Quantum Opportunity Program Demonstration: Short-Term Impacts*. Washington, D.C.: Mathematica Policy Research, Inc.

- Schochet, P. Z. (2010). The Late Pretest Problem in Randomized Control Trials of Education Interventions. *Journal of Educational and Behavioral Statistics*, 35(4), 379-406.
- Schoenwald, S., Garland, A., Chapman, J., Frazier, S., Sheidow, A., & Southam-Gerow, M. (2011). Toward the Effective and Efficient Measurement of Implementation Fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 32-43.
- Scholz, U., Dona, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*.
- Schulz, A. J., Israel, B. A., Zimmerman, M. A., & Checkoway, B. N. (1995). Empowerment as a multi-level construct: perceived control at the individual, organizational and community levels. *Health Education Research*, 10(3), 309-327.
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine*, 8(18).
- Schulz, K. F., & Grimes, D. A. (2005). Sample size calculations in randomised trials: mandatory and mystical. *The Lancet*, 365(9467), 1348-1353.
- Schulz, S., & Shulman, G. (1993). *The development and assessment of the job empowerment instrument (JEI)*. Paper presented at the Joint Central States Communication Association and Southern States Communication Association annual convention.
- Schwandt, T. (2000). Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics, and social constructionism. In N. Denzin & Y. Lincoln (Eds.), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: Nfer-Nelson.
- Schweinhart, L. J., & Weikart, D. P. (1998). High/Scope Perry Preschool Program Effects at Age Twenty-Seven. In J. Crane (Ed.), *Social Programs That Work* (pp. 148-). New York, NY: Russell Sage Foundation.
- Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow up study of antisocial children into. *BMJ*, 323(7306), 191-.
- SealedEnvelope. (2000-2010). SealedEnvelope. Retrieved September 1, 2010, from <http://www.sealedenvelope.com>
- Search Institute. (2008). 40 Developmental Assets. Retrieved February 15, 2010, from <http://www.search-institute.org/content/40-developmental-assets-adolescents-ages-12-18>
- Selznick, P. (1957). *Leadership in Administration*. New York: Harper & Row.
- Sen, A. (1999). *Development as Freedom*. New York, NY: Anchor Books.
- Shapiro, J. P., Welker, C. J., & Jacobson, B. J. (1997). The youth client satisfaction questionnaire: Development, construct validation, and factor structure. *Journal of Clinical Child & Adolescent Psychology*, 26(1), 87-98.
- Sharp, E., Caldwell, L., Graham, J., & Ridenour, T. (2006). Individual Motivation and Parental Influence on Adolescents' Experiences of Interest in Free Time: A Longitudinal Examination. *Journal of Youth and Adolescence*, 35(3), 340-353.

- Sharp, J. M. (2010a). *Jordan: Background and U.S. Relations* (No. RL33546). Washington, D.C.: Congressional Research Service.
- Sharp, J. M. (2010b). *U.S. Foreign Assistance to the Middle East: Historical Background, Recent Trends, and the FY2011 Request*. Washington, D.C.: Congressional Research Service.
- Shelton, D. (2009a). Leadership, education, achievement, and development: A nursing intervention for prevention of youthful offending behavior. *Journal of the American Psychiatric Nurses* 14, 429-441.
- Shelton, D. (2009b). Leadership, Education, Achievement, and Development: A Nursing Intervention for Prevention of Youthful Offending Behavior. *Journal of the American Psychiatric Nurses Association*, 14(6), 429-441.
- Shelton, S. H. (1990). Developing the construct of general self-efficacy. *Psychological Reports*, 66, 987-994.
- Sherer, M., & Adams, C. H. (1983). Construct validation of the Self-Efficacy Scale. *Psychological Reports*, 53, 899-902.
- Shier, H. (2001). Pathways to participation: openings, opportunities and obligations. *Children & Society*, 15(2), 107-117.
- Shiner, M. (1999). Defining peer education. *Journal of Adolescence*, 22(4), 555-566.
- Shinn, M., & Yoshikawa, H. (Eds.). (2008). *Toward Positive Youth Development: Transforming Schools and Community Programs*. New York, NY: Oxford University Press.
- Shulman, G., & Luechauer, D. (1993). The empowering educator: A CQI approach to classroom leadership. In D. Hubbard (Ed.), *Continuous quality improvement: Making the transition to education* (pp. 424-453). Maryville, MO: Prescott Publishing.
- Sinclair, R. (2000). *Young People's Participation* (No. 3). Staffordshire: Making Research Count, Keele University.
- Slymen, D. J., & Hovell, M. F. (1997). Cluster versus individual randomization in adolescent tobacco and alcohol studies: illustrations for design decisions. *International Journal of Epidemiology*, 26(4), 765-771.
- Small, S., & Memmo, M. (2004). Contemporary models of youth development and problem prevention: Toward an integration of terms, concepts, and model. *Family Relations*, 51(3), 3-11.
- Smith, C., & Akiva, T. (2008). Quality accountability: Improving fidelity of broad developmentally focused interventions. In M. Shinn & H. Yoshikawa (Eds.), *Toward Positive Youth Development: Transforming Schools and Community Programs*. New York: Oxford University Press.
- Smith, C., & Hohmann, C. (2005). *Full findings from the Youth PQA Validation Study*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Smith, P. B. (2004). Acquiescent Response Bias as an Aspect of Cultural Communication Style. *Journal of Cross-Cultural Psychology*, 35(1), 50-61.
- Smyth, R. M. D., Kirkham, J. J., Jacoby, A., Altman, D. G., Gamble, C., & Williamson, P. R. (2011). Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *BMJ*, 342.
- Snook, S. (2000). *Friendly Fire*. Princeton: Princeton University Press.

- Sobel, M. E. (2006). What Do Randomized Studies of Housing Mobility Demonstrate? *Journal of the American Statistical Association*, 101(476), 1398-1407.
- Speer, P. W. (2008). Altering patterns of relationship and participation: Youth organizing as a setting-level intervention. In M. Shinn & H. Yoshikawa (Eds.), *Toward Positive Youth Development: Transforming Schools and Community Programs*. New York: Oxford University Press.
- Spence, S. H. (1995). *Social skills training: Enhancing social competence in children and adolescents*. Windsor, UK: The NFER-NELSON Publishing Company Ltd.
- Spence, S. H. (2003). Social skills training with children and young people: Theory, evidence and practice. *Child and Adolescent Health*, 8(2), 84-96.
- Spence, S. H. (2010) Email correspondence. 'Re: Social Skills Questionnaire'. September 5. September 5.
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124(2), 240-261.
- Steinberg, L. (2004). Risk Taking in Adolescence: What Changes, and Why? *Annals of the New York Academy of Sciences*, 1021(1), 51-58.
- Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, 28(1), 78-106.
- Steinberg, L., Lamborn, S. D., Darling, N., Mounts, N. S., & Dornbusch, S. M. (1994). Over-Time Changes in Adjustment and Competence among Adolescents from Authoritative, Authoritarian, Indulgent, and Neglectful Families. *Child Development*, 65(3), 754-770.
- Storck, J. (2000). Knowledge Diffusion through Strategic Communities. *Sloan Management Review*, 41(2).
- Streiner, D. (2008). Missing data and the trouble with LOCF. *Evidence-Based Mental Health*, 11, 3-5.
- Streiner, D. L. (2003). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 80(1), 99-103.
- Suhonen, R., Berg, A., Idvall, E., Kalafati, M., Katajisto, J., Land, L., et al. (2010). Adapting the Individualized Care Scale for cross-cultural comparison. *Scandinavian Journal of Caring Sciences*, 24(2), 392-403.
- Suleiman, A., Soleimanpour, S., & London, J. (2006a). Youth action for health through youth-led research. *Journal of Community Practice*, 14(1/2).
- Suleiman, A. B., Soleimanpour, S., & London, J. (2006b). Youth Action for Health Through Youth-Led Research. *Journal of Community Practice*, 14(1), 125 - 145.
- Tarnoff, C., & Lawson, M. L. (2009). *Foreign Aid: An Introduction to U.S. Programs and Policy*. Washington, D.C.: Congressional Research Service.
- Tayler, E., M.B., C., Heptinstall, E., & Danckaerts, M. (1996). Hyperactivity and conduct problems as risk factors for adolescent development. *Journal of the American Academy of Child & Adolescent Psychiatry* 35(9).
- Taylor, M. J. (2000). The influence of self-efficacy on alcohol use among American Indian. *Cultural Diversity and Ethnic Minority Psychology*, 6(2), 152-167.
- Thabet, A. A., Stretch, D., & Vostanis, P. (2000). Child Mental Health Problems in Arab Children: Application of the Strengths and Difficulties Questionnaire. *International Journal of Social Psychiatry*, 46(4), 266-280.

- The Economist. (2009, July 23). All change, no change. *Economist*. Retrieved from <http://www.economist.com/node/14027674>
- The Economist. (2011, February 3, 2011). Regional reverberations. *The Economist*.
- The Nuffield Foundation. (2009). *Time trends in adolescent well-being: Update 2009*. London: The Nuffield Foundation.
- Thomas, K., & Velthouse, B. (1990). Cognitive elements of empowerment: An "interpretive" model of intrinsic task motivation. *Academy of Management Review*, 15, 666-681.
- Tierney, J., Grossman, J., & Resch, N. (2000). *Making a difference: An impact evaluation of Big Brothers Big Sisters*. Philadelphia: Public/Private Ventures.
- Torgerson, D. J. (2001). Contamination in trials: is cluster randomisation the answer? *BMJ*, 322(7282), 355-357.
- Treseder, P. (1997). *Empowering children and young people: Promoting involvement in decision-making*: Save the Children.
- Tu, Y. K., Blance, A., Clerehugh, V., & Gilthorpe, M. S. (2005). Statistical power for analyses of changes in randomized controlled trials. *Journal of Dental Research*, 84(3), 283-287.
- Tyler, F. (2004). The role of prosocial communities in youth development. *PSYKHE*, 13(2), 3-15.
- UKDCSF. (2005). *Youth Matters - Green Paper*. London: UK Dept for Children, Schools and Families.
- UN. (2005). *United Nations World Youth Report*. Retrieved from <http://www.un.org/esa/socdev/unyin/wyr05.htm>.
- Underhill, K., Operario, D., & Montgomery, P. (2007). Abstinence-only programs for HIV infection prevention in high-income countries. *Cochrane Database of Systematic Reviews*(4).
- UNDP. (2010). International Human Development Indicators. Retrieved May 2, 2011, from <http://hdrstats.undp.org/en/countries/profiles/JOR.html>
- UNESCO. (1994). *The Salamanca Statement on Principles, Policy and Practice in Special Needs Education*. Salamanca, Spain: UNESCO.
- UNFPA. (2003). *Making 1 billion count: investing in adolescents' health and rights*. New York: UNFPA.
- UNFPA. (2008). Supporting Adolescents and Youth. Retrieved February 1, 2010 from <http://www.unfpa.org/adolescents/overview.htm>
- Ungar, M. (2001). The Social Construction of Resilience Among "Problem" Youth in Out-of-Home Placement: A Study of Health-Enhancing Deviance. *Child and Youth Care Forum*, 30(3), 137-154.
- UNHCR. (2004). Chronology for Palestinians in Jordan. Retrieved May 2, 2011, from <http://www.unhcr.org/refworld/country,,CHRON,JOR,,469f38aa1e,0.html>
- UNICEF. (2002). *Jordanian Youth: Their Lives and Views*. New York: UNICEF.
- UNICEF. (2005). *Promoting Healthy Lives: A World Fit for Children Indicator Guide*. New York: UNICEF.
- UNICEF. (2007). *Children in Jordan: Situation Analysis 2006/2007*. Amman: UNICEF.
- UNODC. (2002). *Street Children of Cairo and Alexandria: Drug Abuse Trends, Consequences and Responses*. Cairo: UNODC.

- UNPD. (2009). World Population Prospects: The 2008 Revision Population Database. Retrieved October 20, 2010, from <http://esa.un.org/unpp/index.asp?panel=2>
- USCRI. (2004). U.S. Committee for Refugees World Refugee Survey 2004. Retrieved May 1, 2011
- USDHHS. (2001). *Toward a Blueprint for Youth: Making positive youth development a national priority*. Washington, DC: US Department of Health and Human Services.
- USDOS. (2011). Background Note: Jordan. Retrieved May 2, 2011
- Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59(9), 920-925.
- Van Maanen, J. (1998). Different Strokes: Qualitative research in the Administrative Science Quarterly from 1956 to 1996. In J. Van Maanen (Ed.), *Qualitative Studies of Organizations* (pp. ix-xxxii). Thousand Oaks, CA: SAGE.
- Vickers, A. J., & Altman, D. G. (2001). Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ*, 323(7321), 1123-1124.
- Villarruel, F., Perkins, D., Keith, J., & Borden, L. (2003). *Community Youth Development: Programs, Policies, and Practices*. Thousand Oaks, CA: Sage Publications.
- Wald, M. a. M., T. (2003). *Connected by 25: Improving the live chances of the country's most vulnerable 14-24 year olds*. Stanford: Stanford University.
- Walker, K. E., & Arbretton, A. J. A. (2001). *Working Together to Build Beacon Centers in San Francisco: Evaluation findings from 1998-2000*. Philadelphia: Public/Private Ventures.
- Wallerstein, N. (2006). *What is the evidence on effectiveness of empowerment to improve health?* Copenhagen: World Health Organisation Regional Office for Europe (Health Evidence Network).
- Wallerstein, N., Sanchez-Merki, V., & Verlade, L. (2005). Freirian praxis in health education and community organizing: A case study of an adolescent prevention program. In M. Minkler (Ed.), *Community Organizing and Community Building for Health* (2nd ed.). New Brunswick, NJ: Rutgers University Press.
- Weick, K. E. (2003). Theory and practice in the real world. *The Oxford Handbook of Organization Theory: Meta-theoretical perspectives* (pp. 453-475). Oxford: Oxford University Press.
- Weil, D. N. (1997). Chapter 17 The economics of population aging. In R. R. Mark & S. Oded (Eds.), *Handbook of Population and Family Economics* (Vol. Volume 1, Part 2, pp. 967-1014): Elsevier.
- Weinberger, M., Oddone, E. Z., Henderson, W. G., Smith, D. M., Huey, J., Giobbie-Hurder, A., et al. (2001). Multisite Randomized Controlled Trials in Health Services Research: Scientific Challenges and Operational Issues. *Medical Care*, 39(6), 627-634.
- Weinhardt, L. S., Forsyth, A. D., Carey, M. P., Jaworski, B. C., & Durant, L. E. (1998). Reliability and validity of self-report measures of HIV-related sexual

- behavior: Progress since 1990 and recommendations for research and practice. *Archives of Sexual Behavior*, 27(2), 155-180.
- Weiss, C. H. (1995). Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families. In J. P. Connell, A. C. Kubisch, L. B. Schorr & C. H. Weiss (Eds.), *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*. Washington, DC: Aspen Institute.
- Weiss, H. B. (2004). *Understanding and Measuring Attendance in Out-of-School Time Programs*. Cambridge, MA: Harvard Family Research Project.
- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55(4), 542-549.
- Weller, A., Florian, V., & Mikulincer, M. (1995). Adolescents' Reports of Parental Division of Power in a Multicultural Society. *Journal of Research on Adolescence*, 5(4), 413-429.
- Werner, E., & Smith, R. (1992). *Overcoming the odds: High risk children from birth to adulthood*. Ithaca: Cornell University Press.
- West, S. G., Aiken, L. S., & Krull, J. L. (1996). Experimental Personality Designs: Analyzing Categorical by Continuous Variable Interactions. *Journal of Personality*, 64(1), 1-48.
- White, D., & Pitts, M. (1998). Educating young people about drugs: a systematic review. *Addiction*, 93(10), 1475-1487.
- WHO/UN. (1999). *Programming for adolescent health and development*. Geneva: World Health Organisation.
- Wiggins, M., Bonell, C., Sawtell, M., Austerberry, H., Burchett, H., Allen, E., et al. (2009). Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ*, 339(jul07_2), b2534-.
- Wiktorowicz, Q. (2002). The Political Limits to Nongovernmental Organizations in Jordan. *World Development*, 30(1), 77-93.
- Williamson, W., & Hakki, H. (2010). *Mapping study of non-state actors in Jordan*. Amman: The European Union's MED - Mediterranean Programme for the Hashemite Kingdom of Jordan.
- Willig, C. (2001). *Introducing qualitative research in psychology: Adventures in theory and method*. Buckingham: Open University Press.
- Wills, T. A., Vaccaro, D., & McNamara, G. (1992). The role of life events, family support, and competence in adolescent substance use: A test of vulnerability and protective factors. *American Journal of Community Psychology*, 20(3), 349-374.
- Wilson, N., Dasho, S., Martin, A. C., Wallerstein, N., Wang, C. C., & Minkler, M. (2007). Engaging Young Adolescents in Social Action Through Photovoice: The Youth Empowerment Strategies (YES!) Project. *The Journal of Early Adolescence*, 27(2), 241-261.
- Wilson, N., Minkler, M., Dasho, S., Carrillo, R., Wallerstein, N., & Garcia, D. (2006). Training Students as Facilitators in the Youth Empowerment Strategies (YES!) Project. *Journal of Community Practice*, 14(1), 201-217.

- Wilson, W. J. (1987). *The Truly Disadvantaged*. Chicago: University of Chicago Press.
- Winkleby, M. A., Feighery, E., Dunn, M., Krole, S., Ahn, D., & Killen, J. D. (2004). Effects of an advocacy intervention to reduce smoking among teenagers. *Arch Pediatr Adolesc Med*, 158(3), 269-275.
- Witherspoon, D., Schotland, M., Way, N., & Hughes, D. (2009). Connecting the Dots: How Connectedness to Multiple Contexts Influences the Psychological and Academic Adjustment of Urban Youth. *Applied Developmental Science*, 13(4), 199 - 216.
- Wong, N., Zimmerman, M., & Parker, E. (2010). A typology of youth participation and empowerment for child and adolescent health promotion. *American Journal of Community Psychology*, 46(1), 100-114.
- Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management Review*, 14(3), 361-384.
- Wooldredge, J. D. (1988). Differentiating the Effects of Juvenile Court Sentences on Eliminating Recidivism. *Journal of Research in Crime and Delinquency*, 25(3), 264-300.
- World Bank. (2003). Adolescent Nutrition. Retrieved February 1, 2010, 2010, from <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTHEALTHNUTRITIONANDPOPULATION/EXTPHAAG/0,,contentMDK:20587649~menuPK:1314796~pagePK:64229817~piPK:64229743~theSitePK:672263,00.html>
- World Bank. (2004). *Jordan Poverty Assessment*. Washington, D.C.: The World Bank.
- World Bank. (2006a). Investment Climate Survey Online. Retrieved May 2, 2011, from <http://www.enterprisesurveys.org/ExploreEconomies/?economyid=99&year=2006>
- World Bank. (2006b). *World development report 2007: Development and the next generation*. Washington, D.C.: The World Bank Group.
- World Bank. (2008). *Hashemite Kingdom of Jordan Resolving Jordan's Labor Market Paradox of Concurrent Economic Growth and High Unemployment* (No. Report No. 39201-JO). Washington, DC: World Bank.
- World Health Organisation. (2009). Adolescent Health and Development. Retrieved February 1, 2010, 2010, from http://www.searo.who.int/EN/Section13/Section1245_4980.htm
- Worrall, S. (2000). *Young People as Researchers: A learning and resource pack*. London: Save the Children UK and Joseph Rowntree Foundation.
- Yalom, I. D. (1995). *The theory and practice of group psychotherapy* (4th ed.). New York: Harper Collins.
- Yohalem, N., Wilson-Ahlstrom, A., Fischer, S., & Shinn, M. (2009). *Measuring Youth Program Quality: A Guide to Assessment Tools*. Washington, D.C.: The Forum for Youth Investment.
- Zeldin, S., McDaniel, A., Topitzes, D., & Callvert, M. (2000). *Youth in decision making: A study on the impacts of youth on adults and organizations*. Washington, DC: National 4H Council.
- Zeldin, S., Petrokubi, J., & Camino, L. (2008). *Youth-Adult Partnerships in Public Action: Principles, Organizational Culture & Outcomes*. Washington, D.C.: The Forum for Youth Investment.

- Zief, S., Lauver, S., & Maynard, R. (2006). Impacts of after-school programs on student outcomes: A systematic review for the Campbell Collaboration.
- Zimmerman, M. (1995). Psychological empowerment: Issues and illustrations. *American Journal of Community Psychology*, 23(5), 581-599.
- Zimmerman, M. (2000). Empowerment theory: Psychological, organizational and community levels of analysis. In J. Rappaport & E. Seidman (Eds.), *Handbook of community psychology*. New York: Kluwer Academic/Plenum Publishers.
- Zvoch, K. (2009). Treatment Fidelity in Multisite Evaluation. *American Journal of Evaluation*, 30(1), 44-61.

Appendices

Appendix 2.1. Databases and websites searched	459
Appendix 2.2. Typical professional outreach correspondence	460
Appendix 2.3. Screening guide.....	461
Appendix 2.4. Coding form	462
Appendix 2.5. Guide for appraising study quality	466
Appendix 2.6. Excluded studies.....	469
Appendix 2.7. Outcomes tables with excluded-school based studies.....	476
Appendix 4.1. Impact study informed consent form.....	478
Appendix 4.2. Distribution of age across RCT sample.....	481
Appendix 4.3. Impact study questionnaire	482
Appendix 4.4. Empowerment measure questionnaire	488
Appendix 5.1. Mann-Whitney U baseline statistics	491
Appendix 5.2. Confidence intervals for primary analysis	492
Appendix 5.3. Diagnostics and transformations of outcomes data	495
Appendix 5.4. Mann-Whitney U primary analysis statistics.....	501
Appendix 5.5. Alternative analysis with peer problems item dropped	502
Appendix 5.6. Full subgroup analysis statistics	504
Appendix 5.7. Moderator analysis results for age.....	514
Appendix 6.1. Qualitative interview informed consent form.....	516
Appendix 6.2. Qualitative interview guides	517

Appendix 2.1. Databases and websites searched

Database	Dates of coverage
Applied Social Science Index and Abstracts	1987 to June 20, 2010
Australian Educational Index	1979 to June 20, 2010
British Educational Index	1975 to June 20, 2010
CINAHL	Earliest to June 20, 2010
Cochrane Library (CENTRAL)	1950 to June 20, 2010
Dissertation and Theses Abstracts	Earliest to June 20, 2010
EMBASE	Earliest to June 20, 2010
ERIC	1966 to June 20, 2010
Medline	1950 to June wk 2, 2010
PsycInfo	1967 to June wk 3, 2010
Social Service Abstracts	1979 to June 20, 2010
Sociological Abstracts	1952 to June 20, 2010

Website database/publications page	Last date searched (coverage included earliest to latest publications available on website)
Chapin Hall (University of Chicago)	August 6, 2010
Out-of-School Time Program Research & Evaluation Database (Harvard Family Research Project)	August 6, 2010
Innovation Center	August 7, 2010
National Clearinghouse on Families & Youth (US Administration of Children & Families)	August 7, 2010
Public/Private Ventures	August 7, 2010
Search Institute	August 7, 2010
UNICEF Evaluation and Research Database (ERD)	August 7, 2010
Australian Clearinghouse for Youth Studies (ACYS)	August 7, 2010
National Council for Voluntary Youth Services (NCVYS) Publications	August 7, 2010
UK DCSF Inclusion Development Programme (IDP) Publication Catalogue	August 7, 2010
World Bank Poverty Impact Evaluations Database	August 7, 2010

Appendix 2.2. Typical professional outreach correspondence

Subject: Seeking evaluations of youth empowerment/participation programs

Dear [contact]:

On behalf of an Oxford University and Campbell Collaboration research initiative, I am writing to see if you or your colleagues know of any **evaluations that have been conducted on youth empowerment (sometimes called 'youth participation') programs**. We are in the process of conducting a systematic review of the effects of such interventions on outcomes for youth and adolescents. Your organization was identified as a good professional contact for this.

We are searching international databases, but we ask your help to make sure we do not miss any relevant studies. **Please could you send or refer us to any relevant impact evaluations regardless of their findings—including those published or not published; showing positive effects, no effects, or negative effects?** Looking at all studies on youth empowerment programs will help us better understand this topic and its complexities and provide the most useful analysis for the field of youth services.

For the purpose of this review, we define youth empowerment programs as programs or projects that:

- focus primarily on strengths-building activities
- regularly involve youth participation in decision-making processes for program planning or implementation
- provide supportive relationships between youth and adults or older youth mentors

If you are unsure as to whether an evaluation qualifies, please send it over. Youth empowerment programs could be found in a range of program areas such as community service, recreational, cultural, peer education, advocacy, youth councils, informal/non-formal education, and job/life skills preparation.

Please do not hesitate to contact me should you have any thoughts or questions. We would also be very grateful for any other relevant contacts or institutions that you think could be helpful in this search.

Thank you very much for your help!

Best wishes,
Matthew Morton
& Paul Montgomery

Centre for Evidence Based Intervention,
University of Oxford

Appendix 2.3. Screening guide

Criteria	Y/N
1. More than 75% of participants are adolescents (10-19)?	
2. Eligible setting and duration? Takes place primarily outside of formal education Provides a physically safe environment Convenes regularly (i.e., not a one-off activity) NOT: a juvenile justice program, residential program, therapeutic intervention, conference or workshop	
3. Formally integrates youth participation into program decision-making? Are youth intentionally involved in democratic decision-making processes, boards, advisory boards, workgroups, committees, councils, positions, or staffing roles that directly and regularly influence program decision-making?	
4. Supportive relationship with adult or older youth leader? Do participants have regular access to at least one adult or older young person (e.g., college volunteer) designated to work with the young people in the program?	
5. Focus <i>primarily</i> on capacity-building strategies (e.g., skill-building, assets development, or leadership development) Does not focus primarily on 'treating' existing problem-behaviors (e.g., punitive-based programs or therapy for a specific problem)	
6. Appropriate methodology? Is there a prospectively assigned control group that used randomization, matching, or statistical methods to establish a credible comparator?	

Appendix 2.4. Coding form

Study

Author queries

Status

Reason for exclusion

Location

Published?

Method of recruitment

Recruitment dates

Socioeconomic status

Ethnic/racial characteristics

Other participant details

Mean age (Int)

SD age (Int)

Min age (Int)

Female (Int)

Mean age (Con)

SD age (Con)

Min age (Con)

Female (Con)

Study design:

- (a) Randomized controlled trial
(b) Non-randomized trial w/comp group(s)
(c) No control group*

If (a):

Unit of randomization

Method of randomization

If (b):

Unit of allocation

Method of allocation

Total # Assigned

To Intv Group

To Cont Group

To Other Group

**Type of intervention
participation:**

Volunteer

**Class
Required**

Other

Paid/Staff

**Court
Required**

Baseline differences between groups

Intervention setting – school, community, etc.

Methods of empowerment (if*/how youth were involved in regular decision-making processes)

Hart's ladder classification (according to intvtn description)

Formal training/leadership preparation component?

**Service/advocacy
component?**

**Empowerment processes involving all study
participants or subgroup(s)?**

Type of adult presence (check all that apply)

No regular adult pres.*

Program facilitator

One-one mentoring

☐☐☐

Staff/paid

Volunteer

Teacher

☐☐☐

Other

☐Trained in
facilitating
empower't☐**Frequency of intervention****Duration of intervention****Intervention content and delivery, types of activities****Comparison type****Comparison detail – services, frequency,
exposure****Number and schedule of data collections****Notes****Outcomes****measured:**

Self-efficacy

Self-esteem

☐☐

Type(s)

Social supports/connect.'s

Social skills

Emotional intelligence

Coping & problem-solving

Civic engagement

Academic achievement

☐☐☐☐☐☐

Antisocial behavior

Outcome measures, validity, reliability

*Indicates exclusion criteria.

Appendix 2.5. Guide for appraising study quality

Topic	Item #	Descriptor	Comments
INTRODUCTION			
Title and abstract	1	Study design	
Background	2	Relationship of evaluator to intervention	
	3	Relationship of study sponsor to intervention	
	4	Explanation of the rationale for the study intervention	
Objectives	5	Specific goals/objectives and hypotheses	
	6	Logic model or theory of change	
METHODS			
Participants	7	Eligibility criteria for participants (i.e. target population)	
	8	Explanation of recruitment procedures	
Intervention	9	Precise details of the intended intervention	
	10	Precise details on the implementation of the intervention	
	11	Information about the activities of the control group	
	12	Information on possible contamination	
Outcomes	13	Clearly defined primary and secondary outcome measures	
	14	Outcome measures aligned with the goals of the intervention	
	15	Explanation of measurement instruments and information regarding their validity and reliability	

	16	Methods used to enhance the quality of the data (supplemental studies, multiple evaluations, training of data collectors)	
Sample size	17	Size of treatment and control groups	
	18	Use of power analysis to determine sample size	
Randomization (if applicable)	19	Explanation of the method used to generate the random allocation sequence, including details of any restrictions (e.g. blocking, stratification)	
	20	Parental consent for study participation received prior to random assignment	
	21	Explanation of allocation concealment	
	22	Groups were equated on pretest data for outcomes measures and other characteristics suspected of confounding the results	
Blinding	23	Researchers and assessors were blind as to which group participants belonged	
Statistical methods	24	Statistical methods used to compare groups for primary outcome(s) and for additional analyses, such as subgroup analyses	
	25	Appropriateness of methods chosen	
	26	Pretest measures of outcomes and other important variables collected at baseline and incorporated into the analysis	
RESULTS			
Attrition	27	Number in each group who withdrew from study	
	28	Number in each group who were lost to follow-up	

	29	Number excluded from analysis (give reason)	
	30	Attrition >20%: Completers statistically compared to non-completers	
	31	Attrition >20%: Baseline equivalence of analytic sample demonstrated	
Intention-to-treat	32	Whether the analysis was by "intention-to-treat"	
Outcomes and data reporting	33	For each outcome, a summary of results per group	
	34	Means and SDs reported	
	35	<i>p</i> -values and degrees of freedom reported	
	36	Effect sizes reported	
	37	Other value reported (specify)	
CONCLUSIONS			
Interpretation	38	Interpretation of the results, taking into account study hypotheses and sources of potential bias or imprecision	
	39	Use of observational/qualitative data to understand impact results	
External validity	40	Generalizability of results	
	41	Replicability of intervention	
Overall evidence	42	General interpretation of the results in the context of current evidence.	

Note: Adapted from Zief et al (2006) and *What Works Clearinghouse Evidence Standards for Reviewing Studies* (US Department of Education, Revised 2008).

Appendix 2.6. Excluded studies

Sixty-five studies reached the final level of review and were ultimately excluded. All study papers were fully screened (not just abstracts). Reasons for exclusion follow.

Study	Reason for exclusion
Allen et al 1997	Formal education-based
Baker & Hultsman 1998	Inadequate intervention details available (author contacted); not clear systematic empowerment
Branch et al 1987	Not systematic empowerment; no control group
Brieger et al 2001	Not systematic empowerment
Calabrese & Schumer 1986	Inadequate methods for establishing credible control group
Campbell et al 2008	Formal education-based; not systematic empowerment
Cater 2006	Inadequate methods for establishing credible control group
Cheadle et al 2001	Not systematic empowerment
Clarke et al 1986	Formal education-based; not systematic empowerment
Collum 2003	Not systematic empowerment
D'Onofrio et al 2002	Not systematic empowerment for observed youth
Drolet 1997	Not systematic empowerment
Ebreo et al 2002	Formal education based; not systematic empowerment (peer educators engaged in content delivery but not systematically in program decision-making)
Ferguson et al 1996	No control group (a large-scaled randomized evaluation of YouthBuild is planned to begin recruitment in Spring 2011)
Fertman & Chubb 1992	Formal education-based; not systematic empowerment
Forneris et al 2010	Not systematic empowerment for observed youth; formal education-based
Fors & Jarvis 1995	Not systematic empowerment for observed youth
Gabriel et al 1996	Formal education-based; over 25% below age 10; not clear systematic empowerment
Gottfredson et al 2004	Not systematic empowerment
Grolnick et al 2007	Not systematic empowerment
Hahn et al 1994	Not systematic empowerment
Hahn et al 1996	Data and full report unavailable (author contacted)
IDRA 1995	No control group
Johannes 2004	Inadequate methods for establishing credible control group; not systematic empowerment

Kahne & Bailey 1999	Not systematic empowerment
Komro et al 2001	Formal education-based; not systematic empowerment
Komro et al 2008	Majority of participants not involved in empowerment-based component and largely formal education-based
Kovatseff & Power 2005	No control group
Laird 2009	Formal education based; not systematic empowerment
Lakin & Mahoney 2006	Formal education-based
Langberg et al 2006	Not systematic empowerment
Lauver 2002	Not systematic empowerment
Litrownik et al 2000	Not systematic empowerment (parent-child intervention)
LoSciuto et al 1997	Not systematic empowerment for study participants
LoSciuto et al 1999	Not systematic empowerment
Mackey 2007	Formal education-based; no control group
Maro et al 2009	Not systematic empowerment for observed youths
Martin 2008	Formal education-based; not systematic empowerment
Mason & Chuang 2001	Below age; not systematic empowerment
Mcloughlin 2009	Formal education-based; inadequate methods for establishing credible control group
Melchior 1998	Formal education based; not systematic empowerment
Moody et al 2003	No control group; not systematic empowerment
Naar-King et al 2010	Not systematic empowerment
Patro 1999	Not systematic empowerment
Pearlman et al 2002	Adequate data and intervention details unavailable (author contacted)
Perry 1989	No control group; formal education-based; not systematic empowerment
Philliber et al 2002	Not systematic empowerment
Prince 1995	Not systematic empowerment for observed youth
Quane & Rankin 2006	Not a controlled trial
Saitzyk & Poorman 1994	Inadequate methods for establishing credible control group; not clear systematic empowerment
Schirm et al 2003	Not systematic empowerment
Shelton 2009	Not systematic empowerment for timeframe measured
Simmons & Parsons 1983	Not systematic empowerment
Singer & Garcia 1988	Inadequate methods to establish credible control group; inadequate details

	available on intervention and data (author contacted)
Stone 1994	Formal education-based; not systematic empowerment for observed youth
Tebes et al 2007	Inadequate methods for prospectively establishing credible control group; not clear systematic empowerment; further details unavailable (author contacted)
Thomas 2004	Formal education-based; not systematic empowerment
Valentine 1990	No control group; not systematic empowerment
Walker & Arbretton 2001	No prospective control group
Weiss et al 1998	Not systematic empowerment
White 2010	One-off training retreat intervention
Wiggins et al 2009	Not systematic empowerment
Winkleby et al 2001	No control group
Winkleby et al 2004	Formal education-based
Wright et al 2006	Not systematic empowerment

References

- Allen JP, Philliber S, Herrling S, Kuperminc GP. 1997. Preventing teen Pregnancy and Academic Failure: Experimental Evaluation of a Developmentally Based Approach. *Child Development* 68:729-42
- Baker D, Hultsman J. 1998. Thunderbirds Teen Center program evaluation, Michigan State University
- Branch AY, Liebenberg S, Smith TJ. 1987. Youth Conservation and Service Corps: Findings from a National Assessment, Public/Private Ventures, Philadelphia, PA
- Brieger WR, Delano GE, Lane CG, Oladepo O, Oyediran KA. 2001. West African Youth Initiative: outcome of a reproductive health education program. *Journal of Adolescent Health* 29:436-46
- Calabrese RL, Schumer H. 1986. The effects of service activities on adolescent alienation. *Adolescence* 21:675-87
- Campbell R, Starkey F, Holliday J, Audrey S, Bloor M, et al. 2008. An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): a cluster randomised trial. *Lancet* 371:1595-602
- Cater MDW. 2006. *How voice affects perceptions of relationship with adults, ownership, and engagement in youth*. Ph.D. Louisiana State University and Agricultural & Mechanical College, United States -- Louisiana. 132 pp.
- Cheadle A, Wagner E, Walls M, Diehr P, Bell M, et al. 2001. The effect of neighborhood-based community organizing: results from the Seattle Minority Youth Health Project. *Health Services Research* 36:671-89
- Clarke JH, MacPherson B, Holmes DR, Jones R. 1986. Reducing adolescent smoking: a comparison of peer-led, teacher-led, and expert interventions. *Journal of School Health* 56:102-6

- Collum RC, Jr. 2003. *An investigation of the impact of participation in an afterschool program on the academic achievement and social behavior of adolescent Black males defined as at-risk*. Ph.D. The Florida State University, United States -- Florida. 290 pp.
- D'Onofrio CN, Moskowitz JM, Braverman MT. 2002. Curtailing tobacco use among youth: evaluation of project 4-health. *Health Education & Behavior* 29:656-82
- Drolet M. 1997. Empowerment according to gender: Dramatisation as a tool for developing positive relationships among adolescents. [References]. *Social Work and Social Sciences Review* 7:170-83
- Ebreo A, Feist-Price S, Siewe Y, Zimmerman RS. 2002. Effects of peer education on the peer educators in a school-based HIV prevention program: where should peer education research go from here? *Health Education & Behavior* 29:411-23
- Ferguson RF, Clay PL, Snipes JC, Roaf P. 1996. YouthBuild in Developmental Perspective. A Formative Evaluation of the YouthBuild Demonstration Project, Massachusetts Inst. of Tech., Cambridge. Dept. of Urban Studies and Planning
- Fertman CI, Chubb NH. 1992. The effects of a psychoeducational program on adolescents' activity involvement, self-esteem, and locus of control. *Adolescence* 27:517-26
- Forneris T, Fries E, Meyer A, Buzzard M, Uguy S, et al. 2010. Results of a rural school-based peer-led intervention for youth: goals for health. *Journal of School Health* 80:57-65
- Fors SW, Jarvis S. 1995. Evaluation of a peer-led drug abuse risk reduction project for runaway/homeless youths. *Journal of Drug Education* 25:321-33
- Gabriel RM, Hopson T, Haskins M, Powell KE. 1996. Building relationships and resilience in the prevention of youth violence. *American Journal of Preventive Medicine* 12:48-55
- Gottfredson DC, Gerstenblith SA, Soule DA, Womer SC, Lu S. 2004. Do after school programs reduce delinquency? *Prevention Science* 5:253-66
- Grolnick WS, Farkas MS, Sohmer R, Michaels S, Valsiner J. 2007. Facilitating Motivation in Young Adolescents: Effects of an After- School Program. *Rep.* 28
- Hahn A, Aaron P, Anderson R. 1996. An External Evaluation of REAL: Rural Entrepreneurship through Action Learning, Center for Human Resources, Brandeis University, Waltham, MA
- Hahn A, Leavitt T, Aaron P. 1994. Evaluation of the Quantum Opportunities Program: Did the program work?, Brandeis University, Heller Graduate School, Waltham, MA
- IDRA. 1995. Youth Leadership, Intercultural Development Research Association, San Antonio, TX
- Johannes EM. 2004. Effects of paths(c) after school program on children's social environment and behavior. *Dissertation Abstracts International Section A: Humanities and Social Sciences* 64:4643
- Kahne J, Bailey K. 1999. The role of social capital in youth development: The case of "I have a dream" programs. *Educational Evaluation and Policy Analysis* 21:321-43
- Komro KA, Perry CL, Veblen-Mortenson S, Farbachsh K, Toomey TL, et al. 2008. Outcomes from a randomized controlled trial of a multi-component alcohol use preventive intervention for urban youth: Project Northland Chicago. [References]. *Addiction* 103:606-18
- Komro KA, Perry CL, Williams CL, Stigler MH, Farbachsh K, Veblen-Mortenson S. 2001. How did Project Northland reduce alcohol use among young adolescents? Analysis of mediating variables. *Health Education Research* 16:59-70

- Kovatseff N, Power T. 2005. Talking realities... young parenting : a peer education program. *Report*
- Laird M. 2009. An Evaluation Report of Student Attitude and Behavior Changes to the Tennessee Department of Education and Volunteer Tennessee on the Learn and Serve America School-Based Program, Lions-Quest, Nashville, TN
- Lakin R, Mahoney A. 2006. Empowering youth to change their world: Identifying key components of a community service program to promote positive development. *Journal of School Psychology* 44:513-31
- Langberg JM, Smith BH, Bogle KE, Schmidt JD, Cole WR, Pender CA. 2006. A pilot evaluation of small group Challenging Horizons Program (CHP): A randomized trial. *Journal of Applied School Psychology* 23:31-58
- Lauver SC. 2002. *Assessing the Benefits of an After-School Program for Urban Youth: An Impact and Process Evaluation*. Accessed on
- Litrownik AJ, Elder JP, Campbell NR, Ayala GX, Slymen DJ, et al. 2000. Evaluation of a tobacco and alcohol use prevention program for Hispanic migrant adolescents: promoting the protective factor of parent-child communication. *Preventive Medicine* 31:124-33
- LoSciuto L, Freeman MA, Harrington E, Altman B, Lanphear A. 1997. An outcome evaluation of the Woodrock Youth Development Project. *The Journal of Early Adolescence* 17:51-66
- LoSciuto L, Hilbert SM, Fox M, Porcellini L, Lanphear A. 1999. A two-year evaluation of the Woodrock Youth Development Project. *The Journal of Early Adolescence* 19:488-507
- Mackey JP. 2007. *The effects of pathways youth development program curriculum in enhancing the emotional and behavioral functioning of fifth and eighth grade urban youth*. Mackey, Jocelyn Patrice: Howard U , US
- Maro C, Roberts G, Sorensen M. 2009. Using sport to promote HIV/AIDS education for at-risk youths: An intervention using peer coaches in football. *Scandinavian Journal of Medicine & Science in Sports* 19:129-41
- Martin AJ. 2008. Enhancing student motivation and engagement: The effects of a multidimensional intervention. *Contemporary Educational Psychology* 33:239-69
- Mason MJ, Chuang S. 2001. Culturally-based after-school arts programming for low-income urban children: Adaptive and preventive effects. *Journal of Primary Prevention* 22:45-54
- McCloughlin CS. 2009. Positive Peer Group Interventions: An Alternative to Individualized Interventions for Promoting Prosocial Behavior in Potentially Disaffected Youth. *Electronic Journal of Research in Educational Psychology* 7:1131-56
- Melchior A. 1998. National Evaluation of Learn and Serve America School and Community-Based Programs. Final Report.
- Moody KA, Childs JC, Sepples SB. 2003. Intervening with at-risk youth: evaluation of the Youth Empowerment and Support Program. *Pediatric Nursing* 29:263-70
- Naar-King S, Parsons JT, Murphy D, Kolmodin K, Harris DR. 2010. A multisite randomized trial of a motivational intervention targeting multiple risks in youth living with HIV: initial effects on motivation, self-efficacy, and depression. *Journal of Adolescent Health* 46:422-8
- Patro ET. 1999. *The effects of a community service program on moral development among adolescent males*. Patro, Edward Theodore: Wilmington Coll (Delaware), US

- Pearlman DN, Camberg L, Wallace LJ, Symons P, Finison L. 2002. Tapping youth as agents for change: evaluation of a peer leadership HIV/AIDS intervention. *Journal of Adolescent Health* 31:31-9
- Perry CL. 1989. Prevention of alcohol use and abuse in adolescence: teacher- vs peer-led intervention. *Crisis: Journal of Crisis Intervention & Suicide* 10:52-61
- Philliber S, Kaye JW, Herrling S, West E. 2002. Preventing pregnancy and improving health care access among teenagers: an evaluation of the children's aid society-carrera program. *Perspectives on Sexual & Reproductive Health* 34:244-51
- Prince F. 1995. The relative effectiveness of a peer-led and adult-led smoking intervention program. *Adolescence* 30:187-94
- Quane JM, Rankin BH. 2006. Does it pay to participate? Neighborhood-based organizations and the social development of urban adolescents. [References]. *Children and Youth Services Review* 28:1229-50
- Saitzyk AR, Poorman M. 1994. Transition to Adolescence Program: A Program To Empower Early Adolescent Girls
- Schirm A, Stuart E, McKie A. 2003. The Quantum Opportunity Program Demonstration: Final Impacts. *Mathematica Policy Research, Inc.*
- Shelton D. 2009. Leadership, education, achievement, and development: A nursing intervention for prevention of youthful offending behavior. *Journal of the American Psychiatric Nurses Association* 14:429-41
- Simmons CH, Parsons RJ. 1983. Empowerment for role alternatives in adolescence. *Adolescence* 18:193-200
- Singer M, Garcia R. 1988. From Research to Intervention: Substance Abuse Prevention among Hispanic Adolescents. Report No. 3, Hispanic Health Council, Hartford, CT
- Stone CS. 1994. Critique of Changing adolescent propensities to use drugs: results from project ALERT [original article by Ellickson P et al appears in HEALTH EDUC Q 1993;20(2):227-42]. *Nursing Scan in Research* 7:13-4
- Tebes JK, Feinn R, Vanderploeg JJ, Chinman MJ, Shepard J, et al. 2007. Impact of a Positive Youth Development Program in Urban After-School Settings on the Prevention of Adolescent Substance Use. *Journal of Adolescent Health* 41:239-47
- Thomas O. 2004. *Young Empowered Sisters: Promoting psychological and behavioral well being among African American young women through a culturally relevant school-based intervention*. Ph.D. dissertation. Michigan State University, Lansing, MI
- Valentine NL. 1990. Youth-At-Risk and 4-H: Partners in Success. An Evaluation of a County 4-H Drop-Out Prevention Model
- Walker KE, Arbreton AJA. 2001. Working Together to Build Beacon Centers in San Francisco: Evaluation findings from 1998-2000, Public/Private Ventures, Philadelphia
- Weiss FL, Nicholson HJ. 1998. Friendly PEERsuasion(sm) against substance use: The Girls Incorporated(sm) model and evaluation. *Drugs and Society* 12:7-22
- White D. 2010. *The effects of youth participatory evaluation and youth community action training on positive youth development*. Ph.D. Oregon State University, United States - Oregon. 213 pp.
- Wiggins M, Bonell C, Sawtell M, Austerberry H, Burchett H, et al. 2009. Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ* 339:b2534

- Winkleby MA, Feighery E, Dunn M, Kole S, Ahn D, Killen JD. 2004. Effects of an advocacy intervention to reduce smoking among teenagers. *Archives of Pediatrics & Adolescent Medicine* 158:269-75
- Winkleby MA, Feighery EC, Altman DA, Kole S, Tencati E. 2001. Engaging ethnically diverse teens in a substance use prevention advocacy program. *American Journal of Health Promotion* 15:433-6
- Wright R, John L, Alaggia R, Sheel J. 2006. Community-based arts program for youth in low-income communities: A multi-method evaluation. *Child & Adolescent Social Work Journal* 23:635-52

Appendix 2.7. Outcomes tables with excluded-school based studies

Primary outcomes from included studies and excluded-school based studies ($k = 6$)

Study	Outcome	Results						
Outcome	# Studies (Combined N)	# Measured Outcomes	Outcomes favoring intervention		Outcomes favoring comparison		Null effects ($p > .10$)	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
General self-efficacy ^a	3 (n=210)	3	0	1	0	0	2	66
Specific self-efficacy ^b	2 (n=1,129)	3	1	0	0	0	2	66
Self-esteem ^c	1 (n=316)	1	0	0	0	0	0	n/a

^aOlson-Merichko (2006), Morton (2011), and Lakin & Mahoney (2006)*

^bBerg et al (2009): drug prevention and sexual behavior; Winkleby et al (2004): tobacco-related advocacy***

^cBerg et al (2009)** (had baseline differences)

* $p < .10$, ** $p < .05$, *** $p < .01$ intervention effect

Bolded included studies

Secondary outcomes from included studies and excluded-school based studies ($k = 6$)

Study	Outcome	Results						
Outcome	# Studies (Combined N)	# Measured Outcomes	Outcomes favoring intervention		Outcomes favoring comparison		Null effects ($p > .10$)	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
Social supports and connections ^a	2 (n=443)	5	1	0	0	0	4	80
Social skills ^b	3 (n=483)	5	1	0	0	0	2	66
Emotional intelligence ^c	1 (n=43)	1	1	0	0	0	0	0
Coping and problem-solving skills ^d	1 (n=40)	1	1	0	0	0	0	0
Civic engagement ^e	2 (n=856)	3	0	0	0	0	0	n/a
Academic performance ^f	1 (n=695)	1	1	0	0	0	0	0
Antisocial behavior ^g	5 (n=1,991)	11	6	1	0	0	4	36

^aBerg et al (2009): social connectedness and school bonding (ladder had baseline differences); Morton (2011): social supports of friends, social supports of family, and adult connectedness

^bBerg et al (2009): social assertiveness skills and social skills; Olson-Merichko (2006): team skills; Morton (2011): social skills and prosocial behavior

^cLakin & Mahoney (2006): empathy***

^dOlson-Merichko (2006): proactive coping**

^eLakin & Mahoney (2006): sense of civic responsibility; Winkleby et al (2004): tobacco-related community advocacy***

^fAllen et al (1997): academic failure***

^gBerg et al (2009): alcohol use*, marijuana use**, had sex, and number of sex partners**; Olson-Merichko (2006): delinquency, drug use, and alcohol use; Morton (2011): conduct problems**; Winkleby et al (2004): smoking***; Allen et al (1997): pregnancy** and academic suspension***

* $p < .10$, ** $p < .05$, *** $p < .01$ intervention effect

Bolded included studies

Appendix 4.1. Impact study informed consent form

This appendix included the English-translation of the Arabic informed consent form used in the impact study (the following two pages).



Consent for Nonformal Education Study

Please take time to understand this sheet carefully and decide whether you do or don't want to take part. Ask the Questscope research team if there is anything that is not clear or if you have questions. Thank you for reading this.



What is this study about? This study aims to learn what works best for helping out-of-school youth in Jordan to succeed in their lives and communities. Youth in Jordan deserve the very best. Questscope wants to learn from you so that they can give the best opportunities and supports for young Jordanians in the years to come.

Do I have to take part? Not at all. It is wonderful to us if you join the Questscope Nonformal Education research, but the decision is yours. Your decision to take part or not will not affect any help or support you get from anyone.

Note: If you decide to join and later on you want to stop, you are free to do so.

What do I do? If you decide to take part, you and your parent will first sign a consent form (on the next page), and then spend about an hour talking together and answering questions with a researcher. You will then either start Questscope Nonformal Education (QS NFE) right away, or you will be able to participate in a biweekly Club while waiting to start QS NFE in 4 months.

A computer will randomly determine whether you start right away or wait 4 months and get to take part in the Club. Either way, in 4 months, you will again spend about an hour talking together and answering questions with a researcher to check up on how you're doing and learn from your views about youth programs.

We will also check back on you and your progress 16 months and 24 months after you begin.

Why do I get randomly assigned? Randomly assigning youth to two groups, one that starts QS NFE right away and one that starts later, helps us to compare difference in effects between the two groups of youth that are very similar and better understand how successful the program is.

What if the study upsets me or makes me feel uncomfortable? If you have anything that annoys you, upsets you, or makes you feel uncomfortable, contact immediately your facilitator or club coordinator for problem solving. Or you can contact [research staff] at [mobile number].

What if I have a complaint? If there is anything to do with this research which you are unhappy with, you can complain to [staff name] at [mobile number] or you can come directly to Questscope at Al-Webedeh, Paris Circle, near the Arab Bank.

Will what I say be kept confidential? Anything you tell us about yourself will be kept strictly confidential and will not be told to anyone else. Any information about you would have your name changed so that you cannot be recognized from it.

What will happen to the results of the research study? The results of this study will be used to help Questscope strengthen its programs for youths and help inform Ministry of Education about what youth want and need to live the best lives possible.

Who is organizing and reviewing the research? This research is being organized by Questscope Social Development in the Middle East in partnership with the Centre for Evidence-Based Intervention at the University of Oxford. The University of Oxford's ethics group has approved the study.

Contact for further information: [Research staff] or [investigator], Questscope Social Development in the Middle East, P.O. Box 910729 Amman 11191 Jordan. [email, mobile number]

Thank you for reading this sheet. If you feel comfortable with everything, you can fill in the boxes and sign on the next page.

DO I WANT TO TAKE
PART?

- ☐ I have read and understand the information sheet for this study and have had the chance to ask questions.
- ☐ I understand that I have chosen to take part and that I am free to stop at any time, without giving any reason. This will not change any support or help I am getting.
- ☐ I agree that words I may say during the interview can be used, without giving my name, in the presentation of the research.
- ☐ I understand that I have an equal chance of starting Questscope Nonformal Education right away or participating in a recreational Club before starting Nonformal Education in 4 months. A computer randomly determines which option I get.

I agree to take part in the study.

Name of participant: _____

Participant signature: _____

Date: _____

Name of guardian: _____

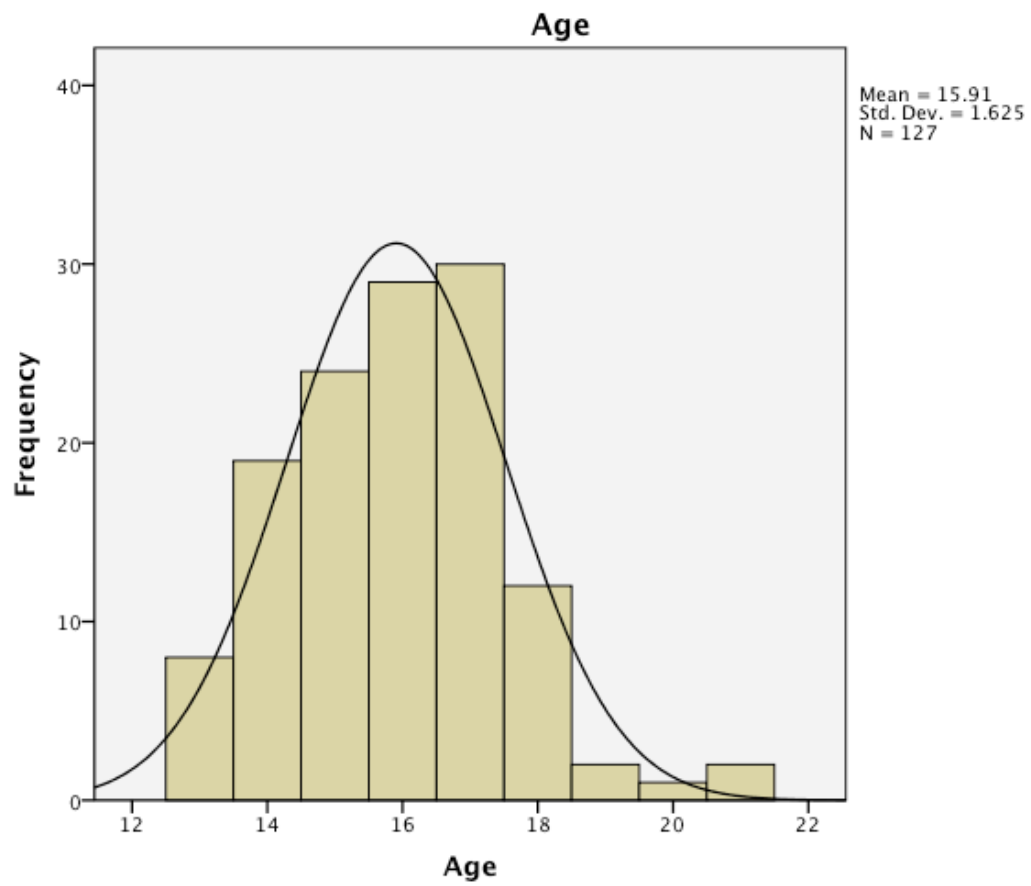
Guardian signature: _____

Date: _____



Contact for further information:
[Research staff] or [investigator],
Questscope Social Development in
the Middle East, P.O. Box 910729
Amman 11191 Jordan. [email, mobile
number]

Appendix 4.2. Distribution of age across RCT sample



Appendix 4.3. Impact study questionnaire

#	Scale [▲]	Item			
			Not true	Some what true	True
1.	GSE	I can always manage to solve difficult problems if I try hard enough.			
2.	SDQ-Pro	I try to be nice to other people. I care about their feelings			
3.	GSE	If someone opposes me, I can find the means and ways to get what I want.			
4.	SDQ-Pro	I usually share with others, for example food, clothes and games			
5.	SDQ-Con	*I get very angry and often lose my temper			
6.	GSE	It is easy for me to meet my needs and accomplish my dreams.			
7.	SDQ-PPr	*I would rather be alone than with people of my age			
8.	GSE	I am confident that I could deal in a good way with unexpected events.			
9.	SDQ-Con	I usually do as I am told			
10.	SDQ-Emo	*I worry a lot			
11.	SDQ-Pro	I am helpful if someone is hurt, sad or feeling ill			
12.	SDQ-PPr	I have one good friend or more			
13.	SDQ-Con	*I fight a lot. I can make other people do what I want			
14.	SDQ-Emo	*I am often unhappy, depressed or tearful			
15.	SDQ-PPr	Other people my age generally like me			
16.	GSE	Thanks to my skills and creativity, I know how to handle unforeseen situations.			
17.	SDQ-Hyp	*I am easily distracted, I find it difficult to concentrate			

18.	SDQ-Emo	*I am nervous in new situations. I easily lose confidence			
19.	GSE	I can solve most problems if I paid the necessary effort.			
20.	SDQ-Pro	I am kind to younger children			
21.	SDQ-Con	*I am often accused of lying or cheating			
22.	GSE	I can remain calm when facing difficulties because I can rely on my coping abilities.			
23.	SDQ-PPr	*Other young people pick on me or bully me			
24.	SDQ-Pro	I often offer to help others (parents, older people, children)			
25.	SDQ-Hyp	I think before I do things			
26.	GSE	When I am confronted with a problem, I can usually find several solutions.			
27.	SDQ-Con	*I take things that are not mine from home or elsewhere			
28.	SDQ-PPr	*I get along better with older people than with people my own age			
29.	GSE	If I am in trouble, I can usually think of a solution.			
30.	SDQ-Emo	*I have many fears, I am easily scared			
31.	SDQ-Hyp	I finish the work I'm doing, my attention is good			
32.	GSE	I can usually handle whatever comes my way.			
			Very little	Some what	A lot
33.	Ad-Conn	How much do adults in the neighborhood care about you?			
			Rarely	Somet imes	Most of the time
34.	SSup-Frd	*Some youth feel left out by their friends, but other youth don't. Do you feel left out by your friends?			
35.	SSup-Frd	Some youth are well-liked by their friends, but other youth aren't. Are you well-liked by your			

		friends?			
36.	SSup-Frd	*Some youth get picked on and teased by their friends, but other youth don't. Do you get picked on and teased by your friends?			
37.	SSup-Frd	*Some youths' friends make fun of them, but other youths' friends don't. Do your friends make fun of you?			
38.	SSup-Frd	Some youths' friends like to hear their ideas, but other youths' friends don't. Do your friends like to hear your ideas?			
39.	SSup-Frd	Some youth and their friends do a lot of things for each other, but other youth and their friends don't. Do you and your friends do a lot of things for each other?			
40.	SSup-Frd	Some youth feel very close to their friends, but other youth don't. Do you feel very close to your friends?			
41.	SSup-Frd	Some youth can count on their friends for help or advice when they have problems, but other youth can't. Can you count on your friends for help or advice when you have problems?			
42.	SSup-Frd	Some youth think their friends really care about them, but other youth don't. Do you think your friends care about you?			
43.	SSup-Frd	*Some youths' friends make them feel bad, but other youths' friends don't. Do your friends make you feel bad?			
44.	SSup-Fam	Some youth can count on their family for help or advice when they have problems, but other youth cannot. Can you count on your family for help or advice when you have problems?			
45.	SSup-Fam	Some youth and their families do a lot of things for each other, but other youth and their families don't. Do you and your family do a lot of things for each other?			
46.	SSup-Fam	*Some youths' families make them feel bad, but other youths' families don't. Does your family make you feel bad?			
47.	SSup-Fam	Some youth share a lot with their family, but other youth don't. Do you share a lot with your family?			

48.	SSup-Fam	*Some youth have a hard time talking to their family, but other youth don't. Do you have a hard time talking to your family?			
49.	SSup-Fam	Some youth feel like their family is there when they need them, but other youth don't feel this way. Do you feel like your family is there when you need them?			
50.	SSup-Fam	*Some youth feel left out by their family, but other youth don't. Do you feel left out by your family?			
51.	SSup-Fam	*Some youths' families ignore their ideas, but other youths' families don't. Does your family ignore your ideas?			
52.	SSup-Fam	Some youth are an important member of their family, but other youth are not. Are you an important member of your family?			
53.	SSup-Fam	Some youth think their families really care about them, but other youth think their families don't. Do you think your family cares about you?			
54.	SSup-Fam	Some youth feel like they belong in their family, but other youth feel like they don't belong. Do you feel like you belong in your family?			
55.	SSup-Fam	*Some youth think their families are unkind to them, but other youth don't. Do you think your family is unkind to you?			
'Please give the answer that best describes you <u>over the last 4 weeks.</u>'					
			Not true	Some what true	True
56.	SocSk	I listen to other people's point of view during arguments			
57.	SocSk	I make requests from my parents in a polite way			
58.	SocSk	I control my temper when I lose in a game or competition			
59.	SocSk	I control my temper if other youth tease me or say unkind things			
60.	SocSk	I ask other youth in a nice way if I want to join in			

		their activities			
61.	SocSk	I show other people if I feel affectionate or good toward them			
62.	SocSk	I do kind things for other people without having to be asked			
63.	SocSk	I say nice things to others when they deserve it			
64.	SocSk	I control my temper when I am told off or criticized by parents			
65.	SocSk	I ask permission before I borrow or use other people's things			
66.	SocSk	I share things with other youth			
67.	SocSk	I control my temper during disagreements with other youth			
68.	SocSk	I ask other youth if I can join in their activities			
69.	SocSk	The expression on my face is usually right (not always angry or grinning)			
70.	SocSk	I say I am sorry when I do something wrong			
71.	SocSk	I join in family activities			
72.	SocSk	I invite other youth to join in my games or activities			
73.	SocSk	I tell a parent or older person if I have a problem or need help			
74.	SocSk	I show that I am worried or that I care if someone is hurt or upset			
75.	SocSk	I follow the rules in games or sports			
76.	SocSk	I take part in games and activities with other youth			
77.	SocSk	I have conversations with adults			
78.	SocSk	I look people in the eye when we are talking			
79.	SocSk	My voice is usually nice (normal and not aggressive)			
80.	SocSk	I control my temper when I do not get my own way with my parents or older people			

81.	SocSk	I laugh or smile when I am happy or hear funny things			
82.	SocSk	I have conversations with other youth			
83.	SocSk	I show other people that I am listening when we are talking			
84.	SocSk	I can show people when I am angry without losing my temper			
85.	SocSk	I stand up for myself if other youth behave badly toward me without losing my temper			
* Reverse-scored ♣ Scale: GSE (General Self-Efficacy scale); SDQ (Strengths & Difficulties Questionnaire); -Con (conduct problems subscale); -Emo (emotional symptoms subscale); -Hyp (hyperactivity subscale); -PPr (peer problems subscale); -Pro (prosocial subscale); Ad-Conn (local adult connectedness); SSUp (Social Supports Appraisal Scale); -Frd (friends subscale); -Fam (family subscale); SocSk (Social Skills Questionnaire)					

Appendix 4.4. Empowerment measure questionnaire

#	Domain [*]	Item		
			Yes	No
1.	CO	In the program, I feel confident that I can do what I am expected to do.		
2.	IM	I have the power to make a difference in how things are done in the group of the non-formal education program.		
3.	IM	My participation is important to the success of the group.		
4.	CO	*The facilitators that are working with us make me feel like I am not good enough.		
5.	IM	I actively take part in the tasks required of the program.		
6.	ME	I usually do more work than what I have to do in the program.		
7.	CO	*I am overwhelmed by all of the work the program requires.		
8.	ME	I work hard for the program because I want to, not because I have to.		
9.	IM	I have a choice in the way I go about doing my work for the program.		
10.	ME	The things I have to do for the program mean a lot to me personally.		
11.	ME	I like to talk about what I am doing in the program with friends or family.		
12.	CO	*I feel nervous about things asked to be done within the program.		
13.	ME	I look forward to meeting the group everyday.		
14.	IM	*The facilitators who are working with us believe that they must control how I do my tasks for the program.		
15.	IM	Expressing my own attitudes and ideas is rewarded in the program.		

16.	ME	I agree with the things I have to do in the program.		
17.	CO	I had or learned the skills I needed to do well in the program.		
18.	IM	I am given enough opportunity to succeed in the program.		
19.	IM	*The facilitators who are working with us feel they are always right.		
20.	ME	I find the program to be exciting and energizing.		
21.	ME	I find the program to be interesting.		
22.	ME	The things I have to do in the program are valuable to me.		
23.	ME	The things I learned and did in the program will be helpful for my future.		
24.	IM	This program is one of the first times I feel like I am able to be in control of something important.		
25.	CO	This program taught me how to work well together with others.		
26.	CO	This program helped me better understand some of the problems in my neighborhood and in my country.		
27.	IM	This program gave me the chance to do something about the problems in my neighborhood and in my country.		
28.	IM	This program helped me connect with people in my class and community in order to help others.		
29.	IM	My opinion matters in determining program learning topics.		
30.	IM	My opinion matters in establishing program rules and expectations.		
31.	IM	My opinion matters in determining program activities.		
32.	ME	I would advise other youth to join the program.		
33.	CO	The facilitators say positive and supportive things to me.		

* Reverse-scored

♣ Domain: CO (competence); ME (meaningfulness); IM (impact)

Single-item addition included in the survey but not part of the empowerment measure:

34.		The non-formal education program increases my belief in the value of education.		

Appendix 5.1. Mann-Whitney U baseline statistics

The following tables include Mann-Whitney U test outputs on differences between intervention and control groups at baseline according to the study's outcome measures. No significant differences were found.

Test Statistics^a

	Pretest GSE	Pretest Conn	Pretest SSupFrd	Pretest SSupFam	Pretest SocSk
Mann-Whitney U	1709.500	1787.000	1826.500	1942.500	1909.000
Wilcoxon W	3539.500	4065.000	3656.500	3772.500	3739.000
Z	-1.459	-1.145	-.891	-.327	-.488
Asymp. Sig. (2-tailed)	.145	.252	.373	.744	.625

a. Grouping Variable: group

Test Statistics^a

	Pretest Conduct	Pretest PeerPr	Pretest EmotSym	Pretest Hyperact	Pretest Difficulties
Mann-Whitney U	1751.000	1891.500	1822.000	1679.000	1827.500
Wilcoxon W	3581.000	4169.500	3652.000	3509.000	3657.500
Z	-1.270	-.581	-.919	-1.661	-.883
Asymp. Sig. (2-tailed)	.204	.562	.358	.097	.377

a. Grouping Variable: group

Test Statistics^a

	Pretest Prosoc
Mann-Whitney U	1915.000
Wilcoxon W	4193.000
Z	-.467
Asymp. Sig. (2-tailed)	.640

a. Grouping Variable: group

Appendix 5.2. Confidence intervals for primary analysis

Pairwise Comparisons for General Self-Efficacy

Dependent Variable: Posttest GSE

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-.545	.596	.363	-1.725	.635
intervention	control	.545	.596	.363	-.635	1.725

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Social Skills

Dependent Variable: Soc Skills Post Transf

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-24.957	20.694	.230	-65.917	16.003
intervention	control	24.957	20.694	.230	-16.003	65.917

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Social Support of Friends

Dependent Variable: Posttest SSupFrd

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-.128	.552	.818	-1.219	.964
intervention	control	.128	.552	.818	-.964	1.219

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Social Support of Family

Dependent Variable: Posttest SSupFam

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	.021	.685	.975	-1.335	1.378
intervention	control	-.021	.685	.975	-1.378	1.335

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Local Adult Connectedness

Dependent Variable: Posttest Conn

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-.088	.140	.531	-.364	.189
intervention	control	.088	.140	.531	-.189	.364

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Overall Difficulties, SDQ

Dependent Variable: Difficulties Post Transf

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-2.141	2.080	.305	-6.257	1.976
intervention	control	2.141	2.080	.305	-1.976	6.257

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Conduct Problems, SDQ

Dependent Variable: Posttest Conduct

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-.802*	.342	.021	-1.479	-.126
intervention	control	.802*	.342	.021	.126	1.479

Based on estimated marginal means

*. The mean difference is significant at the

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Hyperactivity, SDQ

Dependent Variable: Hyperact Post Transf

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-.032	.419	.940	-.861	.798
intervention	control	.032	.419	.940	-.798	.861

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Emotional Symptoms, SDQ

Dependent Variable: Emot Sym Post Transf

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-1.738	1.266	.172	-4.244	.768
intervention	control	1.738	1.266	.172	-.768	4.244

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Peer Problems, SDQ

Dependent Variable: Posttest PeerPr

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	.342	.284	.231	-.220	.904
intervention	control	-.342	.284	.231	-.904	.220

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Pairwise Comparisons for Prosocial, SDQ

Dependent Variable: Posttest Prosoc

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
control	intervention	-.012	.318	.969	-.643	.618
intervention	control	.012	.318	.969	-.618	.643

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

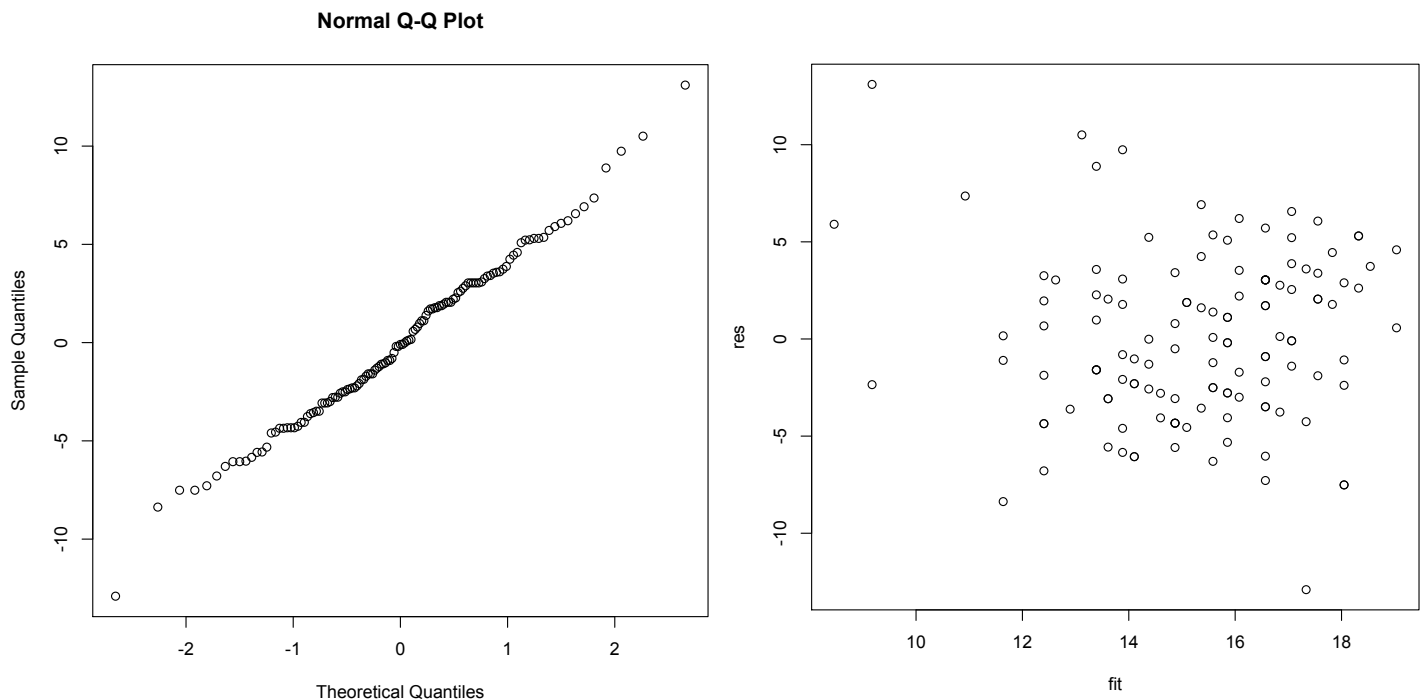
Appendix 5.3. Diagnostics and transformations of outcomes data

This appendix includes normal Q-Q plots and scatter plots for all outcome measures as guide to visual diagnostics of whether data meet parametric test assumptions. Where data appeared to deviate from a normal distribution, a Box-Cox transformation was conducted on the data, an image of the normal Q-Q plot with transformed data is superimposed over a normal Q-Q plot with the original data in order to give an easy visual reference for the effect of transformations. The following diagnostics and transformations were conducted in R Console™.

General Self-Efficacy

Decision: Original data acceptable

Regression⁵: `outGSE <- lm(GSE ~ group + GSE_A)`



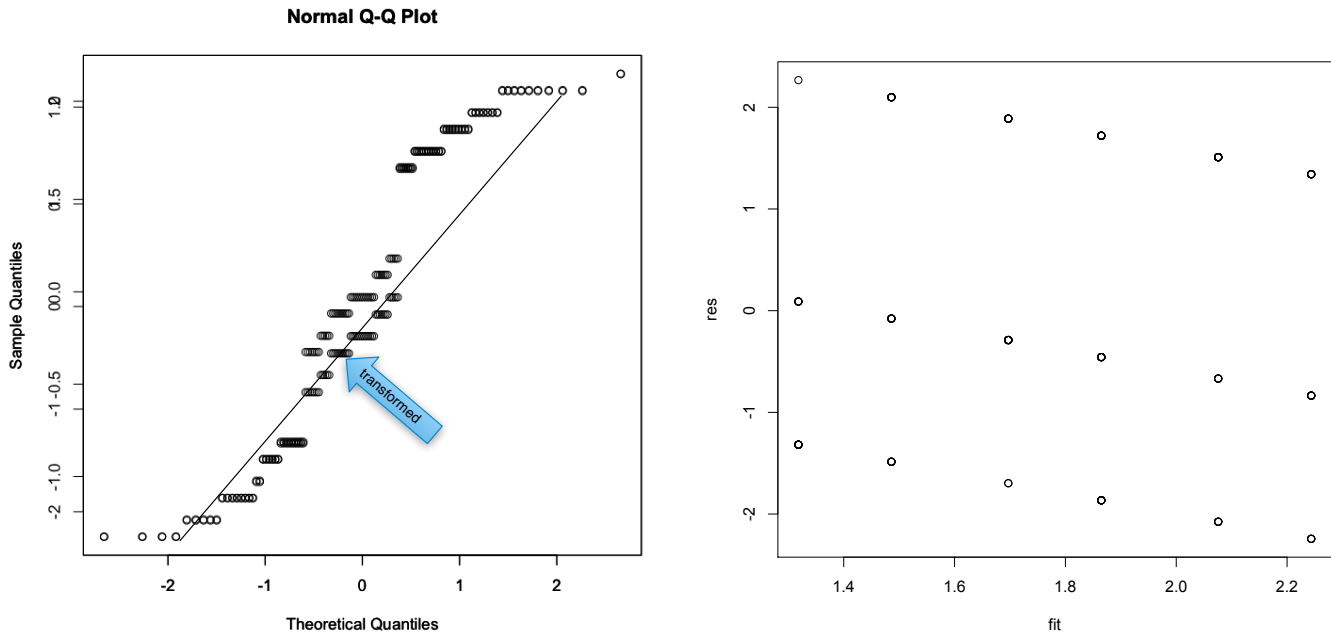
⁵ All diagnostics are conducted on the same linear regression model fit to the data via ANCOVA for primary analysis. Variable_A: pretest covariate. If variableBC: Box-Cox transformation applied.

Local Adult Connectedness

Decision: Original and transformed data unacceptable (single-item measure); Mann-Whitney U test used.

```
ConnBC <- (Conn^1.85 - 1)/1.85
```

```
outConnBC <- lm(ConnBC ~ group + Conn_A)
```

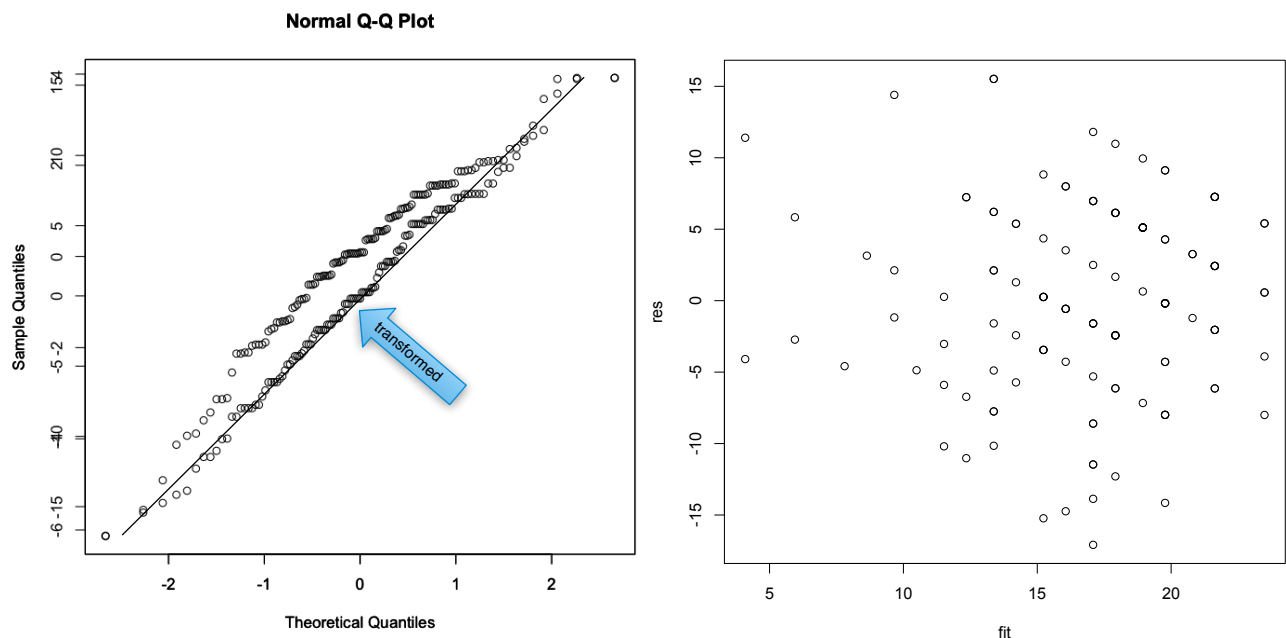


Conduct problems - SDQ

Decision: Use transformed data

```
ConductBC <- (Conduct^1.7 - 1)/1.7
```

```
outConductBC <- lm(ConductBC ~ group + Conduct_A)
```



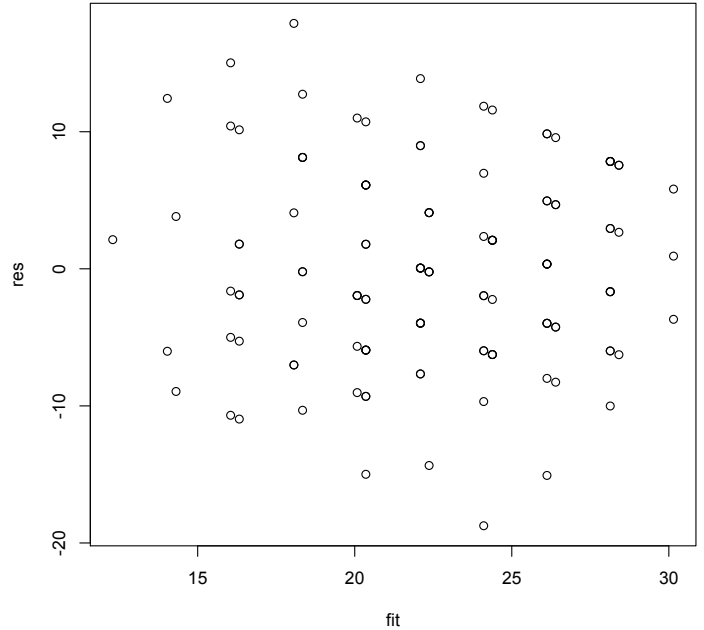
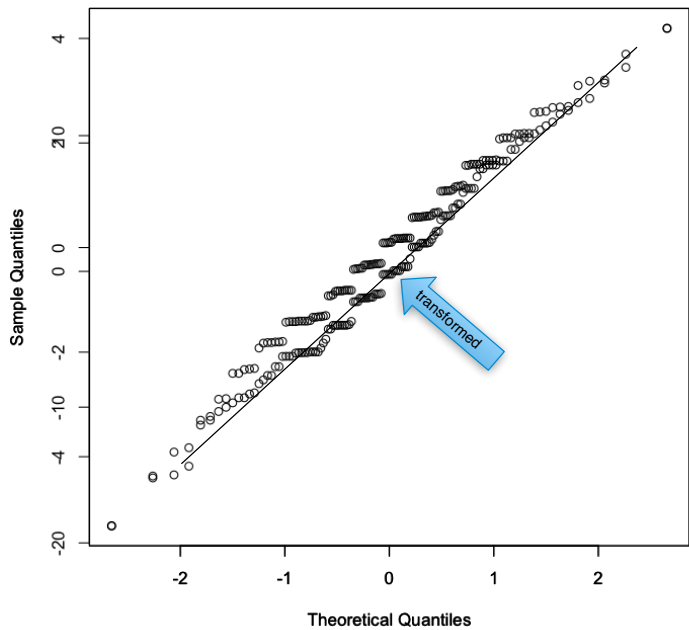
Emotional Symptoms – SDQ

Decision: Use transformed data

```
EmotsymBC <- (Emotsym^1.65 - 1)/1.65
```

```
outEmotsymBC <- lm(EmotsymBC ~ group + Emotsym_A)
```

Normal Q-Q Plot



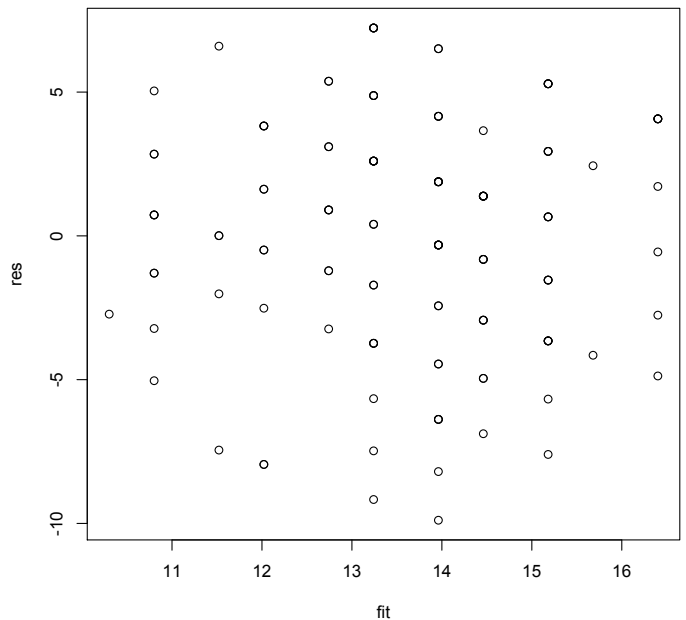
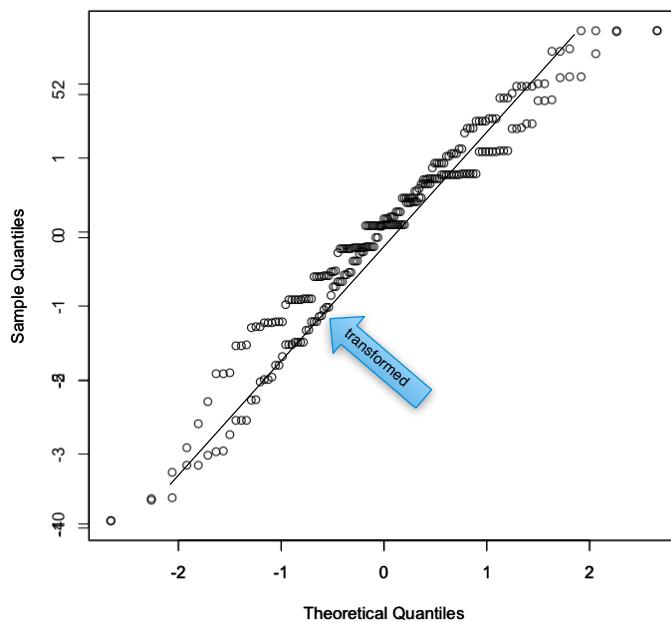
Hyperactivity – SDQ

Decision: Use transformed data

```
HyperactBC <- (Hyperact^1.35 - 1)/1.35
```

```
outHyperactBC <- lm(HyperactBC ~ group + Hyperact_A)
```

Normal Q-Q Plot

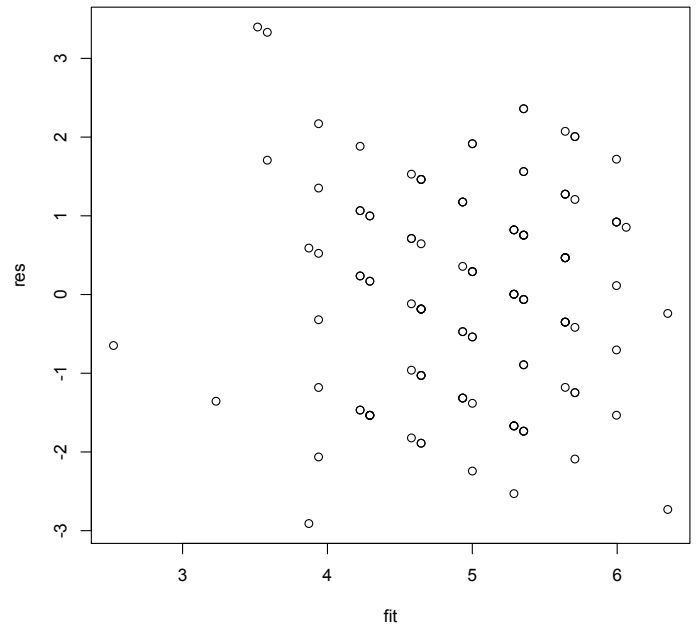
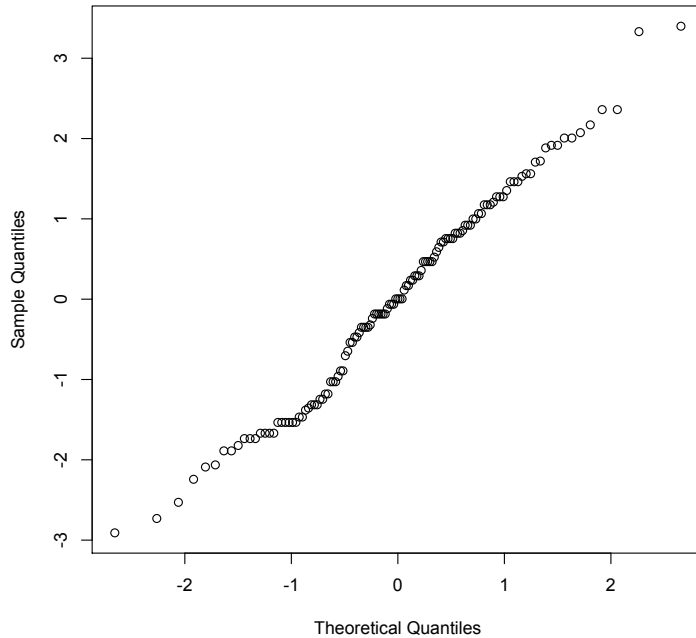


Peer Problems – SDQ

Decision: Original data acceptable

```
outPeerpr <- lm(Peerpr ~ group + Peerpr_A)
```

Normal Q-Q Plot



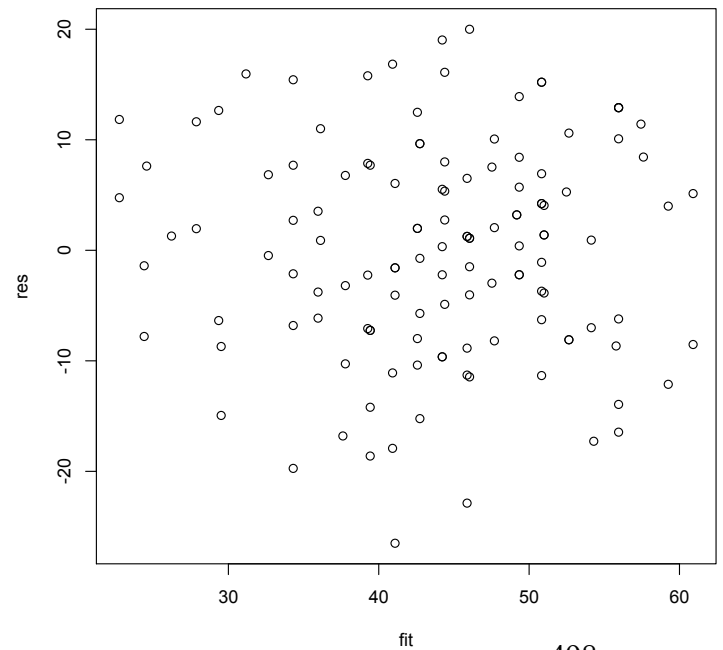
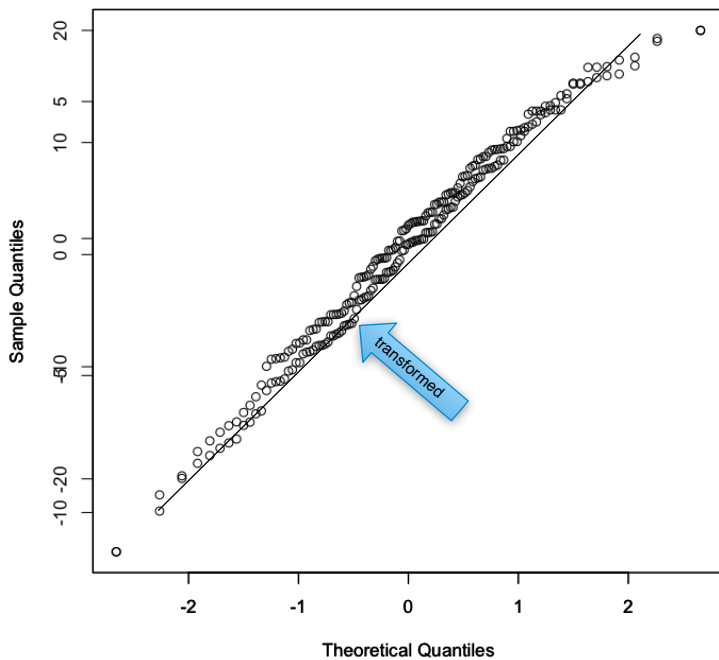
Overall Difficulties – SDQ

Decision: Use transformed data

```
DiffBC <- (SDQ_Diffic^1.3 - 1)/1.3
```

```
outDiffBC <- lm(DiffBC ~ group + SDQ_Diff_A)
```

Normal Q-Q Plot



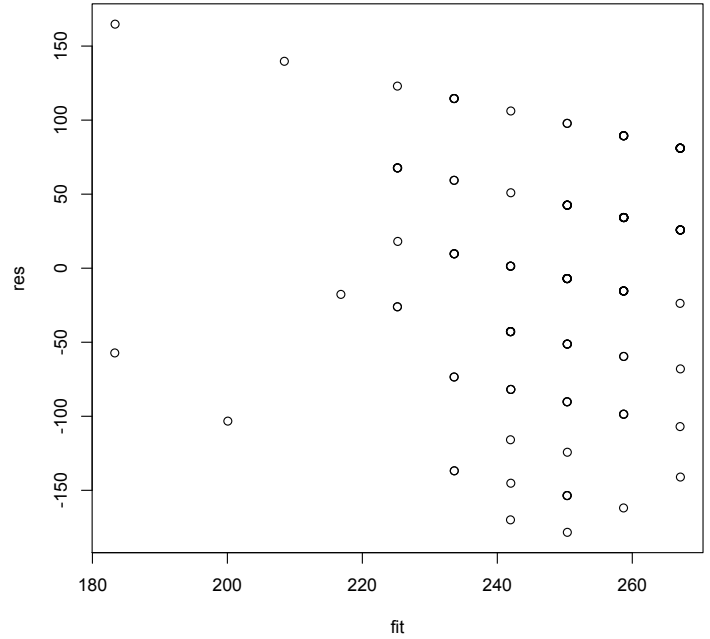
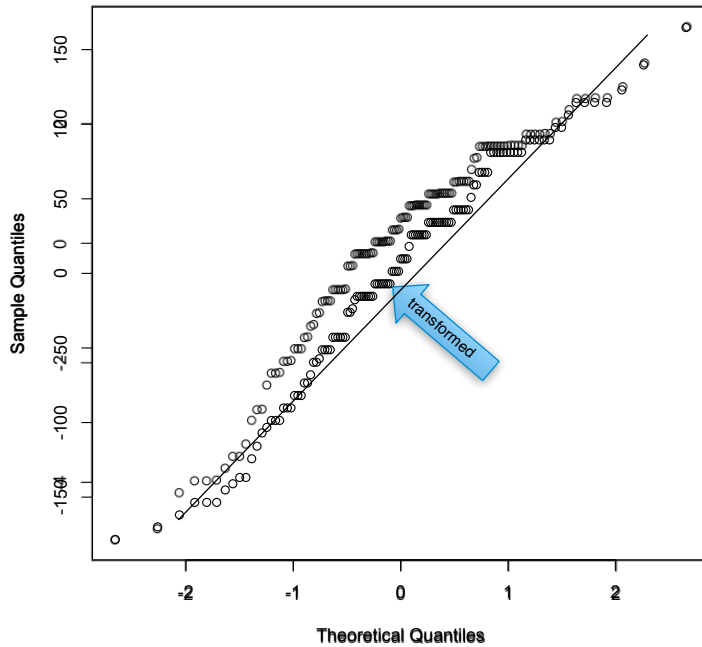
Prosooc – SDQ

Decision: Original and transformed data unacceptable; Mann-Whitney U test used.

```
ProsocBC <- (Prosoc^2.5 - 1)/2.5
```

```
outProsBC <- lm(ProsocBC ~ group + Prosoc_A)
```

Normal Q-Q Plot

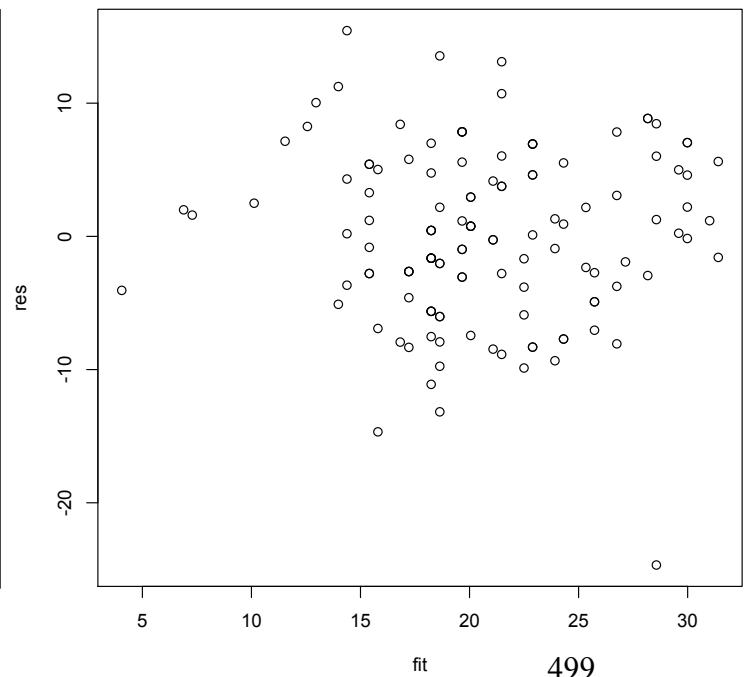
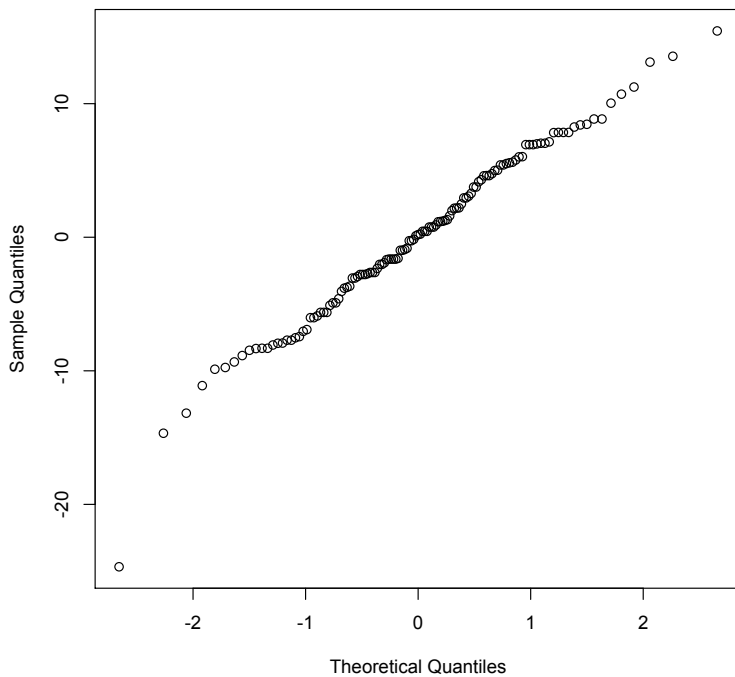


Soc Sup – Friends

Decision: Original data acceptable

```
outSSupFrd <- lm(SSupFrd ~ group + SSupFrd_A)
```

Normal Q-Q Plot

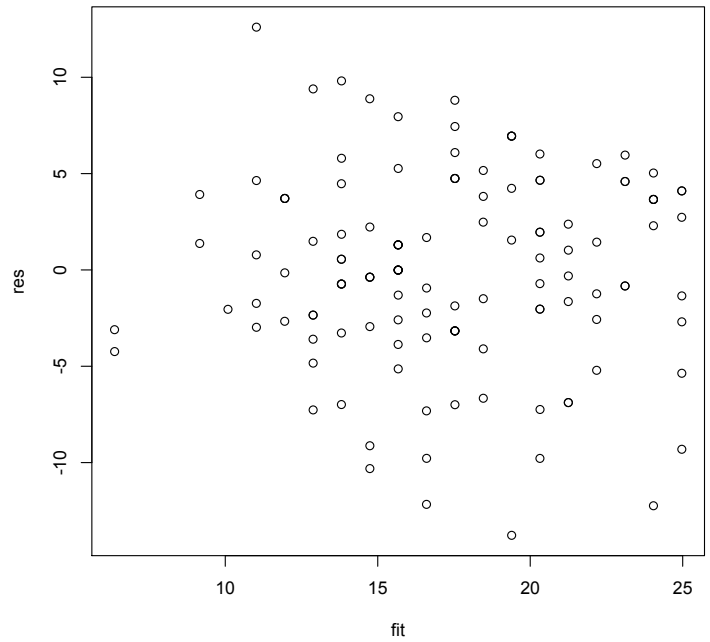
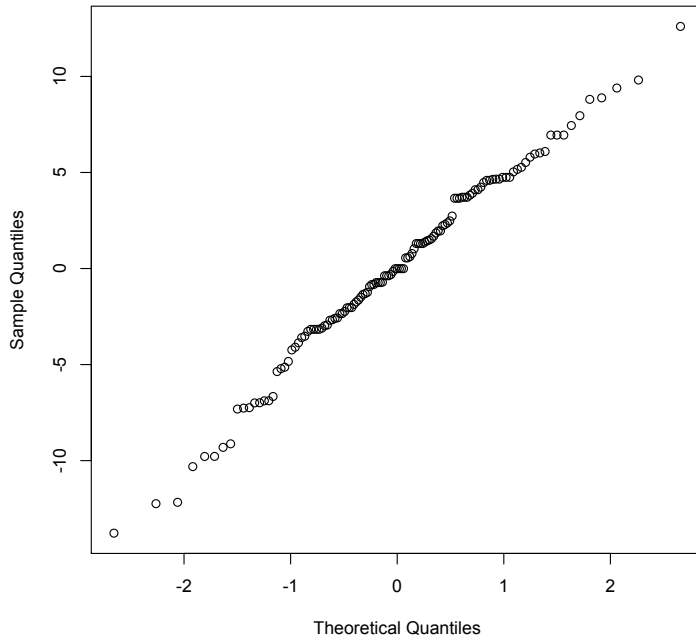


Soc Supp – Fam

Decision: Original data acceptable

```
outSSupFam <- lm(SSupFam ~ group + SSupFam_A + Age)
```

Normal Q-Q Plot



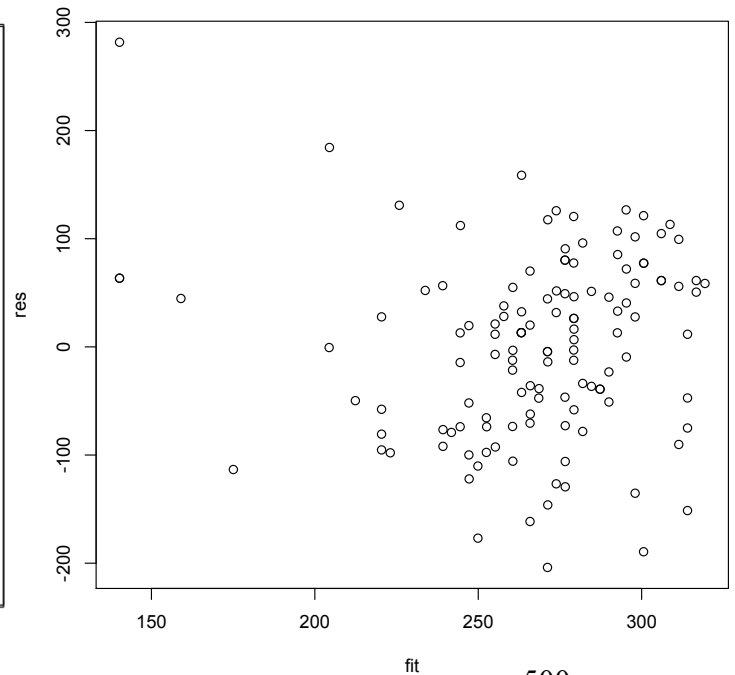
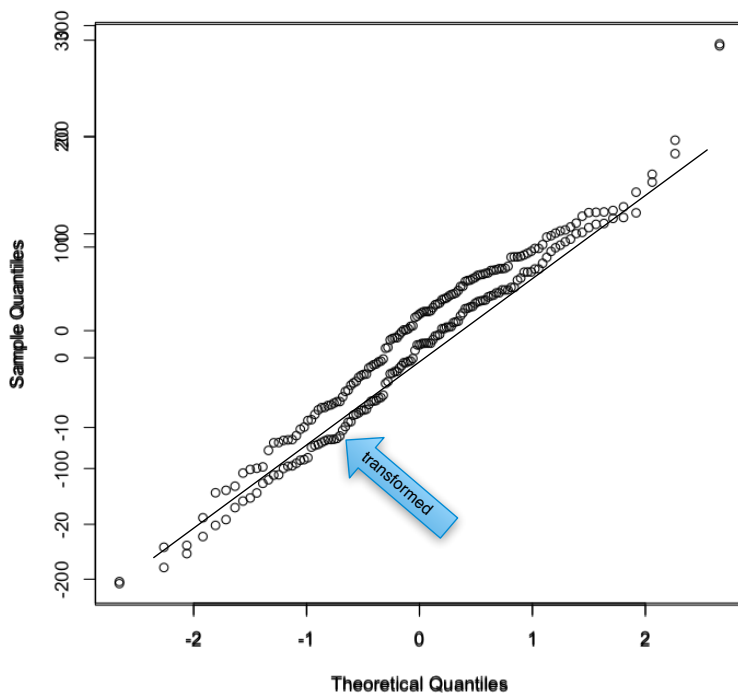
Social Skills

Decision: Use transformed data

```
SSkBC <- (SSk^1.59 - 1)/1.59
```

```
outSSkBC <- lm(SSkBC ~ group + SSk_A)
```

Normal Q-Q Plot



Appendix 5.4. Mann-Whitney U primary analysis statistics

The following tables include Mann-Whitney U test outputs on differences between intervention and control groups at posttest according to the study's outcome measures. Posttest scores are used given randomization and no statistically significant differences between groups at baseline. A significant difference between groups (intervention effect) is detected for one outcome measure: the Strength and Difficulties Questionnaire's conduct problems subscale.

Test Statistics^a

	Posttest GSE	Posttest Conn	Posttest SSupFrd	Posttest SSupFam	Posttest SocSk
Mann-Whitney U	1763.500	1933.000	1845.000	1946.500	1705.500
Wilcoxon W	3593.500	3763.000	3675.000	3776.500	3535.500
Z	-1.195	-.396	-.799	-.307	-1.472
Asymp. Sig. (2-tailed)	.232	.692	.424	.759	.141

a. Grouping Variable: group

Test Statistics^a

	Posttest Conduct	Posttest PeerPr	Posttest EmotSym	Posttest Hyperact	Posstest Difficulties
Mann-Whitney U	1504.500	1699.000	1712.500	1890.000	1731.000
Wilcoxon W	3334.500	3977.000	3542.500	3720.000	3561.000
Z	-2.472	-1.522	-1.453	-.603	-1.351
Asymp. Sig. (2-tailed)	.013	.128	.146	.546	.177

a. Grouping Variable: group

Test Statistics^a

	Posttest Prosoc
Mann-Whitney U	1953.000
Wilcoxon W	4231.000
Z	-.281
Asymp. Sig. (2-tailed)	.779

a. Grouping Variable: group

Appendix 5.5. Alternative analysis with peer problems item dropped

The SDQ peer problems subscale had one item that was particularly problematic both for in this study's scale reliability assessments as well as testing of the SDQ with Arab children in Gaza: 'I get a long better with older people than with people my own age.' This item could be viewed as a neutral indicator rather than positive or negative. As such, analyses were conducted again for the peer problems subscale and for the SDQ's overall difficulties scale with the item dropped, and the outputs are given below for both ANCOVA (using baseline scores as covariates) and the Mann-Whitney U test on differences between groups at posttest. The basic findings did not change—neither outcomes moved into statistical significance. Notably, however, the removal of the item did move the peer problems subscale further away from the prospect of a significant negative intervention effect, and it moved overall difficulties closer to the prospect of a significant positive intervention effect.

	Pretest (Mean, SD)		Posttest (Mean, SD)	
	Intervention	Control	Intervention	Control
Peer problems	4.6, 1.7	4.8, 1.7	4.6, 1.6	4.9, 1.6
Overall difficulties	21.8, 5.1	20.9, 4.9	22.3, 4.9	20.9, 5.1

Univariate Tests

Dependent Variable: Posttest PeerPr

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	1.690	1	1.690	.914	.341	.007
Error	229.299	124	1.849			

The F tests the effect of group. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Univariate Tests

Dependent Variable: Posttest Difficulties

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	18.772	1	18.772	1.376	.243	.011
Error	1692.060	124	13.646			

The F tests the effect of group. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Test Statistics^a

	Posttest PeerPr	Posstest Difficulties
Mann-Whitney U	1763.500	1690.500
Wilcoxon W	4041.500	3520.500
Z	-1.217	-1.546
Asymp. Sig. (2-tailed)	.224	.122

a. Grouping Variable: group

Appendix 5.6. Full subgroup analysis statistics

Age

This analysis compares differences in change scores between younger youth (13 to 15) and older youth (16 to 21) subgroups within the intervention group.

Group Statistics

Age Group		N	Mean	Std. Deviation	Std. Error Mean
ChgSc GSE	13-15	32	1.44	4.635	.819
	– 16-21	35	.03	3.535	.598
ChgSc Conn	13-15	32	.09	1.228	.217
	– 16-21	35	.23	1.003	.169
ChgSc SSupFrd	13-15	32	.50	3.654	.646
	– 16-21	35	-.20	3.795	.641
ChgSc SSupFam	13-15	32	-.09	4.402	.778
	– 16-21	35	-1.17	4.835	.817
ChgSc SocSk	13-15	32	3.56	13.884	2.454
	– 16-21	35	-1.29	8.298	1.403
ChgSc Prosoc	13-15	32	.22	2.599	.460
	– 16-21	35	.26	2.091	.353
ChgSc Conduct	13-15	32	.63	1.773	.314
	– 16-21	35	-.06	1.830	.309
ChgSc PeerPr	13-15	32	.19	2.007	.355
	– 16-21	35	-.17	1.855	.314
ChgSc EmotSym	13-15	32	1.09	2.277	.403
	– 16-21	35	-.09	1.853	.313
ChgSc Hyperact	13-15	32	-.25	1.368	.242
	– 16-21	35	-.17	1.382	.234
ChgSc Difficulties	13-15	32	1.66	3.989	.705
	– 16-21	35	-.49	4.415	.746

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
ChgSc GSE	Equal variances assumed	.991	.323	1.406	65	.164	1.409	1.002	-.592	3.410
	Equal variances not assumed			1.389	57.834	.170	1.409	1.014	-.621	3.439
ChgSc Conn	Equal variances assumed	3.186	.079	-.494	65	.623	-.135	.273	-.680	.410
	Equal variances not assumed			-.490	59.993	.626	-.135	.275	-.686	.416
ChgSc SSupFrd	Equal variances assumed	.218	.642	.768	65	.446	.700	.912	-1.121	2.521
	Equal variances not assumed			.769	64.816	.445	.700	.910	-1.118	2.518
ChgSc SSupFam	Equal variances assumed	.161	.690	.951	65	.345	1.078	1.133	-1.186	3.341
	Equal variances not assumed			.955	64.999	.343	1.078	1.129	-1.176	3.332
ChgSc SocSk	Equal variances assumed	3.262	.076	1.752	65	.084	4.848	2.767	-.677	10.374
	Equal variances not assumed			1.715	49.719	.093	4.848	2.827	-.831	10.527
ChgSc Prosoc	Equal variances assumed	.817	.369	-.067	65	.947	-.038	.574	-1.185	1.108
	Equal variances not assumed			-.066	59.535	.947	-.038	.580	-1.198	1.121
ChgSc Conduct	Equal variances assumed	.261	.611	1.547	65	.127	.682	.441	-.199	1.563
	Equal variances not assumed			1.549	64.770	.126	.682	.440	-.198	1.562
ChgSc PeerPr	Equal variances assumed	1.131	.292	.761	65	.450	.359	.472	-.583	1.301
	Equal variances not assumed			.758	63.187	.451	.359	.473	-.587	1.305
ChgSc EmotSym	Equal variances assumed	.892	.348	2.334	65	.023	1.179	.505	.170	2.189
	Equal variances not assumed			2.313	59.888	.024	1.179	.510	.159	2.200
ChgSc Hyperact	Equal variances assumed	.047	.829	-.234	65	.816	-.079	.336	-.750	.593
	Equal variances not assumed			-.234	64.582	.816	-.079	.336	-.750	.593
ChgSc Difficulties	Equal variances assumed	.256	.614	2.077	65	.042	2.142	1.031	.082	4.202
	Equal variances not assumed			2.086	64.993	.041	2.142	1.027	.091	4.192

Attendance

This analysis compares differences in change scores between high (two or more days weekly average) and low (less than two days weekly average) attendance subgroups within the intervention group.

Group Statistics

AdjAttendGrp		N	Mean	Std. Deviation	Std. Error Mean
ChgSc GSE	< 2 days/wk avg	39	.44	4.309	.690
	> 2 days/wk avg	28	1.07	3.906	.738
ChgSc Conn	< 2 days/wk avg	39	.15	1.182	.189
	> 2 days/wk avg	28	.18	1.020	.193
ChgSc SSupFrd	< 2 days/wk avg	39	-.56	4.044	.648
	> 2 days/wk avg	28	1.11	3.010	.569
ChgSc SSupFam	< 2 days/wk avg	39	-.59	4.070	.652
	> 2 days/wk avg	28	-.75	5.393	1.019
ChgSc SocSk	< 2 days/wk avg	39	1.26	12.981	2.079
	> 2 days/wk avg	28	.71	9.233	1.745
ChgSc Prosoc	< 2 days/wk avg	39	.18	2.553	.409
	> 2 days/wk avg	28	.32	2.019	.382
ChgSc Conduct	< 2 days/wk avg	39	.31	1.962	.314
	> 2 days/wk avg	28	.21	1.641	.310
ChgSc PeerPr	< 2 days/wk avg	39	-.15	2.122	.340
	> 2 days/wk avg	28	.21	1.618	.306
ChgSc EmotSym	< 2 days/wk avg	39	.56	2.210	.354
	> 2 days/wk avg	28	.36	2.059	.389
ChgSc Hyperact	< 2 days/wk avg	39	-.41	1.371	.220
	> 2 days/wk avg	28	.07	1.331	.252
ChgSc Difficulties	< 2 days/wk avg	39	.31	4.455	.713
	> 2 days/wk avg	28	.86	4.187	.791

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
ChgSc GSE	Equal variances assumed	.066	.798	-.619	65	.538	-.636	1.027	-2.687	1.416
	Equal variances not assumed			-.629	61.461	.532	-.636	1.010	-2.656	1.385
ChgSc Conn	Equal variances assumed	.960	.331	-.089	65	.929	-.025	.277	-.578	.528
	Equal variances not assumed			-.092	62.720	.927	-.025	.270	-.565	.515
ChgSc SSupFrd	Equal variances assumed	2.549	.115	-1.848	65	.069	-1.671	.904	-3.477	.135
	Equal variances not assumed			-1.939	64.891	.057	-1.671	.862	-3.393	.050
ChgSc SSupFam	Equal variances assumed	1.790	.186	.139	65	.890	.160	1.156	-2.148	2.468
	Equal variances not assumed			.132	47.904	.895	.160	1.210	-2.272	2.593
ChgSc SocSk	Equal variances assumed	.662	.419	.189	65	.851	.542	2.867	-5.183	6.267
	Equal variances not assumed			.200	64.999	.842	.542	2.714	-4.878	5.962
ChgSc Prosoc	Equal variances assumed	.206	.652	-.244	65	.808	-.142	.581	-1.303	1.019
	Equal variances not assumed			-.254	64.332	.800	-.142	.559	-1.259	.975
ChgSc Conduct	Equal variances assumed	1.132	.291	.205	65	.838	.093	.455	-.815	1.001
	Equal variances not assumed			.212	63.407	.833	.093	.441	-.789	.976
ChgSc PeerPr	Equal variances assumed	.952	.333	-.771	65	.444	-.368	.478	-1.322	.586
	Equal variances not assumed			-.805	64.723	.424	-.368	.457	-1.281	.545
ChgSc EmotSym	Equal variances assumed	.344	.559	.389	65	.699	.207	.532	-.856	1.270
	Equal variances not assumed			.394	60.657	.695	.207	.526	-.845	1.259
ChgSc Hyperact	Equal variances assumed	.617	.435	-1.435	65	.156	-.482	.336	-1.152	.188
	Equal variances not assumed			-1.442	59.331	.154	-.482	.334	-1.150	.186
ChgSc Difficulties	Equal variances assumed	.958	.331	-.510	65	.612	-.549	1.077	-2.699	1.601
	Equal variances not assumed			-.516	60.383	.608	-.549	1.065	-2.680	1.582

Implementation

The following outputs were produced from independent samples t-tests used to assess intervention effects according to differences in change scores between control and intervention groups for the male QS NFE center with the highest empowerment survey scores and the male QS NFE center with lowest empowerment survey scores as separate analyses.

QS NFE center with higher rated implementation

Group Statistics

group		N	Mean	Std. Deviation	Std. Error Mean
ChgSc GSE	control	17	.18	3.340	.810
	intervention	19	.37	2.385	.547
ChgSc Conn	control	17	.24	.903	.219
	intervention	19	.05	1.079	.247
ChgSc SSupFrd	control	17	.94	2.193	.532
	intervention	19	1.00	3.543	.813
ChgSc SSupFam	control	17	.12	2.891	.701
	intervention	19	.37	3.975	.912
ChgSc SocSk	control	17	-1.53	5.843	1.417
	intervention	19	.58	5.571	1.278
ChgSc Prosoc	control	17	-.47	1.841	.447
	intervention	19	.68	1.668	.383
ChgSc Conduct	control	17	.18	2.481	.602
	intervention	19	.42	1.710	.392
ChgSc PeerPr	control	17	.29	1.160	.281
	intervention	19	.74	1.695	.389
ChgSc EmotSym	control	17	.29	1.993	.483
	intervention	19	1.16	1.708	.392
ChgSc Hyperact	control	17	-.18	1.380	.335
	intervention	19	.32	1.336	.306
ChgSc Difficulties	control	17	.59	4.214	1.022
	intervention	19	2.63	3.370	.773

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
ChgSc GSE	Equal variances assumed	.533	.470	-.200	34	.843	-.192	.960	-2.142	1.758
	Equal variances not assumed			-.196	28.636	.846	-.192	.978	-2.192	1.808
ChgSc Conn	Equal variances assumed	.492	.488	.547	34	.588	.183	.334	-.496	.861
	Equal variances not assumed			.553	33.868	.584	.183	.331	-.489	.855
ChgSc SSUpFrd	Equal variances assumed	2.931	.096	-.059	34	.953	-.059	.997	-2.084	1.966
	Equal variances not assumed			-.061	30.435	.952	-.059	.971	-2.042	1.924
ChgSc SSUpFam	Equal variances assumed	.413	.525	-.214	34	.832	-.251	1.171	-2.630	2.129
	Equal variances not assumed			-.218	32.713	.829	-.251	1.150	-2.592	2.091
ChgSc SocSk	Equal variances assumed	.092	.763	-1.108	34	.276	-2.108	1.903	-5.976	1.759
	Equal variances not assumed			-1.105	33.130	.277	-2.108	1.908	-5.990	1.774
ChgSc Prosoc	Equal variances assumed	.284	.597	-1.975	34	.056	-1.155	.585	-2.343	.034
	Equal variances not assumed			-1.963	32.536	.058	-1.155	.588	-2.352	.042
ChgSc Conduct	Equal variances assumed	1.615	.212	-.348	34	.730	-.245	.704	-1.675	1.186
	Equal variances not assumed			-.341	27.997	.736	-.245	.718	-1.716	1.227
ChgSc PeerPr	Equal variances assumed	2.337	.136	-.904	34	.373	-.443	.490	-1.438	.553
	Equal variances not assumed			-.923	31.937	.363	-.443	.480	-1.420	.535
ChgSc EmotSym	Equal variances assumed	1.926	.174	-1.400	34	.170	-.864	.617	-2.117	.390
	Equal variances not assumed			-1.388	31.755	.175	-.864	.622	-2.132	.404
ChgSc Hyperact	Equal variances assumed	.014	.908	-1.087	34	.285	-.492	.453	-1.413	.428
	Equal variances not assumed			-1.085	33.279	.286	-.492	.454	-1.415	.431
ChgSc Difficulties	Equal variances assumed	1.112	.299	-1.615	34	.116	-2.043	1.265	-4.615	.528
	Equal variances not assumed			-1.594	30.634	.121	-2.043	1.282	-4.658	.572



QS NFE center with lower rated implementation

Group Statistics

group		N	Mean	Std. Deviation	Std. Error Mean
ChgSc GSE	control	15	2.60	4.356	1.125
	intervention	16	.88	3.649	.912
ChgSc Conn	control	15	.07	.458	.118
	intervention	16	-.44	1.031	.258
ChgSc SSupFrd	control	15	-.07	3.305	.853
	intervention	16	-.69	3.092	.773
ChgSc SSupFam	control	15	-.53	3.270	.844
	intervention	16	.13	3.757	.939
ChgSc SocSk	control	15	8.47	22.277	5.752
	intervention	16	.63	13.145	3.286
ChgSc Prosoc	control	15	.80	3.121	.806
	intervention	16	-.06	2.792	.698
ChgSc Conduct	control	15	-.47	3.248	.839
	intervention	16	.38	1.746	.437
ChgSc PeerPr	control	15	.53	1.922	.496
	intervention	16	.06	2.351	.588
ChgSc EmotSym	control	15	-.27	1.870	.483
	intervention	16	.88	2.553	.638
ChgSc Hyperact	control	15	.60	1.682	.434
	intervention	16	-.38	1.204	.301
ChgSc Difficulties	control	15	.40	3.719	.960
	intervention	16	.94	4.281	1.070

Adult Connectedness^b

	ChgSc Conn
Mann-Whitney U	80.500
Wilcoxon W	216.500
Z	-1.681
Asymp. Sig. (2-tailed)	.093
Exact Sig. [2*(1-tailed Sig.)]	.119 ^a

a. Not corrected for ties.

b. Grouping Variable: group

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
ChgSc GSE	Equal variances assumed	.814	.374	1.198	29	.241	1.725	1.440	-1.219	4.669
	Equal variances not assumed			1.191	27.410	.244	1.725	1.448	-1.244	4.694
ChgSc Conn	Equal variances assumed	16.489	.000	1.739	29	.093	.504	.290	-.089	1.097
	Equal variances not assumed			1.778	20.979	.090	.504	.284	-.085	1.094
ChgSc SSupFrd	Equal variances assumed	.000	.987	.540	29	.593	.621	1.149	-1.729	2.971
	Equal variances not assumed			.539	28.495	.594	.621	1.151	-1.736	2.978
ChgSc SSupFam	Equal variances assumed	.660	.423	-.519	29	.608	-.658	1.269	-3.253	1.937
	Equal variances not assumed			-.521	28.852	.606	-.658	1.263	-3.242	1.925
ChgSc SocSk	Equal variances assumed	2.813	.104	1.203	29	.239	7.842	6.518	-5.490	21.173
	Equal variances not assumed			1.184	22.403	.249	7.842	6.624	-5.882	21.566
ChgSc Prosoc	Equal variances assumed	.059	.810	.812	29	.423	.863	1.062	-1.310	3.035
	Equal variances not assumed			.809	28.116	.425	.863	1.066	-1.321	3.046
ChgSc Conduct	Equal variances assumed	2.825	.104	-.907	29	.372	-.842	.928	-2.740	1.057
	Equal variances not assumed			-.890	21.165	.383	-.842	.946	-2.807	1.124
ChgSc PeerPr	Equal variances assumed	.074	.788	.608	29	.548	.471	.774	-1.113	2.055
	Equal variances not assumed			.612	28.494	.545	.471	.769	-1.104	2.046
ChgSc EmotSym	Equal variances assumed	.398	.533	-1.412	29	.168	-1.142	.808	-2.795	.511
	Equal variances not assumed			-1.427	27.447	.165	-1.142	.800	-2.782	.499
ChgSc Hyperact	Equal variances assumed	2.788	.106	1.865	29	.072	.975	.523	-.094	2.044
	Equal variances not assumed			1.845	25.247	.077	.975	.528	-.113	2.063
ChgSc Difficulties	Equal variances assumed	.710	.406	-.372	29	.713	-.538	1.445	-3.492	2.417
	Equal variances not assumed			-.374	28.843	.711	-.538	1.438	-3.479	2.404

Higher vs. lower implementation

This analysis compares change scores between the intervention group of the male QS NFE center with the highest implementation scores and the intervention group of the male QS NFE center with the lowest implementation scores.

Group Statistics

NFE Cent		N	Mean	Std. Deviation	Std. Error Mean
ChgSc GSE	low impl.	19	.37	2.385	.547
	high impl.	16	.88	3.649	.912
ChgSc Conn	low impl.	19	.05	1.079	.247
	high impl.	16	-.44	1.031	.258
ChgSc SSupFrd	low impl.	19	1.00	3.543	.813
	high impl.	16	-.69	3.092	.773
ChgSc SSupFam	low impl.	19	.37	3.975	.912
	high impl.	16	.13	3.757	.939
ChgSc SocSk	low impl.	19	.58	5.571	1.278
	high impl.	16	.63	13.145	3.286
ChgSc Prosoc	low impl.	19	.68	1.668	.383
	high impl.	16	-.06	2.792	.698
ChgSc Conduct	low impl.	19	.42	1.710	.392
	high impl.	16	.38	1.746	.437
ChgSc PeerPr	low impl.	19	.74	1.695	.389
	high impl.	16	.06	2.351	.588
ChgSc EmotSym	low impl.	19	1.16	1.708	.392
	high impl.	16	.88	2.553	.638
ChgSc Hyperact	low impl.	19	.32	1.336	.306
	high impl.	16	-.38	1.204	.301
ChgSc Difficulties	low impl.	19	2.63	3.370	.773
	high impl.	16	.94	4.281	1.070

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
ChgSc GSE	Equal variances assumed	1.395	.246	-.493	33	.625	-.507	1.027	-2.596	1.582
	Equal variances not assumed			-.476	25.036	.638	-.507	1.064	-2.697	1.684
ChgSc Conn	Equal variances assumed	.131	.720	1.366	33	.181	.490	.359	-.240	1.220
	Equal variances not assumed			1.372	32.436	.180	.490	.357	-.237	1.218
ChgSc SSupFrd	Equal variances assumed	.296	.590	1.486	33	.147	1.688	1.135	-.622	3.997
	Equal variances not assumed			1.504	32.945	.142	1.688	1.122	-.595	3.970
ChgSc SSupFam	Equal variances assumed	.025	.876	.185	33	.854	.243	1.316	-2.433	2.920
	Equal variances not assumed			.186	32.524	.854	.243	1.309	-2.422	2.908
ChgSc SocSk	Equal variances assumed	1.952	.172	-.014	33	.989	-.046	3.315	-6.791	6.699
	Equal variances not assumed			-.013	19.509	.990	-.046	3.526	-7.413	7.321
ChgSc Prosoc	Equal variances assumed	1.582	.217	.978	33	.335	.747	.763	-.806	2.300
	Equal variances not assumed			.938	23.599	.358	.747	.796	-.898	2.391
ChgSc Conduct	Equal variances assumed	.026	.874	.079	33	.938	.046	.586	-1.146	1.238
	Equal variances not assumed			.078	31.751	.938	.046	.587	-1.150	1.242
ChgSc PeerPr	Equal variances assumed	.644	.428	.984	33	.332	.674	.685	-.720	2.069
	Equal variances not assumed			.957	26.728	.347	.674	.705	-.772	2.121
ChgSc EmotSym	Equal variances assumed	1.675	.205	.391	33	.699	.283	.724	-1.190	1.756
	Equal variances not assumed			.378	25.432	.709	.283	.749	-1.258	1.824
ChgSc Hyperact	Equal variances assumed	.146	.705	1.594	33	.121	.691	.433	-.191	1.573
	Equal variances not assumed			1.608	32.822	.117	.691	.430	-.183	1.565
ChgSc Difficulties	Equal variances assumed	.855	.362	1.310	33	.199	1.694	1.293	-.937	4.325
	Equal variances not assumed			1.283	28.313	.210	1.694	1.320	-1.009	4.397

Appendix 5.7. Moderator analysis results for age

Moderator: Age group (dichotomous: 13-15, 16-21)

	Outcome (change scores)											
	General self-efficacy			Social skills*			Social supports - friends			Social supports - family		
	β	ΔR^2	p	β	ΔR^2	p	B	ΔR^2	p	β	ΔR^2	p
Step 1		.008	.59		.008	.59		.005	.74		.016	.36
Experimental group	.062		.93	.491		.83	-.141		.81	-.292		.68
Age group (moderator)	-.307		.31	-.887		.34	-.188		.44	-.421		.16
Step 2		.007	.34		.013	.19		.001	.66		.000	.96
Experimental group	.083		.91	.579		.79	-.133		.82	-.291		.69
Age group (moderator)	.025		.95	.479		.73	-.068		.85	-.406		.37
Interaction term	-.591		.34	-2.42		.19	-.213		.66	-.027		.96

	SDQ <u>prosocial</u>			SDQ conduct problems			SDQ emotional symptoms*			SDQ peer problems			SDQ hyperactivity		
	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p
Step 1		.007	.64		.033	.12		.027	.18		.008	.62		.014	.42
Experimental group	.276		.51	.461		.23	.195		.59	-.259		.45	-.184		.48
Age group (moderator)	.135		.44	-.227		.14	-.247		.10	-.107		.45	.103		.33
Step 2		.005	.43		.001	.73		.024	.08		.001	.76		.005	.44
Experimental group	.286		.50	.465		.22	.214		.55	-.256		.46	-.178		.49
Age group (moderator)	.288		.27	-.167		.48	.046		.83	-.058		.78	.195		.22
Interaction term	-.273		.43	-.107		.73	-.520		.08	-.086		.76	-.164		.44

	SDQ overall difficulties*		
	β	ΔR^2	p
Step 1		.021	.27
Experimental group	.213		.78
Age group (moderator)	-.477		.13
Step 2		.015	.16
Experimental group	.245		.74
Age group (moderator)	.016		.97
Interaction term	-.877		.16

* Effect size $\geq .10$

Coding: Experimental group dummy codes (0, 1); Age group weighted effects codes (-1.49, 1)

Moderator: Age (continuous, range 13-21)

	Outcome (change scores)														
	General self-efficacy			Social skills*			Social supports - friends			Social supports - family			Local adult connectedness		
	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p
Step 1		.002	.87		.009	.57		.002	.57		.021	.26		.027	.18
Experimental group	.162		.82	.70		.75	-.086		.75	-.200		.78	.230		.19
Age group (moderator)	-.107		.64	-.687		.32	-.095		.32	-.362		.10	.076		.16
Step 2		.005	.41		.018	.13		.001	.13		.000	.90		.007	.33
Experimental group	.165		.82	.719		.74	-.085		.74	-.199		.78	.229		.19
Age group (moderator)	.093		.78	.438		.66	-.021		.66	-.332		.31	.020		.80
Interaction term	-.376		.41	-.211		.13	-.139		.13	-.056		.90	.106		.33

	SDQ prosocial			SDQ conduct problems*			SDQ emotional symptoms			SDQ peer problems			SDQ hyperactivity*		
	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p	β	ΔR^2	p
Step 1		.005	.74		.070	.01		.031	.13		.015	.39		.017	.33
Experimental group	.237		.57	.488		.19	.252		.48	-.244		.47	-.206		.42
Age group (moderator)	.073		.57	-.305		.00	-.202		.07	-.127		.23	.094		.23
Step 2		.009	.29		.012	.19		.009	.29		.000	.97		.010	.27
Experimental group	.240		.57	.485		.19	.254		.48	-.244		.48	-.204		.42
Age group (moderator)	.219		.25	-.463		.00	-.079		.62	-.130		.40	.187		.10
Interaction term	-.275		.29	.296		.19	-.232		.29	.006		.97	-.174		.27

	SDQ overall difficulties		
	β	ΔR^2	p
Step 1		.045	.05
Experimental group	.290		.69
Age group (moderator)	-.540		.02
Step 2		.000	.82
Experimental group	.291		.69
Age group (moderator)	-.485		.15
Interaction term	-.105		.82

* Effect size $\geq .10$

Appendix 6.1. Qualitative interview informed consent form

Informed Consent

Non-Formal Education Interviews



Thank you for considering allowing us to learn from you!

Purpose: The purpose of this research is to better understand the non-formal education program and the perspectives of its stakeholders, including partners, funders, members of the community, program staff, and the youth themselves. Stakeholders can give insight into the contextual factors that influence the methodology and implementation of initiatives like non-formal education. Your contributions will help inform best practices in expanding social and educational programs for disadvantaged youth. This research is a partnership between Oxford University, Questscope, and the Ministry of Education.

Process: The research involves an interview with a researcher and translator. If you choose to participate, we would ask several questions. The interview generally lasts around an hour, but the time can vary.

Compensation: There are no additional benefits to participating in this interview other than contributing to a better understanding of the program and youth development.

Confidentiality: If you do choose to participate, the records of this research will be kept private. In any publications based on this, we will not include any information that will make it possible to identify any participants. Research records will be kept in a locked file, and access will be limited to myself and my research staff.

Voluntary: If you do decide to participate, your participation is completely voluntary. I want to assure you that if you decide not to answer a particular question or wish to discontinue the interview altogether at any point, that is completely fine.

Recording: If you choose to allow it, the interview will be recorded using a voice recorder or video camera. The recordings help the researchers capture your feedback accurately.

Questions: If you have any questions about the study please feel free to contact me or the Questscope research coordinator. All of our information is on the cards given to you.

Many thanks,

Matthew Morton
University of Oxford

Do I want to take part?

I have read and understand this information and have had the chance to ask questions.

I agree to take part in the research.

Print Name: _____

Signature: _____ Date: _____

Appendix 6.2. Qualitative interview guides

QS NFE Facilitator Interview Guide

Personal

1. What are your job title and main responsibilities?
2. When did you become an NFE facilitator?
 - a. And when did you become a teacher?
3. Why did you become a facilitator?

Staff support and empowerment

4. Could you describe the kinds and quality of support you receive?
 - a. From MOE?
 - b. From the school?
 - c. From other facilitators?
 - d. From Questscope?
 - e. From the community?
5. In what areas could you use more training or support?
 - a. What could Questscope do differently to support facilitators?
6. How would you describe collaboration or teamwork between facilitators?
 - a. What are the factors that facilitate or hinder this collaboration?
7. To what extent do you feel like you can have an influence in NFE decision-making or the way NFE programs are implemented?

Influence and implementation

8. Who (or what) do you look to for guidance or ideas in working with youth?
9. (How) has NFE changed since you first got involved with it?
 - a. What factors have influenced the change(s)?

Youth approach

10. Describe the relationship between facilitators and NFE youth.
11. (How) are youth involved in decision-making in NFE planning or implementation?
 - a. Could you give any examples?
 - b. Do you consider this adequate?
 - c. [Interviewer follow-up note:] Do you think youth should have more voice over the way the NFE program or activities are run? (If so,) in what ways?

Impacts

12. How have *you* changed since starting with NFE?
 - a. (How) has your involvement in NFE changed the way you work in formal education?

13. (How) does NFE strengthen family involvement and support?

QS NFE general

14. Could you tell me about a favorite memory with the NFE program?
 - a. Was there a most challenging memory or experience for you?
15. What are the challenges for recruiting Jordanian youth into NFE?
16. What recruitment approaches/techniques have you found most effective?
17. How can NFE be improved?
18. Are there any other comments you would like to make?

QS NFE 1st Cycle Youth Interview Guide

1. What is your age?
2. How long have you been in the NFE program?
3. Could you describe how your first four months of NFE have gone?
4. Could you describe the interaction (or relationship) between the facilitators and students?
 - a. How do you think that interaction could have been different?
5. Is the environment of the group/classroom comfortable? (How?)
6. What's your favorite thing about NFE?
 - a. What's your least favorite
7. Tell me about your relationships with youth in the program.
 - a. And were your relationships with youth your age different before NFE?
8. (How) has your participation in the program affected your family relationships?
9. Do you feel like your voice or opinion matters in NFE?
 - a. In what ways do you feel like you have an influence in NFE?
 - b. Do you think youth should have more voice in the way the NFE program is run? (If so,) in what ways?
10. What would you change about the program, or add to it, if you were an NFE facilitator?
11. Do you think you have changed at all because of NFE?
 - a. (If so,) in what ways? (e.g., skills, behaviors, attitudes, relationships?)
12. What are the most important benefits you get out of the program?
13. Is there anything else you would like to share?

Implementing Organization's Executive Director Interview Guide

1. How would you describe your role in Questscope?
2. How did the NFE program get started?
3. Who (or what) do you look to for guidance or ideas in working with youth?
4. What is Questscope's role in the NFE program?
 - a. What is the role of MOE in the NFE program?
5. Could you describe the program methodology and key components?

- a. How is youth empowerment or participation emphasized in the NFE program?
6. Could you describe the supports and training facilitators receive from Questscope?
7. In what ways is NFE distinct from formal education?
8. Have you noticed changes in the NFE program or implementation of the program since it began?
 - a. What were the causes of those changes?
9. To what extent are you satisfied with the current implementation of NFE programs?
 - a. Are there specific areas you think it could be strengthened?
10. Are there ways that Questscope and MOE can strengthen collaboration together for the benefit of the program?
11. There have been many changes inside of Questscope over the last couple of years. Could you describe how these might have affected the NFE program?
12. What have been the main challenges for Questscope in its quality assurance role with NFE?
13. What is the role of Education Coordinators working with NFE centers?
 - a. Are you satisfied with this role?
14. Would you describe the procedures and expectations put on facilitators as too little, too much, or about right? Could you explain?
15. (How) has the nature of funding for NFE changed over the years?
 - a. What effects does the nature of funding have on the implementation of the NFE program?
16. Do you have a best memory/experience working with the NFE program?
17. What do you think other countries in the Middle East and North Africa could learn from the NFE initiative or experience?
18. What dreams or goals do you have for the program in the months or years to come?
19. Is there anything else you would like to share?

Implementing Organization's Technical Officer Interview Guide

1. How would you describe your role in Questscope?
2. How did the NFE program get started?
3. What is Questscope's role in the NFE program?
 - a. What is the role of MOE in the NFE program?
4. Could you describe the program methodology and key components?
 - a. How is youth empowerment or participation emphasized in the NFE program?
5. In what ways is NFE distinct from formal education?
6. Could you describe the supports and training facilitators receive from Questscope?
7. How was this training developed?
 - a. (How) has the training changed over the years?

8. How would you describe the current quality and amount of implementation of training and coaching for NFE facilitators?
9. To what extent are you satisfied with the current implementation of NFE programs?
 - a. Are there specific areas you think it could be strengthened?
10. Are there ways that Questscope and MOE can strengthen collaboration together for the benefit of the program?
11. There have been many changes inside of Questscope over the last couple of years. Could you describe how these might have affected the NFE program?
12. What have been the main challenges for Questscope in its quality assurance role with NFE?
13. What is the role of Field Coordinators working with NFE centers?
 - a. Are you satisfied with this role?
14. Would you describe the procedures and expectations put on facilitators as too little, too much, or about right? Could you explain?
15. (How) has the nature of funding for NFE changed over the years?
 - a. What effects does the nature of funding have on the implementation of the NFE program?
16. Do you have a best memory/experience working with the NFE program?
17. What do you think other countries in the Middle East and North Africa could learn from the NFE initiative or experience?
18. What dreams or goals do you have for the program in the months or years to come?
19. Is there anything else you would like to share?

Implementing Organization's Field Coordinator Interview Guide

1. How would you describe your role in Questscope?
2. Who (or what) do you look to for guidance or ideas in working with youth?
3. What is Questscope's role in the NFE program?
 - a. What is the role of MOE in the NFE program?
4. Could you describe the program methodology and key components?
 - a. How is youth empowerment or participation emphasized in the NFE program?
 - b. Do you see opportunities for increasing youth participation in QS or NFE?
5. In what ways is NFE distinct from formal education?
6. Could you describe the supports and training *facilitators* receive from Questscope?
7. Could you describe the supports and training that *you* receive from Questscope?
 - a. What kind of training or professional development do you feel you could use more of?
8. To what extent are you satisfied with the current implementation of NFE programs?
 - a. Are there specific areas you think it could be strengthened?

9. To what extent do you feel like *you* can have an influence in decision-making concerning NFE or the way NFE programs are implemented?
10. There have been many changes inside of Questscope over the last couple of years. Could you describe how these might have affected the NFE program?
11. How would describe collaboration between facilitators and Education Coordinators generally?
 - a. How would you describe collaboration between Education Coordinators generally?
12. What have been the main challenges for Questscope in its quality assurance role with NFE?
13. Would you describe the procedures and expectations put on facilitators as too little, too much, or about right? Could you explain?
14. Have you noticed any best practices or promising innovations in the NFE centers? For example, with improving program quality or recruitment?
15. Do you have a best memory/experience working with the NFE program?
 - a. What about a most challenging memory/experience
16. Is there anything else you would like to share?

Ministry of Education Interview Guide

1. What is your title and role with the MOE?
2. How long have you been with the MOE?
3. When did you start working with the dropouts program?
 - a. Could you talk about your role in the history of the program?
4. Could you describe how you see MOE's role with the NFE program?
 - a. How would you describe Questscope's role?
5. How would you describe the program methodology and key components?
 - a. Is it different from formal education methodology? (how?)
6. Have you noticed changes in the NFE methodology or implementation of the program since it began?
 - a. What were the causes of those changes?
7. What do you think are the main challenges for facilitators in implementing the NFE program?
 - a. Are there ways that MOE can address or alleviate these challenges?
8. Would you describe the procedures and expectations put on facilitators as too little, too much, or about right? Could you explain?
9. To what extent are you satisfied with the current implementation of NFE programs?
 - a. Are there specific areas you think it could be strengthened?
10. Are there ways that Questscope and MOE can strengthen collaboration together for the benefit of the program?
11. Could you talk about the situation of funding for NFE?
12. What do you think other countries in the Middle East and North Africa could learn from the NFE program?
13. Do you have a best memory/experience working with the NFE program?

14. What dreams or goals do you have for the program in the months or years to come?
 - a. Does MOE have hopes of expanding the program to other youth and communities?
15. Is there anything else you would like to share?

Community-Based Organization Interview Guide

1. Could you describe what this CBO does and how it got started?
2. What is your title and role with the CBO?
 - a. How long have you been with the CBO?
3. When and how did Questscope and the CBO start working together?
4. How would you describe the informal education (IFE) program?
5. Why might some youth come to IFE before going to NFE rather than doing all of their 3 cycles in the school with NFE?
6. Could you describe the kinds and quality of support you receive?
 - a. From MOE?
 - b. From the school?
 - c. From Questscope?
 - d. From the community?
7. In what areas could IFE facilitators use more support or training?
 - a. What could Questscope do differently to support IFE facilitators?
8. How much collaboration, if any, is there between your CBO and other CBOs or schools concerning the IFE program?
9. (How) has IFE changed since you first got involved with it?
 - a. Has the relationship with Questscope changed at all?
10. Describe the relationship between facilitators and IFE youth.
11. (How) are youth involved in decision-making in IFE planning or implementation?
 - a. Could you give any examples?
 - b. Do you consider this adequate?
 - c. [Interviewer follow-up note:] Do you think youth should have more voice over the way the IFE program or activities are run? (If so,) in what ways?
12. How have *you* changed since starting with IFE/NFE?
13. (How) does IFE strengthen family involvement and support?
14. Could you tell me about a favorite memory with the IFE program?
 - a. Was there a most challenging memory or experience for you?
15. What are the challenges for recruiting Jordanian youth into IFE?
 - a. What recruitment approaches/techniques have you found most effective?
16. How can IFE be improved?
17. Are there any other comments you would like to make?

Funder Interview Guide

1. What is your title and role with UNHCR?

2. How long have you been with UNHCR?
3. When did UNHCR start supporting the NFE program?
 - a. Could you talk about UNHCR's role in the history and development of the NFE program?
4. How would you describe the program methodology and key components of NFE?
 - a. Is it different from formal education methodology? (how?)
5. What are the supports/resources that UNHCR gives, financially or otherwise, to the NFE program?
6. Could you discuss UNHCR's approach as a funder in supporting programs like NFE?
 - a. (e.g., is it hands off? Actively involved with the grantee and the project? Is there much contact?)
7. Why did UNHCR support the NFE program?
8. What effects, positive or negative, do you think UNHCR funding might have on the *quality* of NFE program implementation?
9. Do you have a sense of whether the support UNHCR provides for educating out-of-school Iraqis has any impact, positive or negative, on the program's work with out-of-school Jordanians?
10. What are the results that UNHCR expects from the NFE program?
 - a. How are these defined and assessed?
11. What do you think other countries in Middle East and North Africa could learn from the NFE program?
12. What goals do you think the partners of NFE should have for the program in the months and years to come?
13. Is there anything else you would like to share?

Youth Advisory Council Youth Interview Guide

1. What is your age?
2. How long have you been in the NFE program? (If graduated, when did you graduate?)
3. Why did you join NFE?
4. How would you describe the NFE program to other youth?
5. Could you describe the interaction (or relationship) between the facilitators and students?
 - a. How do you think that interaction could have been even better?
6. Did you have a favorite facilitator (one you connected with the best)?
 - a. What were the qualities that made him/her your favorite?
7. Is the environment of the classroom/group comfortable? (How?)
8. What's your favorite thing about NFE?
 - a. What's your least favorite?
9. Tell me about your relationships with youth in the program.
 - a. And were your relationships with youth your age different before NFE?

10. How has your participation in the program affected your family relationships?
11. Do you feel like your voice or opinions matter in NFE?
 - a. In what ways?
 - b. Do you think youth should have more voice in the way the NFE program is run? (If so,) in what ways?
12. What would you change about the program, or add to it, if you were an NFE facilitator?
13. Do you think you have changed at all because of NFE? (In what ways? – e.g., skills, behaviors, attitudes, relationships?)
14. What are the most important benefits you get out of the program?
15. What are your goals/dreams after graduating NFE?
 - a. Do you think NFE has helped/is helping you to achieve them?
16. Sometimes facilitators have a hard time recruiting new youth into the NFE program. What do you think they could do to get more youth to participate?
17. What are the reasons you joined the Youth Council?
18. As a member of the Youth Council, do you think you can make a difference for other dropout youth? (If so,) in what ways?
19. Is there anything else you would like to share?