An empirical study on implementation quality in prevention programs

Novak, Miranda

Doctoral thesis / Disertacija

2013

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Education and Rehabilitation Sciences / Sveučilište u Zagrebu, Edukacijsko-rehabilitacijski fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:158:306689

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-28**



Repository / Repozitorij:

Faculty of Education and Rehabilitation Sciences - Digital Repository







University of Zagreb

FACULTY OF EDUCATION AND REHABILITATION SCIENCES

Miranda Novak

AN EMPIRICAL STUDY ON IMPLEMENTATION QUALITY IN PREVENTION PROGRAMS

DOCTORAL THESIS

Zagreb, 2013



Sveučilište u Zagrebu

EDUKACIJSKO-REHABILITACIJSKI FAKULTET

Miranda Novak

EMPIRIJSKA STUDIJA KVALITETE IMPLEMENTACIJE PREVENTIVNIH PROGRAMA

DOKTORSKI RAD

Zagreb, 2013



University of Zagreb

FACULTY OF EDUCATION AND REHABILITATION SCIENCES

Miranda Novak

AN EMPIRICAL STUDY ON IMPLEMENTATION QUALITY IN PREVENTION PROGRAMS

DOCTORAL THESIS

Supervisors:

Clemens Hosman
Celene Domitrovich

Zagreb, 2013



Sveučilište u Zagrebu

EDUKACIJSKO-REHABILITACIJSKI FAKULTET

MIRANDA NOVAK

EMPIRIJSKA STUDIJA KVALITETE IMPLEMENTACIJE PREVENTIVNIH PROGRAMA

DOKTORSKI RAD

Mentori:

Clemens Hosman
Celene Domitrovich

Zagreb, 2013

Na temelju članka 46. Statuta Edukacijsko-rehabilitacijskog fakulteta Sveučilišta u Zagrebu, Fakultetsko vijeće Edukacijsko-rehabilitacijskog fakulteta na 4. redovitoj sjednici održanoj dana 28. siječnja 2013. godine donijelo je odluku o imenovanju Povjerenstva za obranu doktorske disertacije u sastavu:

1. **Prof. dr. sc. Josipa Bašić**, redovita profesorica Edukacijsko-rehabilitacijskog fakulteta Sveučilišta u Zagrebu, predsjednica
2. **Prof. dr. sc. Gordana Pavleković**, redovita profesorica na Školi narodnog zdravlja Andrija Štampar, Medicinski fakultet Sveučilišta u Zagrebu, članica
3. **Prof. dr.sc. Damir Ljubotina**, izvanredni profesor Filozofskog fakulteta Sveučilišta u Zagrebu, član

Imenovano povjerenstvo pozitivno je ocijenilo doktorsku disertaciju Mirande Novak pod naslovom „An empirical study on implementation quality on prevention programs“.

Disertacija je uspješno obranjena pred navedenim povjerenstvom 11. ožujka 2013. godine na Edukacijsko-rehabilitacijskom fakultetu Sveučilišta u Zagrebu.

CURRICULUM VITAE OF DOCTORAL STUDY SUPERVISORS

Prof. dr.sc. Clemens Hosman

Clemens Hosman is emeritus Professor of Mental Health Promotion and Prevention of Mental Disorders at the Maastricht University (Health Promotion) and Radboud University Nijmegen (Clinical Psychology) in The Netherlands and former director of their common Prevention Research Centre. He is also guest professor of prevention at the University of Zagreb within the PhD program on Prevention and Prevention Science in the field of mental health. His research in collaboration with many Dutch mental health and public health centres is targeted at prevention of depression, prevention of transmission of mental disorders from parents to children, development of preventive community interventions, principles of program development, effectiveness studies, effect management, and innovation of prevention approaches and policies. He is offering (post)graduate prevention seminars and training at both Dutch universities and in other European countries.

He is involved in prevention and mental health promotion since 1969 and played a crucial role in the development of this field in the Netherlands and Europe. He has occupied many international functions, e.g. chair of the European WHO Task Force on MH Promotion and Prevention, chair of the Clifford Beers Foundation, board member of the international Society for Prevention Research, and chair/member of the Global Consortium for Prevention/Promotion in Mental Health. He was co-project leader of several European prevention projects (e.g. IMHPA, DataPrev) and serves as prevention consultant to NGO's and professional and governmental agencies in many countries. He is co-author of the PREFFI 2.0, an effect management instrument that aims to enhance the implementation of scientific knowledge on effectiveness in prevention policy and practice. He was principle editor of the WHO Report on Evidence-based Prevention of Mental Disorders (Hosman, Jane-Llopis & Saxena, 2004).

A selection of publications for prof. dr. sc. Clemens Hosman

Molleman, , G.R.M., Peters, L.W.H., Hosman, C.M.H., Kok, G., & Oosterveld, P. (2005). Project quality rating by experts and practitioners with Preffi 2.0 as a quality assessment instrument. *Health Education Research, 21 (2),* 219-229.

Patel, V., Araya, R., Chatterjee, S., Chisholm, D., Cohen, A., De Silva, M., Hosman, C., McGuire, H., Rojas, G., &van Ommeren, M. (2007). Treatment and prevention of mental disorders in low-income and middle-income countries. *Lancet, 15*, 370 (9591), 991-1005.

Van Doesum, K., Riksen-Walraven, M., Hosman, C., & Hoefnagels, C. (2008). A Randomized Controlled Trial of an Early Intervention Aimed at Preventing Relationship Problems in Depressed Mothers and their Infants. *Child Development,79 (3),* 547-561.

Kraag, G., Van Breukelen, G.J.P, Kok, G., & Hosman, C. (2009). 'Learn Young, Learn Fair', a Stress Management Program for Fifth and Sixth Graders: Longitudinal Results from an Experimental Study. *Journal of Child Psychology and Psychiatry, 50 (9),* 1185-1195.

Kersten-Alvarez, L.E., Hosman, C.M.H., Riksen-Walraven, M., van Doesum, K.T.M., & Hoefnagels, C. (2010). Which preventive interventions effectively enhance depressed mothers' sensitivity? A meta-analysis. *Infant Mental Health Journal, 32 (3),* 362-376.

Dr.sc. Celene Domitrovich

Celene Domitrovich is research assistant professor of Health and Human Development, College of Health and Human Development, Pennsylvania State University, and a Director of Research at Collaborative Association for Social and Emotional Learning (CASEL) Chicago. Over the past 12 years, she is conducting randomized clinical trials and evaluations of community-based preventive interventions which provided her with extensive knowledge regarding the critical theoretical and methodological issues that are relevant to developing interventions and conducting high-quality prevention research. Her research interests are: developmental psychopathology, interventions, implementation quality, school readiness, mental disorders, development of social and emotional competence in young children, relationship between skills and school success.

Domitrovich is a member of several teams involved in researching the effectiveness of integrated models of evidence-based interventions. For example, she is part of a research team at Johns Hopkins University that is testing the integration into elementary schools of PATHS (Promoting Alternative Thinking Strategies), a program for educators that is designed to facilitate the development of self-control, emotional awareness, and interpersonal problem-solving skills, and the Good Behavior Game, an approach that rewards children for displaying appropriate behaviours during instructional times. She has in-depth knowledge of dissemination and implementation research, and has vast experience in proposing grants and assuring funding. She is one of developers of the PATHS curriculum and certified PATHS trainer.

During 2011, she has won the Joseph E. Zins Award in recognition of her unique contributions to research on social and emotional learning (SEL). The award is reserved for outstanding work of an early-career scientist who promotes social and emotional skills.

A selection of publications for dr. Celene Domitrovich

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs for school aged children. *Journal of Educational and Psychological Consultation, 11,* 193-222.

Domitrovich, C. E., & Greenberg, M. T. (2004). Preventive interventions with young children: Building on the foundation of early intervention programs. *Early Education and development, 15,* 365-370.

Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2004).*The study of implementation in school-based preventive interventions: Theory, research, and practice*. Washington, DC: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

Domitrovich, C.E., Bradshaw, C.P., Poduska, J.M., Hoagwood, K., Buckley, J.A., Olin, S., Hunter Romanelli, L., Leaf, J. P., Greenberg, M.T., Ialongo, S.N. (2008). Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: A Conceptual Framework. *Advances in School Mental Health Promotion, 1 (3),* 6-28.

Domitrovich, C.E., Gest, S.D., Jones, D., Gill, S., Sanford DeRousie, R.M. (2010). Implementation Quality: Lessons Learned in the Context of Head Start REDI Trial. *Early Childhood Research Quarterly, 25,* 284-298.

ACKNOWLEDGEMENTS

I would like to express my warm and sincere gratitude to many people for the support they have given me throughout the course of my Ph.D. In the first place, I have to mention my wonderful colleague Josipa Mihic who has been a part of this project and research, and beyond everything, my life. I would like to acknowledge her energy, enthusiasm, professional but above all, human qualities. She has been a big support, associate and friend. I couldn't have made it without her. She is, among others, one of the factors why I love my work. I am very grateful for the chance to work with two great prevention scientists: Clemens Hosman and Celene Domitrovich. Their knowledge, wisdom, patience, time and efforts invested in feedback and my guidance are a part of this dissertation. It has been a real privilege to work with them and learn from the best, but most of all, they were friends to me, believing that I will do it when I didn't. Professor Hosman has supported the crazy idea of two enthusiasts, while Celene has stepped in, not only because she feels at home in implementation quality – she feels the same in Croatia. Special thank you and a 'I owe you' goes to Rob Nix, statistician and methodologist who has helped me with the SAS and HLM struggles and survived all my stupid questions with a smile. This project also wouldn't happen without Gerard Molleman who has provided us with Preffi 2.0 research licence. My job at a Faculty and work with the Region of Istria wouldn't be possible without long term investments of my boss, professor Josipa Basic who is responsible for international collaboration, founding of our Ph.D study and for development of prevention as a scientific discipline in Croatia.

I would also like to show my gratitude to Sonja Grozic Zivolic from the Region of Istria and to all of the participants of the Training for Prevention! Although we were struggling at first, the experience was wonderful; I had a sense like I was doing something worthy and had a lot of fun during that.

I am also indebt to all my friends, for putting up with me, sending silly messages because I have asked for it, keeping me sane. Big hug 'psiho-curkama'. Thanks to Ivan for being here for me, challenging my soul, reminding me what things in life matter. I couldn't make it without my gestalt training peer group, I have felt you guys, cheering for me all the time. A special thanks goes to my family. Mum for being caring and bringing frozen food, Dad because he said that I can quit if I want, to Boris because he was strict while trying to understand and to my Tena, for being here for me when I was in the worst possible state. I did it, for all of you as well!

SUMMARY AND KEY WORDS

This doctoral research examines the issue of implementation quality in mental health promotion and prevention programs that are being delivered in the community settings in the Region of Istria, taking into account the programs' characteristics and the support system surrounding them. The general aim of this doctoral research was to study implementation processes and their outcomes in prevention programs in Croatia. In order to answer above stated aim, this doctoral research was conducted through pre-research concerning the construction of implementation scales and two studies, 1) study on implementation quality and 2) study of the impact of Training for Prevention. Sample of the study of implementation quality included managers, implementers and participants from the cohort of 24 community-based mental health promotion and prevention programs in Istria. Data for each of 24 included programs were collected from 24 managers and 55 program implementers. Altogether 434 program participants gave their report about the implementation quality at mid-intervention while 744 participants gave their report at post-test implementation assessment.

In general, results have shown that implementation factors are consistently rated lower than indicators of implementation quality. Managers report upon lower levels of implementation factors than implementers and those results could indicate possibilities for future investments. On the general level, average results per indicators of implementation quality support the conclusion that implementation quality in the studied mental health and prevention programs in the County of Istria are satisfactory to high in the perception of implementers and participants. The study on the impact of the Training for Prevention examined if the newly designed Training for Prevention has affected implementation factors and implementation quality in the experimental group. Analyses have shown that program managers and implementers don't report about improved implementation. Post-test differences have shown that program participants report upon the higher level of two indicators of implementation quality, quality of delivery and responsiveness than participants in the control conditions. Regarding the intervention Training for Prevention effectiveness and its implications on the level of implementation quality, it could be concluded that Training seems promising but its impact still has to be further researched.

Key words: mental health promotion and prevention, implementation research, implementation factors, indicators of implementation quality, implementation outcomes

PROŠIRENI SAŽETAK I KLJUČNE RIJEČI

Kroz dugogodišnju i plodnu suradnju Istarske županije s Edukacijsko-rehabilitacijskim fakultetom te međunarodnim istraživačkim centrima, čelnici unutar Upravnog odjela za zdravstvo i socijalnu skrb koji je zadužen za prevenciju problema u ponašanju i promociju mentalnog zdravlja, postupno razvijaju temelje za prevencijsku praksu utemeljenu na dokazima. Prema trenutnom stanju u području, upravo zahvaljujući naporima spomenutog Upravnog odjela, može se zaključiti da je Istarska županija otišla najdalje u ulaganjima u znanstveno utemeljenu prevenciju. Upravni odjel za zdravstvo i socijalnu skrb Istarske županije provodi procjenu potreba, ima svoje preventivne prioritete te jasno usmjerena ulaganja, povezan je s domaćim i internacionalnim istraživačkim centrima te primjerima dobre prakse, zainteresiran je za evaluaciju učinkovitosti programa te kontinuirano ulaže u kvalitetu prevencijske prakse.

Zbog potrebe za donošenjem novog Plana za zdravlje građana Istarske županije, županijska uprava trebala je osigurati transparentnu dodjelu sredstava tj. financiranje onih programa koji poštuju kriterije kvalitete i učinkovitosti. Pokretu za razvoj jasno definiranih kriterija kvalitete pridonijele su i nove znanstvene spoznaje o prediktorima učinkovitosti kao i veća senzibiliziranost provoditelja programa i voditelja organizacija na povratne informacije koje su dobivali kod ocjena projektnih prijava.

Sukladno navedenome, dugogodišnja suradnja Upravnog odjela za zdravstvo i socijalnu skrb Istarske županije s Edukacijsko-rehabilitacijskim fakultetom, Sveučilišta u Zagrebu u okviru projekta Zajednice koje brinu nastavljena je i kroz istraživački podprojekt pod nazivom »Preffi - osiguranje kvalitete u Istarskoj županiji«. Ulaganje u kvalitetu i učinkovitost tj. u implementacijske i programske ishode bilo je planirano kroz primjenu Preffi 2.0 instrumenta u hrvatskim uvjetima te provedbu novo osmišljene intervencije Trening za prevenciju. Preffi 2.0 je mjera za upravljanje učincima promocije zdravlja koja je osmišljena u Nizozemskoj a koristi se za procjenu učinkovitosti i razvoj kvalitete preventivnih programa (Molleman, 2005; Molleman i suradnici, 2005). Instrument se temelji na poduzimanju teorijski utemeljenih koraka te izbora koji trebaju biti slijeđeni tijekom dizajna i primjene programa kako bi se povećala krajnja učinkovitost preventivnih intervencija. S obzirom da su voditelji organizacija i provoditelji programa različitog profesionalnog porijekla, s obzirom na to da je prevencijska znanost nova disciplina koja raspolaže jasnim znanstveno utemeljenim

preporukama a da se upravo trening i stručno osposobljavanje provoditelja intervencija najčešće napominje kao način na koji se može utjecati na kvalitetu programa, Trening za prevenciju osmišljen je kao konkretna metoda ulaganja u kvalitetu. Dugoročni cilj projekta bio je smanjiti mentalne i ponašajne probleme djece i mladih na području Istre kroz osnaživanje prevencijske prakse koja je utemeljena na dokazima. Da bi se postigao ovaj krajnji ishod, kratkoročniji ciljevi projekta bili su:

1. Kroz provedbu intervencije "Trening za prevenciju" poboljšati znanja o načelima znanstveno utemeljene prevencijske prakse kod onih voditelja organizacija, autora i provoditelja preventivnih aktivnosti koji su financirani od strane Istarske županije;
2. Poboljšati kvalitetu prijavnih prijedloga organizacija koje se javljaju za sufinanciranje pri Odjelu za zdravstvo i socijalnu skrb Istarske županije;
3. Poboljšati implementacijske i programske ishode onih promotivnih i preventivnih programa koji su financirani od strane Odjela za zdravstvo i socijalnu skrb Istarske županije;
4. Utemeljiti jasne znanstvene kriterije kvalitete koji će služiti za financiranje preventivnih programa u Istarskoj županiji.

Glavna pretpostavka projekta bila je da je ugradnja principa znanstvene utemeljenosti, tzv. prediktora učinkovitosti, ključna za poboljšanje kvalitete i učinkovitosti programa koji se bave djecom, mladima i obiteljima u Istarskoj županiji.

Ovo doktorsko istraživanje bavi se empirijskom studijom kvalitete implementacije preventivnih programa koji se provode na području Istarske županije, uzimajući u obzir karakteristike sustava te obilježja podrške koja ih okružuje. Opći cilj ovog doktorskog rada bio je proučiti procese implementacije i njihove ishode u preventivnim programima na uzorku programa koji se bave prevencijom problema u ponašanju i promocije mentalnog zdravlja u Istarskoj županiji.

Kako bi odgovorilo na navedeni cilj, ovo doktorsko istraživanje je provedeno u tri koraka: 1) predistraživanje za konstrukciju instrumenata za mjerenje implementacije; 2) studija kvalitete implementacije te 3) studija utjecaja Treninga za prevenciju na kvalitetu implementacije. U svrhu ovog doktorskog istraživanja, na temelju pregleda literature i dosadašnjih teorija implementacije, osmišljen je konceptualni model implementacije koji u odnos stavlja implementacijske faktore te indikatore kvalitete implementacije. Sukladno konceptualnom modelu, konstruirane su i nove mjere koje obuhvaćaju sve aspekte

implementacije u modelu. Ovaj model objedinjuje ekološke, individualne i intervencijske faktore potrebne da bi proces implementacije bio uspješan, a opisani su u radu Fixsen i suradnika (2005, 2009), Domitrovich i suradnika (2008), u radu autora Durlaka i Duprea (2010) te Durlaka (2010). Ovaj konceptualni model implementacije uključuje implementacijske faktore sa dviju razina. Prva razina odnosi se na kapacitet same organizacije da osigura podršku implementaciji programa: osiguravanje adekvatnog treninga i osposobljavanja provoditelja; podrška pojedincima koji provode program; stavovi prema intervenciji te praćenje procesa implementacije programa. Druga razina odnosi se na sam program te obuhvaća čimbenike poput vještina provoditelja programa; stavova provoditelja prema intervenciji te standardizacije programa. Upravo su ovi implementacijski faktori odabrani od cijelog niza faktora prezentiranih u literaturi kao oni koji su najzastupljeniji u svim pregledima jer su povezani s programskim ishodima. Isto tako, prema iskustvu istraživača, ovi su se faktori činili najrelevantnijima za hrvatski kontekst.

Kvaliteta implementacije, primarni ishod prezentiranog konceptualnog modela, zastupljena je kroz pet indikatora kvalitete implementacije o kojima su informacije prikupljane i od provoditelja programa i od samih sudionika intervencije. Četiri od pet tih dimenzija su definirali Durlak i Dupre (2008): vjernost (*fidelity*), kvaliteta provedbe programa (*quality of delivery*), doziranje (*dosage*) te reakcije sudionika programa (*participants' responsiveness*). Mjere koje se odnose na reakcije sudionika programa uključivale su reakcije sudionika na samu intervenciju, reakcije sudionika na provoditelja programa te čestinu pohađanja programa. Uz te četiri već poznate dimenzije implementacije, odlučeno je da će se od provoditelja i sudionika intervencija prikupiti i indirektna mjera kvalitete implementacije, informacije o percepciji učinka samog programa na njihov život (*perceived program impact*). Ova se dimenzija temelji na pretpostavci da ukoliko provoditelji ili sudionici programa izvještavaju o utjecaju intervencije na njihovo ponašanje, vrlo je vjerojatno da je implementacija programa bila uspješna. Osnovna pretpostavka modela odnosi se na povezanost i međuzavisnost implementacijskih faktora i indikatora kvalitete implementacije što je i bio predmet studija u ovom doktorskom radu.

Uzorak ove doktorske disertacije bio je slojevit: činila ga je kohorta od 24 programa prevencije problema u ponašanju i promocije mentalnog zdravlja tj. ukupan uzorak činili su voditelji uključenih organizacija, provoditelji programa i sami sudionici intervencija koje su se provodile. Ovi su programi odabrani iz skupine programa koje financira Upravni odjel za zdravstvo i socijalnu skrb Istarske županije. Uključeni programi podijeljeni su na kontrolnu i

eksperimentalnu skupinu programa metodom putem rezultata na Preffi 2.0 instrumentu te metodom izjednačavanja parova prema karakteristikama programa, broju i tipu sudionika te dužini provedbe. U prikupljanju podataka sudjelovalo je 24 voditelja organizacija te 55 provoditelja programa, iako se broj sudionika mijenjao ovisno o točci mjerenja. Implementacijski faktori i indikatori kvalitete implementacije mjereni su sa četiri novokonstruirana upitnika. Trening za prevenciju u trajanju od 32 sata uz 3 sata individualnih konzultacija proveden je sa voditeljima i provoditeljima 12 programa koji su bili u eksperimentalnoj skupini. U Treningu su bile obuhvaćene sve teme relevantne za učinkovitu implementaciju i postizanje ishoda programa: analiza problema, procjena potreba, važnost teorijske utemeljenosti, izrada logičkog modela, postavljanje ciljeva i rezultata, značaj interaktivnih metoda rada sa korisnicima, važnost standardizacije programa, evaluacija učinaka te metode zagovaranja i osiguravanja podrške.

Kako bi se prikupile informacije i o kvaliteti procesa implementacije te o utjecaju i učinku intervencije Trening za prevenciju, organizirana su mjerenja u dvije vremenske točke. S obzirom na to da program već mora biti u primjeni kako bismo mjerili implementaciju, odlučeno je da prva točka mjerenja bude nakon prve trećine/prve polovice implementacije programa dok je druga točka mjerenja bila nakon same provedbe programa. S obzirom na heterogenost programa u proučavanoj kohorti, posebice razlike u trajanju programa, za neke programe nije bilo moguće organizirati dva mjerenja već samo jedno mjerenje na samom kraju provedbe intervencije. Upravo je stoga nakon prve trećine implementacije u istraživanju sudjelovalo 434 sudionika dok je uzorak sudionika intervencija na kraju provedbe činilo 744 ispitanika. S obzirom na vrstu programa, sudionici intervencija pa i istraživanja bili su i odrasli i djeca. Od 434 sudionika u prvom mjerenju, 297 ih je bilo djece i tinejdžera dok je u drugom mjerenju od 744 sudionika istraživanja 513 djece i tinejdžera. Kako bi se odgovorilo na sve istraživačke probleme, provedena je analiza pouzdanosti i faktorska analiza za odgovor na prvi istraživački problem, deskriptivna analiza za drugi istraživački problem, korelacijska analiza te multipla regresija provedena je za treći i četvrti istraživački problem dok je na peti problem odgovoreno uz pomoć hijerarhijskog linearnog modeliranja.

Preliminarno istraživanje metrijskih karakteristika instrumenata je pokazalo da su četiri novokonstruirane mjere visoke unutarnje konzistencije te da pokazuju dobru konstruktnu valjanost. Rezultati studije kvalitete implementacije su pokazali da su u promatranoj kohorti programa iz Istarske županije implementacijski faktori kontinuirano procijenjivani niže od indikatora kvalitete implementacije. Voditelji organizacija su bili

posebno kritični prema implementacijskim faktorima, te su se posebno kritično osvrnuli na ulaganja u trening i znanja provoditelja programa, standardizaciju samih programa te ulaganja u sustav praćenja. Provoditelji programa implementacijske faktore vide slično kao i voditelji, ali su nešto blaži u ocjenama. Generalno govoreći, prosječni rezultati za indikatore kvalitete implementacije koji su dobiveni iskazom provoditelja i sudionika programa ukazuju na zaključak da je kvaliteta implementacije programa koji su u fokusu ove disertacije zadovoljavajuća.

Studija učinka Treninga za prevenciju na kvalitetu implementacije ispitivala je da li intervencija ima utjecaja na razinu implementacijskih faktora i indikatore kvalitete implementacije. Hijerarhijsko linearno modeliranje pokazalo je da voditelji organizacija i provoditelji programa ne izvještavaju o poboljšanju implementacijskih faktora tj. i kod voditelja organizacija i kod provoditelja programa vidljiv je trend viših rezultata kod kontrolne skupine već u prvoj točci mjerenja. S obzirom da je Trening za prevenciju kojem je bila izložena eksperimentalna skupina sadržavao teme koje su se odnosile upravo na implementacijske faktore, jedna od mogućnosti je da je eksperimentalna skupina upravo zbog Treninga bila senzibiliziranija i kritičnija prema organizacijskim i kontekstualnim faktorima unutar svojih programa. Što se tiče indikatora kvalitete implementacije za koje su samoiskaz davali i provoditelji i ispitanici, analize nisu pokazale utjecaj Treninga na indikatore kvalitete implementacije viđene iz perspektive provoditelja programa. Kada je riječ o samim sudionicima programa, na post-testu su sudionici programa iz eksperimentalne skupine izvještavali pozitivnije o sva četiri ispitivana indikatora kvalitete implementacije. Statistički značajne razlike su pronađene kod dva od četiri indikatora kvalitete implementacije: sudionici iz eksperimentalne skupine izvještavaju o višoj kvaliteti provedbe programa (*quality of delivery*) i većim reakcijama na intervenciju te provoditelja (*participants' responsiveness*). Isto tako, uz pomoć hijerarhijskog linearnog modeliranja provedene su analize moderatora kako bi se provjerilo da li su neki moderatori utjecali na učinak Treninga za prevenciju. Testiran je utjecaj aktivnosti voditelja, dužine programa, tipa ispitanika (djeca, tinejdžeri, odrasli) te tipa programa (roditeljski program, program promocije mentalnog zdravlja te program prevencije ovisnosti). Rezultati su pokazali da su aktivnost voditelja organizacije i dužina programa značajni moderatori učinka Treninga za prevenciju na iskaze od strane ispitanika. Čini se da Trening za prevenciju ima veći učinak na kraće programe: gledano iz perspektive sudionika intervencija, kraći programi iz eksperimentalne skupine čiji su voditelji i provoditelji sudjelovali u Treningu za prevenciju pokazali su veću kvalitetu implementacije

u usporedbi s kontrolnom skupinom programa. Analize moderatora su također pokazale da Trening za prevenciju ima veću učinak na one programe kod kojih je voditelj uključen samo formalno, bez neke istinske podrške provedbi. Kada se takvi programi koji nemaju pravu podršku voditelja organizacije usporede međusobno, programi iz eksperimentalne skupine imali su veću kvalitetu provedbe i bolje reakcije od strane sudionika u usporedbi s kontrolnom skupinom programa.

Ova doktorska disertacija prvi je i jedinstven primjer istraživanja kvalitete implementacije preventivnih programa u Hrvatskoj. S obzirom da je riječ o novom području i u svjetskim okvirima, pregled svih istraživanja također je značajan za područje promocije mentalnog zdravlja i prevencije problema u ponašanju. Znanstveni doprinos predstavlja i konstrukcija četiri instrumenta za mjerenje implementacije te osmišljavanje intervencije koja ima potencijal utjecati na kvalitetu rada praktičara tj. na prevencijsku praksu. Učinkovitost Treninga za prevenciju još treba biti dodatno istražena u novim istraživanjima. Zadnje poglavlje ove disertacije nudi analizu svih manjkavosti i ograda studije te donosi neke preporuke za praksu i buduća istraživanja.

Ključne riječi: promocija mentalnog zdravlja i prevencija, implementacijska istraživanja, implementacijski faktori, indikatori kvalitete implementacije, ishodi implementacije

CONTENTS

1. CHAPTER ONE: INTRODUCTION 4
1.1. Introduction to implementation research agenda 4
1.1.1. Need for implementation research 4
1.1.2. Historical development of implementation research 7
1.1.3. Definition of implementation quality 9
1.1.4. Conceptual framework of implementation 18
1.1.5. Overview of implementation research 26
1.2. Mental health promotion and prevention in Croatia 34
1.2.1. Development of mental health promotion and prevention in Croatia 34
1.2.2. Core problems of mental health promotion and prevention field in Croatia 37
1.3. Research project "Preffi - Quality Assu rance in the County of Istria" 39
1.3.1. The case of Istria 39
1.3.2. Project overview and doctoral research studies 40
1.4. Conceptual framework for the research of implementation quality in the County of Istria 46

2. CHAPTER TWO: AIMS AND RESEARCH QUESTIONS 49

3. CHAPTER THREE: METHODS 54
3.1. Construction of scales for implementation quality assessment 56
3.2. Study on implementation quality 58
3.2.1. Participants 58
3.2.2. Procedure 65
3.3. Study on the impact of the Training for Prevention 65
3.3.1. Participants and matching 65
3.4. Content of the Training for Prevention 72
3.5. Measures 76
3.5.1. The Preffi 2.0 instrument 76
3.5.2. Measures of implementation quality 78
3.6. Ethical considerations 80

4. CHAPTER FOUR: VALIDATION OF MEASURES FOR THE IMPLEMENTATION QUALITY RESEARCH 82
4.1. Reliability and Construct Validity of Measures of Implementation Factors 83

4.1.1. Implementation Factors Questionnaire for Program Managers....83
4.1.2. Implementation Factors Questionnaire for Program Implementers....88
4.2. Reliability and Construct Validity of Measures of Implementation Quality....93
4.2.1. Implementation Quality Questionnaire for Program Implementers....93
4.2.2. Implementation Quality Questionnaire for Program Participants....97
4.3. Discussion regarding the construction and validation of measures....105

5. CHAPTER FIVE: LEVEL OF IMPLEMENTATION QUALITY IN PREVENTIVE PROGRAMS IN ISTRIA....106
5.1. Level and variation of implementation quality in preventive programs in Istria....107
5.1.1. Level of implementation factors seen by program managers and program implementers: findings....108
5.1.2. Level of implementation quality seen by program implementers and program participants: findings....119
5.2. Relationship of implementation factors and indicators of implementation quality: findings....128
5.2.1. Correlation of implementation factors and indicators of implementation quality: findings....128
5.2.2. Hierarchical linear models of association between implementation factors and indicators of implementation quality: findings....131
5.3. Discussion of results describing the level of implementation quality....135

6. CHAPTER SIX: IMPACT OF THE TRAINING FOR PREVENTION ON THE IMPLEMENTATION QUALITY....142
6.1. Impact of Training for Prevention on Implementation Factors....147
6.1.1. Group Comparisons on Manager Ratings of Implementation Factors: Findings....147
6.1.2. Group Comparisons on Manager Ratings of Implementation Factors: Discussion....150
6.1.3. Group Comparisons on Implementer Ratings of Implementation Factors: Findings....152
6.1.4. Group Comparisons on Program Implementer Ratings of Implementation Factors: Discussion....154
6.2. Impact of Training for Prevention on Implementation Quality....157
6.2.1. Group Comparisons for Implementers' Ratings of Implementation Quality: Findings....157
6.2.2. Group Comparisons on Program Implementer Ratings of Indicators of Implementation Quality: Discussion....160
6.2.3. Group Comparisons for Program Participant Ratings of Implementation Quality: Findings....161
6.2.4. Group Comparisons on Participant Ratings of Indicators of Implementation Quality: Discussion....164

6.3. Moderators of the Training for Prevention impact: findings and discussion 165

7. CHAPTER SEVEN: DISCUSSION AND CONCLUDING REMARKS 171
7.1. General discussion 171
7.2. Limitations of the study and suggestions for future research 182
7.3. Practical implications 188
7.3.1. Recommendations for the County of Istria 188
7.3.2. Perspectives of Training for Prevention capacity 191
7.4. Conclusions 193

REFERENCES 197

APPENDIX .....
APPENDIX 1 211
APPENDIX 2 215
APPENDIX 3 219
APPENDIX 4 224
APPENDIX 5 227
APPENDIX 6 232

BIOGRAPHY OF THE CANDIDATE 265

1. CHAPTER ONE: INTRODUCTION

1.1.Introduction to implementation research agenda

1.1.1. Need for implementation research

A 1994 report by the Institute of Medicine, Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research (Mrazek and Haggerty Eds., 1994), highlighted the potential of prevention which has led to serious efforts of investments in mental health promotion and prevention field, not just in the United States but worldwide. Since that report to the 2009 report by the Institute of Medicine, Preventing Mental, Emotional and Behavioral Disorders among Young People: Progress and Possibilities (O'Connell, Boat and Warner, Eds, 2009), the volume and quality of research in the mental health promotion and prevention field have increased dramatically. Over the last two decades, prevention science related to developing and identifying evidence-based practices and programs has improved: prevention scientists worldwide know a lot about interventions which are effective (Kellam and Langevin, 2003; Fixsen, Naoom, Blase, Friedman and Wallace, 2005; Flay and colleagues, 2005; Hosman, 2008). Studies have clearly shown that a number of specific preventive interventions can modify risk and promote protective factors that are linked to important determinants of mental, emotional, and behavioural health, especially in areas such as family functioning, early childhood experiences, and social skills (O'Connell, Boat and Warner, Eds, 2009). There is a respectable number of empirically validated prevention programs that can reduce children's risks to substance abuse (Tobler and Stratton, 1997), mental disorders (Hawkins, Catalano and Arthur, 2002; Hosman, Jané-Llopis and Saxena, Eds, 2005; Jané-Llopis and Barry, 2005; O'Connell, Boat and Warner, Eds., 2009), aggression, delinquency and other risky behaviours (Botvin and Griffin, 2007) as well as promote competencies (Mrazek & Haggerty, Eds., 1994; Greenberg, Domitrovich, Grazyk, Zins, 2005).

In its first report about the research in mental health promotion and prevention, the Institute of Medicine (Mrazek & Haggerty, Eds., 1994) presented a model for prevention research cycle (Figure 1). It is a five step model which incorporates assessment of risk and protective factors, development of program innovations that address identified problems, research of their efficacy and effectiveness and dissemination of those tested innovations into the community. The Feedback Loop presented in Figure 1 stresses the importance of informative epidemiological research after interventions have been delivered to the target

population to check if there are reductions in the incidence of targeted problems (Mrazek & Haggerty, Eds., 1994).

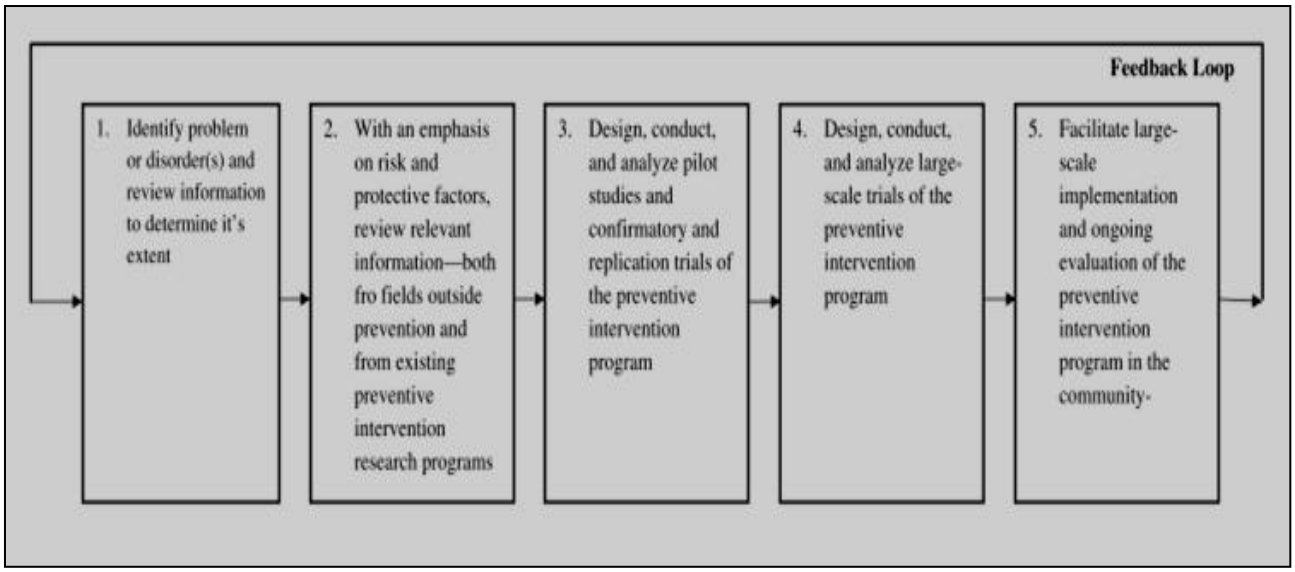


Figure 1.Model of the prevention research cycle presented in the report by the Institute of Medicine (taken over from Mrazek & Haggerty, Eds., 1994, page 16).

Greenberg and colleagues (2005) conclude that until the mid-to-late 1990s, the main research question in the mental health promotion and prevention field was to determine whether the program worked. That movement led to refinement of rigorous methodological procedures of outcome evaluation and internal program validity.Wandersman and colleagues (2008) refer critically to the IOM prevention research cycle model presented in Figure 1 stressing that even though this model supported advancement in the mental health promotion and prevention field, it does not offer information how that jump from research to practice would occur. The same has been stated by other mental health promotion and prevention scientists: although rigorously researched, preventive interventions have not been widely implemented in schools and communities, even in countries where mental health promotion and prevention is well developed (Hallfors and Godette, 2002).

Although the efficacy of various evidence-based interventions has been established through carefully designed trials in control conditions, there is a lack of evidence for its utilization in natural community conditions (Kam, Greenberg and Weiss, 2003; Greenberg et al., 2005; Fixsen et al., 2005, Proctor and Rosen, 2008). It seems that those evidence based interventions are being used but not on a big scale and have done little to reduce behavioural problems of children and youth on public health level (Chinman et al., 2005). The science needed to promote successful implementation of evidence-based practices in real time, under naturally occurring conditions, is poorly developed (Greenberg et al., 2005). Even when

organizations and communities take over empirically supported programs, it may be difficult for them to achieve the same levels of technical assistance, support, resources, and prevention expertise available in well-funded, controlled prevention research trials. A more systematic process is warranted to translate efficacy results into positive participants' outcomes, with special attention to factors that contribute to the quality of program dissemination in a variety of settings.

There is a need for new approaches to supplement the existing approaches of research to practice models and the evolving community-centred models for bridging this gap (Wandersman et al., 2008). Clearly, policies aimed at improving human services on national levels require more effective and efficient methods to translate policy mandates into actions with effective interventions. Fixsen and colleagues (2005), Proctor and Rosen (2008), Wandersman and colleagues (2008), Flaspohler, Duffy, Wandersman, Stillman and Maras (2008) acknowledge the importance of both approaches which will not only answer the question of what has to be done but also of how this will be done in practice. Researchers call for applied research to better understand service delivery processes and contextual factors to improve the efficiency and effectiveness of program delivery at local, state, and national levels (Fixsen et al., 2005). That paradigm shift in the mental health promotion and prevention field is seen in the 2009 IOM report (O'Connell, Boat and Warner, Eds., 2009) which states that the future of prevention requires combined efforts to (1) apply existing knowledge in ways that are meaningful to families and communities and (2) pursue a rigorous research agenda that is aimed at improving both the quality and implementation of interventions across diverse communities.

Wandersmann and colleagues (2008) explain that implementation and dissemination processes and research are often mentioned as a missing puzzle for bridging the gap between prevention science and practice. That is the reason why implementation research is becoming a new direction of translation research: important step towards the advancement of applied prevention science. In each study, there are intervention processes and outcomes and there are implementation processes and outcomes (Greenberg et al., 2005).When implementing evidence-based practices and programs, there is the need to discriminate implementation outcomes from effectiveness outcomes which answer the question if the intervention is resulting in good outcomes. Distinguishing implementation effectiveness from intervention effectiveness is critical for transporting interventions from a laboratory setting to the community (Proctor et al., 2010). Generally speaking, a variety of contextual issues, such as

leadership, the nature of an organization, program deliverer's training and support, may influence both the level and the quality of program implementation. The investment in developing structures to ensure gold standard research evidence has yet to be matched by equal investment in ways of elucidating how organizations change cultures or use different techniques to manage the change process. Past research indicates that when communities replicate programs, the quality of delivery can vary widely and aspects of the program are altered compared to the model to match community characteristics. Thus, research is needed to identify the specific elements of evidence-based programs that are essential to program success and those elements that may be modified while remaining true to the intended purpose or concept underlying the model.

The main impulse for the new movement in mental health promotion and prevention is the question of how to get from prevention efficacy to prevention effectiveness. "Prevention practice will reach its full maturity only when known effective programs are implemented with sufficient integrity" (Greenberg and colleagues, 2005, page 11). One of the main answers to that question was to start improving implementation quality (Domitrovich and Greenberg, 2000; Gottfredson and Gottfredson, 2002; Dusenbury, Brannigan, Hansen, Walsh and Falco, 2005; Proctor and Rosen, 2008).

From an implementation point of view, doing more extensive and better research on a program or practice itself does not lead to more successful implementation. A series of meta-analyses and detailed assessments of the strength of research findings for certain practices and programs may help a consumer, agency, or community to select a program (Derzon, Sale, Springer, Brounstein, 2005; Galavotti, Sebert Kulman, Kraft, Harford, and Petraglia, (2008). However, more data on program outcomes will not help implement that program: implementation is an entirely different process (Greenberg et al., 2005). A paradigm shift occurred with the confirmation that prevention programs work and with the need for identifying factors that influence effects on outcomes. Evaluators realized that even thoroughly researched interventions do not yield positive outcomes unless they were implemented with integrity.

1.1.2. Historical development of implementation research

Fixsen and colleagues (2005) state that current views of implementation are based on the foundations laid down by Pressman & Wildavsky's study of policy implementation (1973, according to Fixsen and colleagues, 2005) and Havelock & Havelock's (1973, according to

Fixsen and colleagues, 2005) classic curriculum for training change agents. Also, it is important to mention the contribution of Rogers (1974, 1976) and his research on diffusion of innovations i.e. factors connected with decisions to choose a given innovation. Within Rogers (1995) classic model, implementation is one of five crucial stages in a wide-scale diffusion of innovations: 1) dissemination (conveying information about the existence of an innovation to potentially interested parties), 2) adoption (an explicit decision by a local unit or an organization to try the innovation), 3) implementation (executing innovation effectively when put in place), 4) evaluation (assessing how well the innovation achieved its intended goals) and 5) institutionalization (the unit then incorporates the innovation into its usual practices). Researching the innovation in health services, Greenhalgh, Robert, MacFarlane, Bate and Kyriakidou (2004) take Rogers classic model (1974, 1976) in the centre from which they present a review of diffusion of innovation field.

Going from the broader field of innovation instalment to more narrow mental health and prevention research field, Greenberg and colleagues (2005) mention three separate but related fields which provide a developmental perspective of program implementation research: education, program evaluation and school prevention programs field. Greenberg and colleagues (2005) state that during the 1970s educators stressed the importance of assessing the degree to which educational programs were implemented as intended. The same authors stress that many educational researchers tried to understand implementation in order to understand the impact of a specific program when compared with the impact of teacher style, wanted to describe how a program is being delivered and were interested in resources necessary for a program to function appropriately.

Regarding the program evaluation field, Greenberg and colleagues (2005) conclude that until the mid-to-late 1990s, the primary aim was accurate measurement of program outcomes. The main research question was to determine whether the program worked. That movement led to refinement of rigorous methodological procedures of outcome evaluation and internal program validity as it was earlier mentioned in the review of the IOM report (Mrazek & Haggerty, Eds., 1994). At first, in the evaluation field, term process evaluation was used for measurement of program delivery, especially to follow participants receiving the program and components delivered. With the development of the implementation research movement, it has become a lot more than process evaluation.

Greenberg and colleagues (2005) also address the important role of school-based prevention and positive youth development field, which has invested the most concentrated efforts into implementation research. To start with, the implementation aspect of prevention began to develop within school based prevention programs, with an emphasis on teacher's characteristics and behaviour as well as school climate (Kam, Greenberg and Walls, 2003; Han and Weiss, 2005). Within these areas, it has been stressed that apart from program characteristics, context and environment where a program is being delivered also have to be a part of implementation research (Durlak, 1998).

1.1.3. Definition of implementation quality

With respect to implementation research, there is no agreed-upon set of terms. The diversity of literature sources, language, definition of concepts, and data collection methods causes many problems in reviewing the literature and making common ground. There are few organized approaches for executing and evaluating implementation practices and outcomes while good research designs are a challenge considering the number of observed variables and taking into account that one program, if in focus, represents the general number of subjects. Fixsen and colleagues (2005) reviewed articles on implementations of which 377 were identified as important and describe that the review was challenging due to the lack of consensus and understanding of implementation as a term or a process. For example, diffusion, dissemination, and implementation sometimes referred to the same general construct and, at other times, different meanings were ascribed to the same terms. For example, "implementation" sometimes means "used" in a general sense or "put into effect" with specific reference to a program or practice. At other times it referred to a set of methods to purposefully help others make use of a program or practice on a broad scale (Fixsen et al., 2005), which is connected with stages of innovation diffusion (Rogers, 1976, 1995) already mentioned above.

Peters and colleagues (2003) add to the discussion more from a European point of view: to some implementation relates to the delivery of a project to the ultimate target group while to others it refers to projects which are carried by intermediaries or purveyors, other people included in health promotion projects. To add to the complexity, in older reviews and literature implementation quality has been referred to as "treatment integrity", "fidelity" and "adherence" (Dane and Schneider, 1998). Those terms can be found in some papers even nowadays, especially "fidelity" as a synonym for the whole implementation process, so readers which are not very familiar with the field could be confused.

For the purpose of this dissertation, several aspects of implementation will be discussed. First definitions of implementation have started as related with implementation of evidence-based programs. That is the reason why most traditional definitions describe implementation quality **as the degree to which an intervention is conducted as it was originally intended** (Durlak, 1995). This definition is based on the assumption that the evaluator and the community of implementers specify the intervention before beginning the program and then measure how the intervention actually is conducted in the field (Greenberg and colleagues, 2005). A few years later, Durlak (1998) defines implementation as the **degree of adequate setting of a program into practice** and it refers to specific program components and the way in which those components are delivered in practice.Fixsen and colleagues (2005) offer a similar definition where implementation is seen **as a specified set of activities designed to put into practice an activity or program of known dimensions**. According to this concept, implementation processes are purposeful and should be described in sufficient detail so that independent observers can detect the presence and strength of the "specific set of activities" related to implementation.

Those definitions were exceeded with efforts of implementation scientists during 2000s, when the intervention aspect of implementation studies was broadened by putting emphasis on the influence of the support system which shouldn't be bypassed. Newer generation definitions in general agree that **implementation refers to what a program consist of when it is delivered in a particular setting** (Durlak and Dupre, 2008). Domitrovich and colleagues (2008) establish the definition of implementation that includes **characteristics of intervention itself** and **characteristics of the intervention support system**. This doctoral dissertation seeks for the approach to implementation which has moved from traditional definitions and could be applicable to community preventive interventions which have not yet been tested i.e. their efficacy and effectiveness is still unknown. This approach can be based on Durlak and Dupre's (2008) definition combined with the definition of Domitrovich and colleagues (2008).

Considering the fact that implementation research is a new and still emerging field of mental health promotion and prevention research, implementation research in countries where effectiveness trials of preventive interventions are still underdeveloped, could develop parallel to traditional research of program impact and be more explorative. Implementation research in these conditions and in countries like Croatia can at the same time be used to support the practice by informing practitioners about the needed investments in the supports system as

well as investments in the development of evidence-based interventions. Newer views also support that: Fixsen, Blase, Naoom and Wallace (2009) state that implementation is the **process of introducing and using interventions in real world setting**.IntheCroatian case, that would mean: assessing implementation quality simultaneously with impact assessment. Meyers, Katz, Chien, Wandersman, Scaccia and Wright (2012) focus on the term **quality implementation** which they define **as putting innovation into practice in such way that it meets necessary standards to achieve innovation's desired outcomes**. Fixsen, Blase, Naoom and Wallace (2009) stress that implementation includes all activities referring **to the "to" in the saying "from science to service"**. That would mean that the Croatian approach to implementation research must start with new views on implementation: incorporating scientific principles in community programs through the training of practitioners while at the same time testing the effectiveness of an intervention. That will be elaborated in depth in the text and chapters that follow.

When defining implementation, it is important to differentiate between **factors affecting implementation quality**, **implementation drivers** (Fixsen et al., 2005, 2009) and **implementation aspects** (Durlak and Dupre, 2008). To elaborate all presented definitions, after aspects of implementation, an overview of factors that are mostly related with context conditions and specifics will be presented. There are eight different **aspects to implementation**, according to Dane and Schneider (1998), and Durlak and Dupre (2008): **fidelity, dosage, quality, participant's responsiveness, program differentiation, monitoring of control/comparison conditions, program reach and adaptation.**

(1) There is **fidelity**, which is the extent to which the innovation corresponds to the originally intended program (also referred as adherence, compliance, integrity, faithful replication). Stith and colleagues (2006) refer to fidelity as a process of delivering a program in the same way in which it was delivered during efficacy and effectiveness trials. If such studies don't exist, the same authors argue that fidelity means delivering a program in the way it was designed to be delivered. Domitrovich and colleagues (2010) define fidelity as a degree to which the core elements of an intervention were conducted as planned. Fidelity measurement should include "core program components" assessment or the amount of time dedicated to each of those core components (Spoth, Guyll, Trudeau and Goldberg-Lillehoj, 2002; Lillehoj, Griffin and Spoth, 2004).

(2) There is **dosage**, which refers to how much of the original program has been delivered (in research also called quantity or intervention strength). Greenberg, Domitrovich and Bumbarger (1999) and Domitrovich and colleagues (2010) define dosage as the amount of exposure participants had to an intervention which is often presented in terms of specific units of an intervention (e.g. number of lessons delivered) or amount of time (e.g. hours of contact with participants). It is crucial, to differentiate here between the part of the original program that was actually implemented by a provider, which is linked to the issue of fidelity; and dosage in terms of the actual duration of a program and exposure time with the participant, regardless of whether the exposure time was equal to the planned exposure time defined by the program designer, or if the actual exposure time was shorter due to poor fidelity or contextual restraints.

(3) **Quality** refers to how well different program components have been conducted (e.g., answer to the question if the main program elements were delivered clearly and correctly). Greenberg and colleagues (2005) and Domitrovich and colleagues (2010) state that quality refers to the qualitative aspect of program content, rather a process-oriented dimension and is examined less frequently in research because the best way to assess it are direct observations. Dusenbury and colleagues (2005) used the 'quality of process' term to underline that an interactive approach to participants is needed for learning and behaviour change. Domitrovich and colleagues (2010) conceptualize it in terms of how the intervention content is delivered and responded to. Quality of delivery aspect of implementation is completely related to the delivery behaviour of the implementer (e.g. interpersonal style, affective engagement, sensitivity to participants' needs, generalization of the content outside the intervention's lessons). Odom and colleagues (2010) describe quality in terms of deliverer's preparation of large group and small group activities, the skill with which lessons were delivered, integration of concepts into other activities during the day (the deliverers were school teachers) and deliverer's responses to participants. Dane and Schneider (1998), Coatsworth, Duncan, Pantin & Szapocznik (2006), and Berkel, Mauricio, Schoenfelder, Sandler (2011) state that enthusiasm and clarity with which the implementer presents program activities, reflective listening of participants as well as fostering cohesion among participants are also dimensions of the quality of delivery.

(4) **Participant responsiveness** refers to the degree to which the program stimulates the interest or holds the attention of participants (e.g., answering questions if students are

engaged and attentive during program lessons). Participant's responsiveness is their reaction not only to the intervention but also to the program deliverer and conditions under which the program is being delivered. Berkel and colleagues (2011) stress that responsiveness can also be defined in the number of sessions a participant has attended (attendance), active participation, satisfaction and home practice completion.

(5) **Program differentiation** involves the extent to which a program's theory and practices can be distinguished from other programs (program uniqueness). Dane and Schneider (1999) as well as Greenberg and colleagues (2005) include this implementation dimension because some studies they examined were highly controlled research evaluations in which an intervention group was compared to a control or comparison group that did not receive the test intervention but that unintentionally may have received another type of intervention.

(6) **Monitoring of control/comparison conditions**, which involves describing the nature and amount of services received by members of these groups i.e. a more accurate view of the value of a new intervention (treatment contamination, usual care, alternative services). It is often incorrectly assumed that controls do not receive any services, but this is almost never the case and several authors who have examined the issue have found that many individuals in their no-intervention control condition received some alternative services (Durlak and Dupre, 2008).

(7) **Program reach** (participation rates, program scope) refers to the rate of involvement and representativeness of program participants. It is concerned with questions relating to the percentage of the eligible population that took part in the intervention, and their characteristics. Durlak and Dupre (2008) give an example of a prevention program potentially suitable for all parents in a diverse community that may only attract less than 5% of eligible parents, most of them being from upper class, motivated and not a minority.

(8) **Adaptation** refers to changes made in the original program during the process of implementation (program modification, reinvention). Durlak (2010) explains that there has been a lot of debate on the role of adaptation: some authors feel that adaptation should be avoided whenever possible because it has negative effects on fidelity; others argue that adaptation is needed in real world settings. Some degrees of adaptation are needed when lengthier and complex interventions are delivered, considering the characteristics of participants, time and setting itself. Durlak (2010) explains that one size rarely fits all and that is the reason why it is so important to

document how an innovation is being modified in a new setting in order to study that process systematically. Odom and colleagues (2006) also stress the importance of program adaptation in order to meet service recipients' needs and possible cultural differences.

Since this doctoral dissertation focuses on community based interventions which are not yet proven as effective and have not been researched, four out of eight defined and well known implementation aspects will be in focus. Those are fidelity, dosage, quality of delivery and participant's responsiveness, which are going to be referred to as indicators of implementation quality.

According to the literature overview offered in this chapter, especially the review of Domitrovich and colleagues (2008) it is possible to conclude that eight aspects of implementation are based on the research about the delivery of program which is science-based and reflect the characteristics of intervention when delivered in specific setting. On the other hand, explanation of program implementation wouldn't be complete if we don't study and assess the characteristics of the system supporting that delivery. **Factors affecting implementation quality** and implementation drivers can be seen as implementation predictors while implementation aspects can be regarded as facets of implementation (Fixsen et al., 2005, 2009). When talking about **factors affecting implementation quality**, both Greenberg and colleagues (2005) and Fixsen and colleagues (2005) who have done a comprehensive overview of the implementation field as well as some other authors as Domitrovich and colleagues (2008) and Wandersman and colleagues (2008) come to the convergent conclusions. Durlak and Dupre (2008) identified at least 23 factors from prior research that have affected the implementation process and most of these have been identified independently in other implementation reviews (Greenhalgh et al., 2005; O'Donnell, 2008; Stith et al., 2006). Durlak (2010) explains that ecological factors affecting implementation are present at the broad community level (e.g. funding, politics, current theory and research), some involve characteristics of intended program (e.g. its complexity and its compatibility with institution and staff), some are related to the characteristics of program providers (e.g. perceived need for program, general skills, proficiency, self-efficacy) and others to the organizational capacity of the host system (work climate, leadership, vision and decision making flowcharts). Durlak (2010) also stresses factors associated with a specific support system that had been elaborated before by Greenberg and colleagues (2005), who believe that implementation support system follows certain stages (see Table 1.1.).

Table 1.1

Planned Implementation Support (prepared according to Greenberg and colleagues 2005, page 22).

PLANNED IMPLEMENTATION SUPPORT SYSTEM
1. Pre-planning • Need for change • Readiness for change • Capacity to effect change • Awareness of the need for change • Commitment/engagement in the process of change • Incentives for change • History of prior implementation efforts
2. Quality of materials • Design of program materials • Format of program materials
3. Technical support available • Structure of training and supervision • Content of training and supervision • Timing of training and supervision • Implementation monitoring system
4. Quality of technical support • Quality of delivery • Quality of the working relationships • Trainer characteristics
5. Implementer's readiness • Perception • Skills • Knowledge • Beliefs

Greenberg and colleagues (2005) while analysing the school system as a place of delivery define **pre-planning** as any preparation made by the school before the implementation of an identified intervention. They stress that little quantitative research has been conducted on ways to assess or improve the readiness of a context for an intervention, or the ways in which readiness assessment may predict the quality of program implementation. The majority of prevention programs do not specify pre-planning steps, even though planning decisions made before program implementation can have a significant impact on the success of program adoption. Context evaluation includes an accurate assessment of the student population and its needs, coupled with a determination of the program's capability to address the needs of the targeted group of students. Without such information, it is difficult to assess the likely fit between what students need and what a program offers. Input evaluations encourage schools to analyse their infrastructure to determine whether it is sufficient to handle program needs. Analyses at this program stage consider factors such as availability of needed personnel and material resources, budgeting issues, and feasibility. Once a problem is targeted

for change, all personnel should be aware of the problem in their school, should be informed about how the suggested program will address the problem (i.e., they must understand the program theory), and should be committed to carrying out the program. All of those issues can be translated to different context, not just schools but wherever the new intervention is being planned for delivery.

According to Greenberg et al (2005), the next set of factors affecting implementation comes from the area of **quality of materials** (Table 1.1.). Successful program implementation is more likely when program materials are visually appealing, user friendly, age appropriate, and culturally sensitive. Instructor manuals are probably the single resource most widely used by teachers implementing positive youth development and other instructional prevention programs (Domitrovich and Greenberg, 2000; Greenberg et al., 2005). Consequently, the design and format of instructor manuals may have a significant impact on the quality of program delivery. They are most helpful when they include a comprehensive scope and sequence chart, provide the theoretical rationale for the program, explain the theory's connection to the lesson content and teaching strategies, clearly state the program objectives, and include detailed, well-organized, and easily understood lesson plans.

The next dimension is connected with the **available technical support**. This support includes the structure, content, and timing of pre-intervention training, and any ongoing support required to deliver the program successfully. It also includes the implementation monitoring system or additional technical assistance materials provided by the program. The structure of the technical support determines who delivers program support and how it is delivered to program implementers (e.g., direct training, trainer of trainers' model or a videotape). Training and supervision should be comprehensive and should prepare the implementers to conduct the program. Therefore, the content of the technical support should include the essential elements of the intervention. The timing of technical support refers to the frequency, duration, and pace at which the support and follow-up should be administered.

The next dimension, **quality of technical support**, includes (a) the quality of delivery during training and supervision, (b) the quality of the working relationship between the trainers and the implementers, and (c) the characteristics of the trainers who provide the assistance. Greenberg and colleagues (2005) explain that it is likely that teachers will be more interested in a program when their training is conducted in a collaborative, engaging manner. It is helpful for the teachers to perceive the trainer as one who respects their individual needs and interests and as one who is sensitive to their skill level and learning style. Adherence to

program protocol can be improved and resistance decreased by creating a supportive, cooperative partnership between trainers and implementers. The technical support provided to program staff should establish and maintain open channels of communication and result in effective problem-solving between school personnel and program staff. Efforts should be made to clarify teacher beliefs and expectations about the implementation process and about intervention options and outcomes.

Foremost, indicators of **implementer readiness** (Table 1.1) include whether they have both adequate skills to carry out the intervention and sufficient knowledge about the theoretical basis of the intervention, feel positive about a program, value what it contributes to the educational setting, and are committed to its goals. While talking about school interventions Greenberg and colleagues (2005) stress that if a teacher does not see the value of fostering a specific skill or conducting lessons about particular topics (e.g., sexuality), he or she may be more likely to skip those lessons, even though they may be core parts of the program. Implementers also need to believe that both the intervention and their role in its delivery will be effective. Deliverers' confidence in the effectiveness of an intervention and in their own knowledge and skills affects the ability to deliver a program successfully. The more confident and comfortable they feel when conducting lessons about a particular topic (e.g., suicide), the more likely they are to cover those lessons that are a necessary part of the program.

Fixsen and colleagues (2005; 2009) described **core implementation components** or **implementation drivers** which will be presented in the conceptual model in the next section (see Figure 3.). To conclude the section on factors affecting implementation quality, description of implementation support system by Greenberg and colleagues (2005) has to be expanded by the description of core implementation components. Fixsen and colleagues (2005; 2009) talk about: **pre-service and in-service training**, **staff evaluation** and **facilitative administration support** which is complementary to Greenberg and colleagues (2005) available technical support (see Table 1.1.). Besides that, Fixsen and colleagues (2005; 2009) see that core implementation components are also **staff evaluation** and **program evaluation**.

Pre-service and **in-service training** are effective ways to provide knowledge of background information, theory, philosophy, and values; introduce the components and rationales of key practices; and provide opportunities to practice new skills and receive

feedback in a safe training environment. Most skills needed by successful practitioners can be introduced in training but really are learned on the job with the help of a consultant/coach. Implementation of evidence-based practices requires behaviour change at the practitioner, supervisory, and administrative support levels. Training and coaching are the principle ways in which behaviour change is brought about for carefully selected staff in the beginning stages of implementation and throughout the life of evidence-based practices and programs. **Staff evaluation** is designed to assess the use and outcomes of the skills that are reflected in the selection criteria, are taught in training, and reinforced and expanded in consultation and coaching processes. Assessments of practitioner performance also provide useful feedback to managers and purveyors regarding the progress of implementation activities and the usefulness of training and coaching. **Program evaluation** (e.g., quality improvement information, organizational fidelity measures) assesses key aspects of the overall performance of the organization to help assure continuing implementation of the core intervention components over time. **Facilitative administration** provides leadership and makes use of a range of data inputs to inform decision making, support the overall processes, and keep staff organized and focused on the desired outcomes. A more comprehensive overview of factors affecting implementation quality is presented in the work of Domitrovich and colleagues (2008) which is described in the next section (see Figure 6).

1.1.4. Conceptual framework of implementation

Considering the background of implementation research, Greenberg and colleagues (2005) note that neither of the fields that contributed to the development of implementation research and awareness of implementation process until the end of 2000s developed a conceptual model of implementation. Until today there is a lack of comprehensive models that systematically clarify factors that predict implementation quality and relationships of those factors with various aspects of implementation. Also, the main problem is that a small number of implementation studies was conducted according to a specific framework or a model i.e. most of the models are just theoretical and conceptual and have not been translated into practice in some kind of assessment tool, questionnaire or a checklist. Even those presented here are not accompanied by validated measures for implementation monitoring: that still needs to be done in future implementation research attempts.

For the field of mental health promotion and prevention to continue growing, greater attention and better understanding of the implementation process and the factors that support it, is essential. Regardless of the mentioned lack of validated implementation assessment

measures, first step to help achieve the clarity of implementation field is through the development of comprehensive implementation theory models that integrate perspectives presented in the literature. Flaspohler and colleagues (2008) state that implementation frameworks are the windows into the key attributes, facilitators and challenges related to promoting implementation. Meyers, Durlak and Wandersman (2012) conclude that frameworks provide an overview of ideas and practices that shape the complex implementation process and can help researchers and practitioners use ideas of others. Regarding all contextual factors that affect implementation quality which are presented in the section above, presentation of conceptual frameworks will at least offer some clarification.

There are just few conceptual frameworks currently being used in the mental health promotion and prevention implementation field: framework offered by Fixsen and colleagues (2005) which is broader, concentrated on scientist's perspective (Figure 2) and afterwards elaborated with the view on practitioners (Figure 3); a model presented by Greenberg and colleagues (2005) and Graczyk, Domitrovich, Small and Zins (2006) coming from the field of school prevention (Figure 4); an ecological model from Durlak and Dupre (2008) incorporating Wandersman and colleagues (2008) Interactive System Framework (Figure 5) and a socio-ecological model of Domitrovich and colleagues (2008) which highlights different levels affecting implementation (Figure 6).

Fixsen and colleagues (2005; 2009) present a conceptual model for implementation of well-defined programs and practices based on their literature review and ideas they got from computer programming. While explaining their model, Fixsen and colleagues (2005; 2009) stress that implementation components and outcomes exist quite independently of the quality of the program or practice being implemented. Research has shown that ineffective programs can be implemented well while effective programs can be implemented poorly. Desirable outcomes are achieved only when effective programs are implemented well.

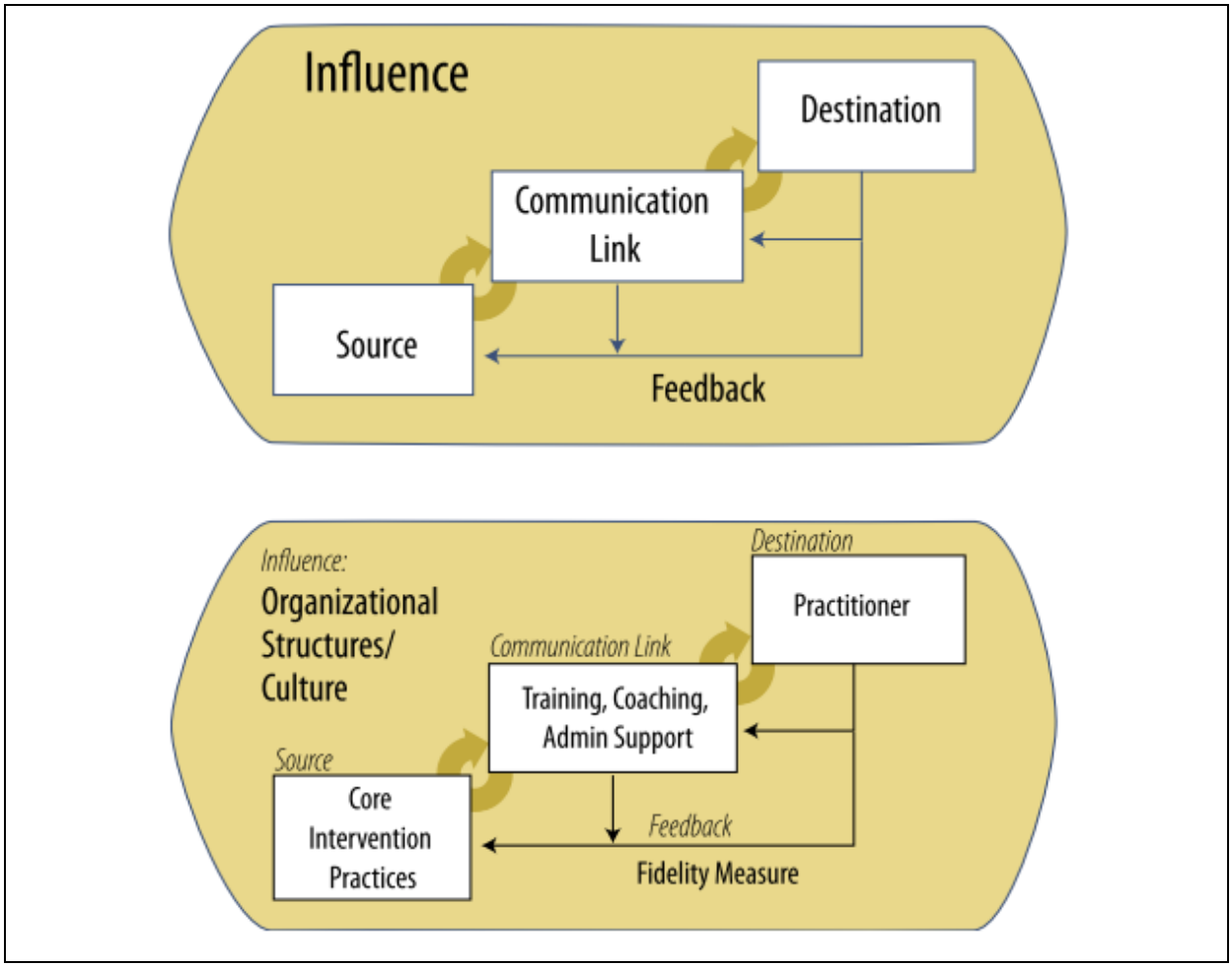


Figure 2. A Conceptual Framework for Implementation of Defined Practices and Programs (Fixen and colleagues, 2005, page 68).

They explain that in its simplest form implementation has five essential components:

1. SOURCE: a "best example," often a composite of the original practice or program that was developed and evaluated and the best features of attempted implementations of that practice or program. Source would present a scientific institution, authors and fist researchers of some kind of program, practice, policy or an innovation in general.
2. DESTINATION: the individual practitioner and the organization that adopts, houses, supports, and funds the installation and ongoing use of an innovation. Destination refers to the micro-community which has chosen to deliver some kind of a program or a practice.
3. COMMUNICATION LINK: an individual or group of individuals, named "purveyors" in Fixsen and colleagues (2005) work, representing a program or practice who actively work to implement the defined program with fidelity and good effect at an implementation site.

4. FEEDBACK mechanism: a regular flow of reliable information about performance of individuals, teams, and organizations acted upon by relevant practitioners, managers, and purveyors.
5. INFLUENCE: connected with social, economic, political, historical, and psychosocial factors that directly or indirectly change people, organizations, or systems. Influence in this model could be interpreted as community level in some other models presented in this chapter.

Fixsen and colleagues (2005; 2009) stress that their conceptual framework is applicable to various fields where evidence-based programs and practices come into contact with practitioners, from manufacturing and human services to health care, agriculture and business. They offer an applied implementation framework to an organization and present most important concepts more clearly (lower part of Figure 2) to analyse their contribution to the stages of implementation, implementation strategies and core implementation components (Figure 3). Core implementation components were already mentioned in the previous section but Figure 3 offers an overview of their relationships.

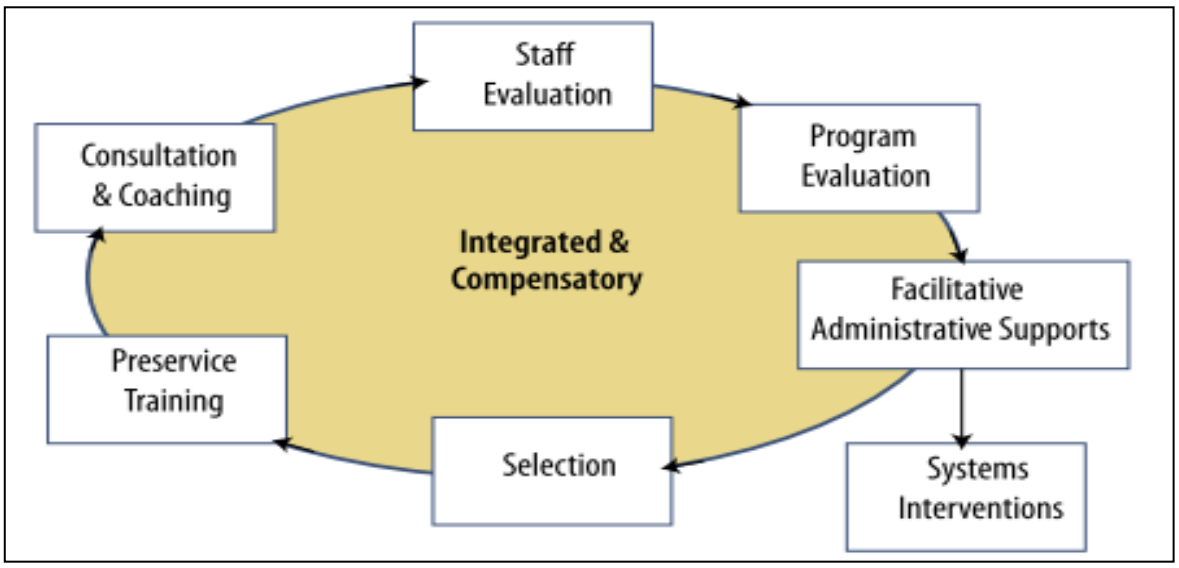


Figure 3. Core Implementation Components that can be used to successfully implement evidence-based practices or practices within evidence-based programs (Fixen and colleagues, 2005, page 29; Fixsen, Blase, Naoom, Wallace, 2009, page 534).

As stated before, based on the commonalities among successfully implemented practices and programs found in their review of literature, Fixsen and colleagues (2005) affirm that the goal of implementation is to have practitioners base their interactions with clients and stakeholders on research findings. To accomplish this, practitioner's behaviour is supported by core implementation components: staff selection, pre-service and in-service

training, ongoing consultation and coaching, staff and program evaluation, facilitative administrative support and systems interventions. **The implementation drivers are interactive and compensate for one another so that a weakness in one component can be overcome by strengths in other components.** These interactive processes are again discussed in the paper of the same authors in 2009 where the authors offer extended analysis of implementation drivers as an attempt to influence science to service delivery and changes in staff behaviour and the organizational culture (Fixsen et al., 2009). Authors stress the fact that practitioners at an implementation site need to learn when, where, how and with whom to use new approaches and new skills. Critical functions of implementation consist of practitioner training, coaching the practitioner on the job, regularly assessing fidelity and using that information to improve the performance of practitioners who are carefully selected for the position (Fixsen and colleagues, 2005, 2009). Authors emphasize that with these core implementation components in place, practitioner behaviour can be routinely changed and improved to assure competent performance of evidence-based practices and programs.

Greenberg and colleagues (2005) try to offer a theory-driven model for implementation studies in a school-based setting. The model differentiates the causative theory inspired by the work of Chen (1999, according to Greenberg et al., 2005) that explains program outcomes from the prescriptive theory that describes how the program should be implemented to reach intended outcomes. Greenberg and colleagues (2005) and Graczyk, Domitrovich, Small and Zins (2006) also base evaluation of implementation quality on both measures of program delivery itself and measures of the support system for training and consultation. In addition, their conceptual model identifies the external influences to the program that may have considerable impact on the quality of program implementation. The same approach can be found in all the models presented in this chapter. The model emphasizes the fact that interventions take place within an implementation system that provides the means and context for delivery of the intervention. Also, that implementation system is embedded within the broader general environment so the implementation system must also be monitored as part of the program evaluation. The shift from former evaluations of implementation quality, which have focused solely on the discrepancy between the program as planned and the program as delivered, in this model comes to the research of discrepancy between the implementation system as planned and the implementation system as delivered (Figure 4).

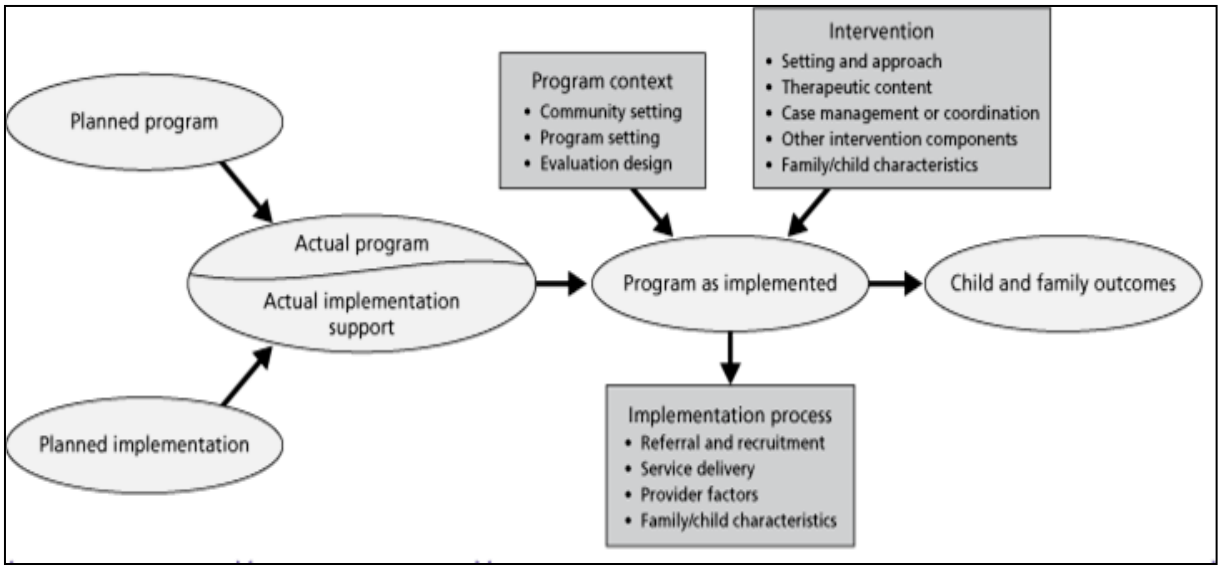


Figure 4.Framework of contextual factors which affect program quality, Greenberg and colleagues (2005), page 19, and Grazcyk, Domitrovich, Small and Zins (2006), page 268.

Wandersman and colleagues (2008) have revised several existing models, including the IOM model of research (Mrazek and Haggerty Eds., 1994), framework of Fixsen and colleagues (2005) and Greenhalgh and colleagues (2004). Wandersman and colleagues (2008) stress that all of the models can be classified as source-based relying on the perspective of the innovation developer, usually scientist (for example, see Fixsen et al., model, 2005) and user based focusing on the awareness of the user that there is the need for change. The same authors argue that the dominant approach until the late 2000s had a research as a starting point and then tried to understand the incorporation of research results into practice. On the other hand, community centred models begin with the world of practice, concentrating on its capacity and asking the community what is needed from science to produce effective interventions. Wandersman and colleagues (2008) appreciate the insights of those two types of models but stress that none of them offers a broad understanding of both perspectives so they developed an Interactive Systems Framework (ISF, left side of the Figure 5). ISF connects three systems which ideally work together for successful dissemination and implementation of effective prevention innovations. ISF has a practical focus on infrastructure, innovation capacities and systems needed to carry out the implementation – the Prevention Synthesis and Translation System, the Prevention Support System and the Prevention Delivery System. Prevention Support System is supporting the work of those who will put the innovations into practice while Prevention Delivery System is implementing the programs in the field. Delivery system is comprised of individuals, organizations and communities while supported by Support System. In a later paper, Meyers, Durlak and

Wandersman, who included 25 frameworks from 27 sources in their synthesis (2012), explain that Support system is building two types of capacities through training, technical assistance and monitoring progress: innovation specific capacity and general capacity regarding organizational functioning. Durlak and Dupre (2008) connect Wandersman and colleagues (2008) Interactive Systems Framework with their multilevel ecological perspective needed for successful implementation, highlighting the role of organization capacity and support that comes from training and technical assistance (right side of Figure 5) which is partly overlapping with the Core implementation component model presented by Fixsen and colleagues (2005, 2009; Figure 3). Durlak and Dupre (2008) hypotheses that implementation is influenced by variables presented in five categories: innovations, providers, communities, prevention delivery system related to organizational capacity and prevention support system, related to training and technical assistance. Under favourable conditions, those variables interact resulting with effective implementation. The main accent of this model is on the capacity which is perceived as community readiness and also as a capacity of organization for conducting some intervention.

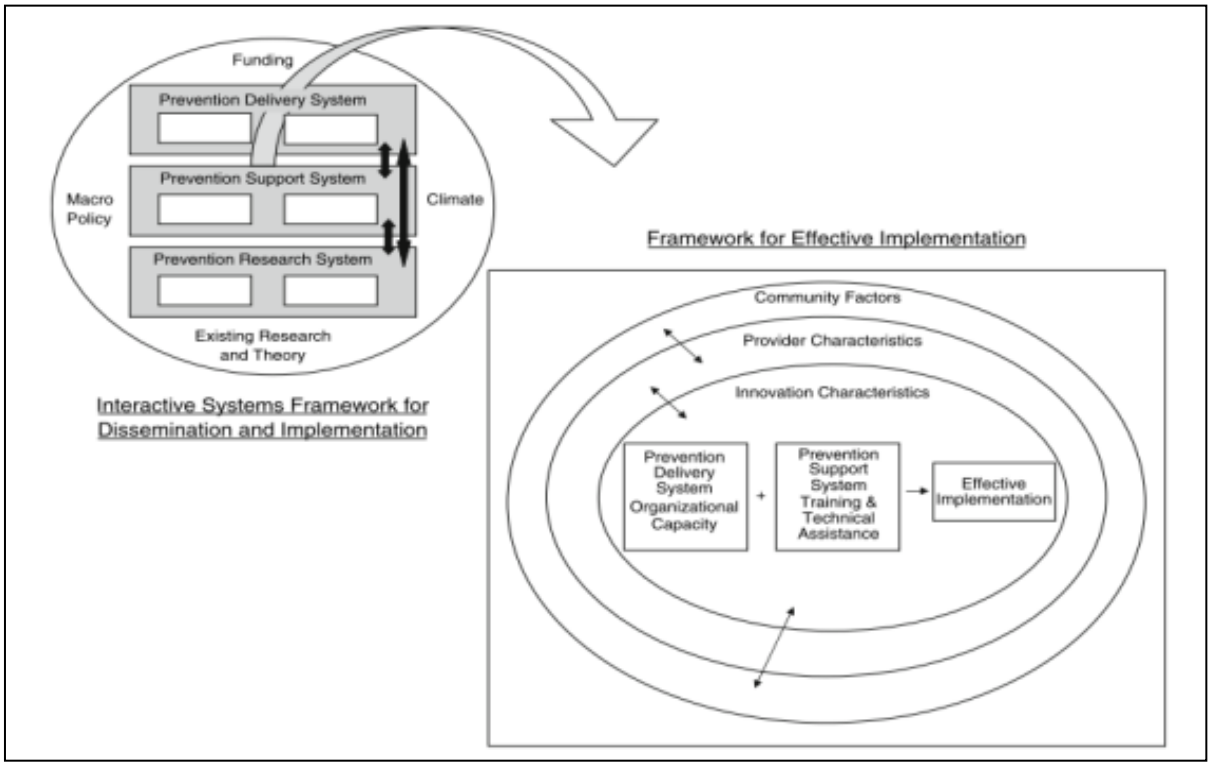


Figure 5.Ecological framework for understanding effective implementation (Interactive Specific Framework, Wandersman and colleagues, 2008, page 174; Durlak and Dupre, 2008, page 335).

Domitrovich and colleagues (2008) presented a multi-level model consistent with the socio-ecological framework offered in Durlak and Dupre's work (2008) but concentrate more

on the implementation quality of school-based interventions. Domitrovich and colleagues (2008) argue that it is necessary to specify which model i.e. implementation theory is used to measure the actual implementation in practice. The multilevel model presented in the Figure 6 takes into consideration the influences of macro-level factors such as policies, school-level factors and individual-level factors. Implementation quality is the outcome of interest, so it is positioned in the centre of the model. In regard to implementation quality, Domitrovich and colleagues (2008) position intervention and a system that supports that intervention as two layers in the centre of the conceptual model. Those two concepts are components of a whole, although they are independent, they affect each other greatly. Both intervention and its support system have to be standardized; their core elements and a delivery model specified (Figure 6). Support system core elements, delivery and standardization are overlapping with the models and information provided by Fixsen and colleagues (2005, 2009) about core implementation components/drivers as well as with Wandersman's (2008) Interactive System Framework and Durlak and Dupre's ecological framework (2008).

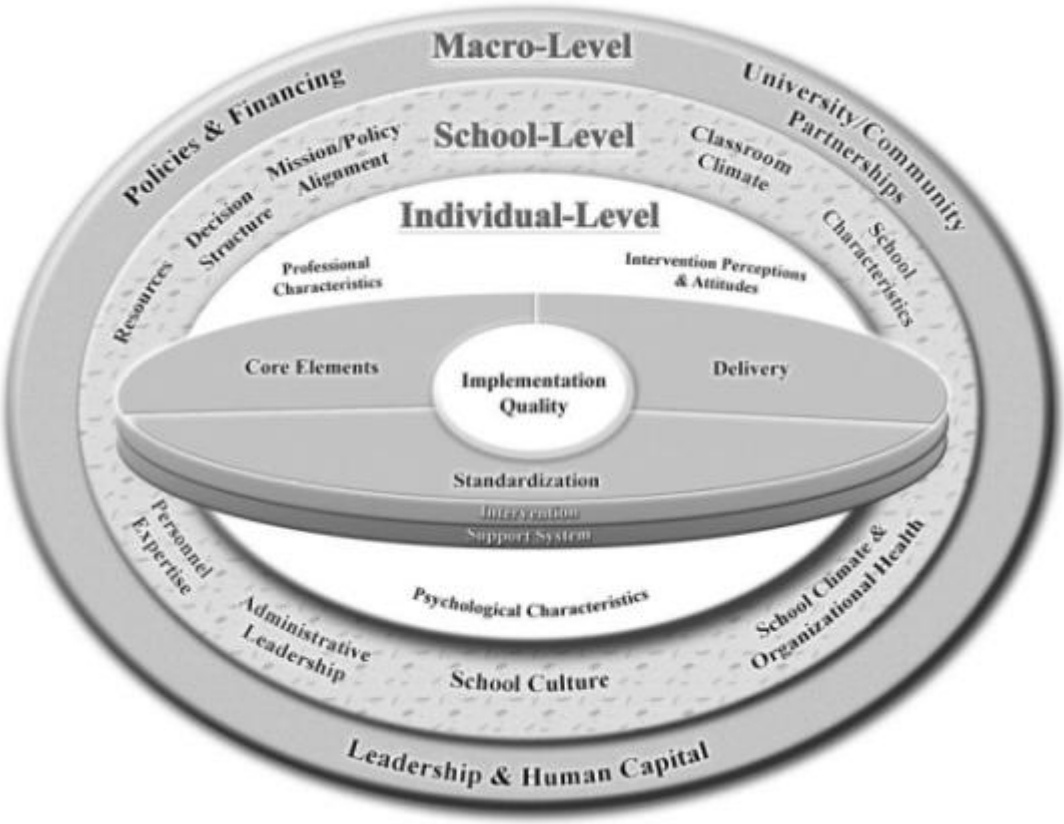


Figure 6. Factors that can affect implementation quality: A Multi-Level models (taken from Domitrovich and colleagues, 2008, page 8).

Setting their implementation framework in school-based interventions, Domitrovich and colleagues (2008) show that intervention is influenced by an array of practices surrounding specific intervention and school. Multilevel influences come from macro level, school level and individual level. Factors at all three levels are co-dependent and have the potential to influence the quality with which interventions are implemented and participants outcomes. The widest, the macro level, can be compared with the community factors stated in the framework of Durlak and Dupre (2008): it also includes community, policy, decision makers, funding, legislative and leadership. The second level of the framework represents the school as an organization in charge of intervention delivery. That is comparable to the Interactive System Framework's (Wandersman, 2008) prevention delivery and support system. School level factors refer to the whole culture in the school, organizational context, school climate, characteristics of the school and classroom as well as to the staff and their experience and training. Current research reviews show that mentioned factors affect not just the quality of implementation but often contribute to the magnitude of outcomes of participants (Gottfredson and Gottfredson, 2002; Kam, Greenberg and Walls, 2003). The third level of Domitrovich and colleagues (2008) framework is an individual level similar to the provider characteristics layer in Durlak and Dupre's model (2008). The individual level comprises professional characteristics (in this model connected with teachers) such as training, skills, experience, psychological characteristics, especially self-efficacy and burnout measures and perceptions and attitudes towards intervention.

The value of the presented models is that they started to incorporate all findings within the field of implementation science and contribute a lot to the better understanding of the processes and defining the theory. Still, all of the presented models have to be incorporated in future research and supported with practical tools for implementation measurement.

1.1.5. Overview of implementation research

As is evident from the previous section, the intent of researches of the implementation field is to identify what it will take to transmit innovative programs and practices to mental health, social services, juvenile justice, education, early childhood education, and substance abuse prevention. All previous researchers agree that thoughtful and effective implementation strategies at multiple levels are essential to any systematic attempt to use the products of science to improve the lives of children, families, and adults. Greenberg and colleagues (2005) state several important reasons for studying and monitoring implementation:

- Effort Evaluation - To know what actually happened
- Quality Improvement - To provide feedback for continuous quality
- Documentation - To document compliance with legal and ethical guidelines
- Internal Validity - To strengthen the conclusions being made about program outcomes
- Program Theory - To examine whether the change process occurred as expected
- Process Evaluation - To understand the internal dynamics and operation of an intervention program
- Diffusion - To advance knowledge regarding best practices for replicating, maintaining, and diffusing the program
- Evaluation Quality - To strengthen the quality of program evaluations by reducing the error in the evaluation.

Kam, Greenberg and Walls (2003) stress that it is very likely that programs show no effect when they are implemented poorly. The same authors also emphasize the other side of the picture: prevention science needs to the question of what is the right context for effective implementation. Durlak and Dupre (2008) continue with the clarification: science tests the potential value of an intervention and gives feedback to the practice. Practitioners from implementation research get a clear answer about the internal and external validity of their interventions, or simpler – which aspects of intervention cause what outcomes. Additionally, accurate interpretation of outcomes depends on knowing what aspects of the intervention were delivered and how well they were conducted. Achieving good implementation not only increases the chances of program success in statistical terms, but can also lead to much stronger benefits for participants and therefore contribute to sustainable outcomes – programs get recognized, respected and well known. While sustained, high quality implementation is essential for achieving greater public health impact if the programs are tested and effective (O'Connell, Boat and Warner, Eds., 2009). At the same time, Durlak and Dupre (2008) add that implementation data are also important in testing the theory behind an innovation.

In the section that follows, an overview of the implementation research and studies conducted in the last two decades will be offered. It can be seen that the research on implementation is evolving gradually as well as understanding the complexity of the field. Meyers, Durlak and Wandersman (2012) address a very important issue in their discussion about the field: because implementation involves studying innovations in real world context, rigorous experimental designs covering all possible systematic variables are almost

impossible to execute. Individual and multiple case studies were so far trying to assess the factors that influence the implementation process in order to learn more about the relationships among factors. Methodological rigor and generalizability of the following research varies, but it is very important to have in mind that these are all pioneer studies.

Issue of implementation measurement

Before research studies are presented, it is crucial to discuss the problem of measurement in implementation field. Schoenwald and colleagues (2010) while describing psychotherapy research refer to the Implementation Methods Research Group (see Table 1.2) which has presented a continuum of quality assurance methods connected with implementation quality research.

Table 1.2

Continuum of quality assurance methods reported in Implementation Methods Research Group literature review (adapted according to Schoenwald and colleagues, 2010).

Rating	Quality assurance method
1	No report of quality control methods or implementation measurement
2	Report on quality control methods only but no measurement (e.g. training, specified manual, ongoing supervision, consultation)
3	Report on some implementation measurement/review, but no specified measure or data reported
4	Measure of implementation reported but no data on reliability or validity measure and no test of implementation quality and outcomes
5	Established measure of implementation quality with established psychometrics used and assessment of relationship between implementation quality and outcomes

While explaining the continuum, Schoenwald and colleagues (2010) stress the importance of defining the purpose of instrument for implementation quality measurement and usage of its scores. The two primary methods of assessing implementation have been **provider self-reports** and **independent behavioural observations**. There are some results that observational data are more likely to be linked to outcomes than self-report data (Lilehoj and colleagues, 2004) but few studies have directly compared these two strategies. From a methodological standpoint, observational methods are preferred because facilitator self-reports may be inflated compared to observer reports (Dusenbury et al., 2005).Adding to this discussion, Schoenwald and colleagues (2010) underline the need for observational approach but highlight the complexity of coding system of conducted observations. Considerable time and expenses are associated with hiring and training observers, developing a protocol,

generating records and turning ordinal and categorical data into interval scale measures. Prevention researchers should seek innovative ways of assessing implementation that are valid and feasible for use in community delivery settings. The level of detail is an important consideration in the design of implementation measures. Besides independent observations, there are non-independent approaches: paper and pencil ratings by program deliverers and participants' self-report on implementation quality.

When deciding about the method of implementation quality assessment, Schoenwald and colleagues (2010) stress the significant implication of available resources and organizational climate, especially time, money, available training and support. Sometimes the context requires different sources of implementation quality data. For example, Cappella, Frazier, Atkins, Schoenwald and Glison (2008) report on the usage of implementation quality checklists in a school mental health service model dealing with children with disruptive behaviour disorders. Cappella and colleagues (2008) have used different referrers: teachers which completed checklist after every professional development meeting, parents reported after parents group and mental health providers reported on the content and perceived quality of supervision. Teachers and parents in this study reported both on the frequency and the content of support provided by the mental health team as well as their own use of recommended tools and strategies.

A final important issue relates to the timing of implementation assessments. Many studies code a random sample of program sessions, assuming that the results generalize to all sessions and that implementation has consistent effects across all program components (e.g., Lochman et al., 2009). A more refined approach is to assess outcome-specific implementation. This approach involves selecting a proximal or distal outcome and assessing implementation for the specific program components that target the selected outcome.

Domitrovich and Greenberg (2000) emphasize the importance of qualitative data assessment, especially recording discussions with program developers and school consultants which reveal a great personal experience and accumulated wisdom which shouldn't be neglected. The same authors suggest that all implementation research should begin with specifying program components and active ingredients to reveal a theory behind an intervention. Chen (1998, according to Domitrovich and Greenberg, 2000) argues that some attention has also to be paid to the implementation system and that research has to include

factors such as characteristics of implementers, the nature of the implementing organization, and the quality of the linkages between this organization and the broader community.

Berkel and colleagues (2011) also discuss challenges of implementation quality research. They address the issue of confound measures: fidelity and adaptation, fidelity and quality and quality and responsiveness. While some measurement work needs to be conducted to distinguish conceptually different dimensions of implementation, efforts also need to be undertaken to understand the best way to combine indicators of these dimensions. Rohrbach and colleagues (2007) measured fidelity of Projects Towards No Drug Abuse program in the way that they wanted to assess the program process. They observed each program implementer delivering the same two highly interactive curriculum sessions to two separate classes in each experimental school. They observed 90 sessions where two observers were present and 27 sessions where only one observer was present (n=207). Their observation instrument assessed adherence, quality of delivery, classroom process, and student acceptance of the curriculum. The specific items were adapted from previous studies and were combined to create four indexes. Besides observations, Rohrbach and colleagues (2010) developed a self-report measure for students where they were asked about overall associations about the program (program acceptance index) but also rated implementers confidence, ability to elicit participation, understanding and respectfulness towards students (evaluation of implementer index). The overall program rating index averaged how much students liked each of the 12 curriculum lessons.

O'Donnell (2008) highlighted the distinction between structural implementation (focusing on the amount of instruction or number of lessons provided) and process implementation (focusing on delivery of key procedural features of the curriculum and quality of delivery). Leaning on that differentiation, Odom and colleagues (2010), while researching early child curriculum, designed implementation measures to capture structural and process features of implementation. Site supervisors visited classrooms every week and measured number of lessons teacher completed during the school year. Implementation quality was assessed with rating scales which were introduced to site supervisors seven times during the year in each classroom at approximately same time points during curriculum delivery.

Flaspohler and colleagues (2012), while reporting on structures and process of a Prevention Support System at Miami University, describe the School-wide Implementation Checklist, a comprehensive implementation fidelity measure. The Checklist was designed to

assess several domains of implementation including buy-in from administrators and educators, participation in program training, procurement of program materials and implementation of core components. Implementation fidelity items were rated on a three-point Likert scale: "completed" which got two points, "making good progress" which got one point and "progress needed" which got zero points. Implementation fidelity percentages were calculated by dividing the total number of implementation fidelity points achieved by the total number of points possible and converting the scale to percentage.

Mihalic, Fagan and Argamaso (2008) in the study of Life Skills Training created a measure of adherence, dosage, quality of delivery and participant responsiveness. Classroom observers of adherence to curriculum have followed 4 of the 15 classroom sessions taught during LST level one, 3 level two lessons and two level three lessons. During each observation, the proportion of objectives and activities taught was identified using a fidelity checklist. An implementation score for each lesson taught was calculated as a percentage of material taught of all required material. Observers were also asked to identify the use of varied techniques, student participation and activity as well as to note if there were any problems such as deviations from the curriculum or student behaviour issues. Quality of delivery in this Mihalic and colleagues (2008) study was assessed as the percentage of the class period spent using interactive teaching techniques such as discussion, skill demonstration and behavioural rehearsal. A summary score was created for each site, based on all site observations in a three year period. Participant responsiveness was measured by means of a teacher self-report survey which was administered in the end of each implementing year.

Proctor and colleagues (2010) define implementation outcomes as the effects to deliberate and purposive action to implement new treatments, practices and services. They serve as indicators of implementation success; they are proximal indicators of implementation processes and are key intermediate outcomes in relation to service system outcomes. Fixsen and colleagues (2005) state that essential implementation outcomes which should be covered in research are:

a. Changes in professional behaviour – knowledge and skills of practitioners and other key staff members within an organization or system,

b. Changes in organizational structures and cultures, both formal and informal which support the changes in professional behaviour – values, philosophies, ethics, policies, procedures, decision making,
c. Changes in relationships to consumers, stakeholders and system partners – location and nature of engagement, inclusion and satisfaction.

Research overview – implementation outcomes vs. program outcomes

Dane and Schneider (1998) conducted a literature review of prevention programs published between 1980 and 1994. They found that only 39 of 162 outcome studies documented the implementation and only 13 used a measure of fidelity as a variable when analysing the results. Although 57 percent of the studies reviewed provided manuals and training, ongoing consultation was provided in less than half of the studies. Only 20 percent of the outcome evaluations used "comprehensive integrity promotion," which included a program manual, formal training, and ongoing consultation. Dane and Schneider (1998) further examined which dimensions of program integrity and dosage were analysed in relation to outcomes. The results confirmed that particularly adherence and exposure played a significant role in explaining outcomes. In some studies, positive outcomes were evident only when a specific proportion of the program content was provided. Intervention effects were found most often when trained observers, rather than service providers, were the source of information. The authors noted that the variability in the sources and aspects of integrity reported limited their ability to draw firm general conclusions about the effect of implementation on program outcomes (Dane and Schneider, 1998).Dusenbury and colleagues (2005) examined several hundred outcome studies covering a 25-year period of drug prevention research but briefly summarized data from only nine reports providing information on relationship between implementation and outcomes. Derzon and colleagues (2005) assessed findings from 46 unpublished drug prevention programs funded by SAMHSA. They have found that factors with strongest effects on outcomes were related to implementation.

Domitrovich and Greenberg (2000) researched 34 programs proven effective and noted that only 13 studies conducted analyses relating implementation to outcomes. The review included programs that produced improvements in specific psychological symptoms (e.g., aggressive behaviour and anxiety) or in factors directly associated with increased risk for child mental disorders (e.g., poor parenting skills or a history of early child behavioural problems). Programs were included if they were evaluated using either a randomized-trial

design or a quasi-experimental design that included a comparison group. Studies were required to have both baseline and post-intervention findings and, ideally, follow-up data to examine the duration and stability of program effects. In addition, programs were required to have a written manual that specified the model and procedures used in the intervention. Only 34 programs met all of these criteria and were subsequently classified as effective and included in the review. Domitrovich and Greenberg (2000) have found that only 7 from 34 programs have researched more than one implementation dimension. Greenberg and colleagues (1999) then classified the 34 programs using a system based on the work of Dane and Schneider (1998). All were examined for specific features related to program integrity. As in Dane and Schneider (1998), a distinction was made between strategies that promote integrity (e.g., manual and staff training), and procedures that verify integrity (e.g., monitoring adherence and dosage). All 34 programs promoted integrity to some degree. Overall, 26 of the 34 effective programs verified program integrity in some way. Twenty programs included some rating of fidelity or adherence in their implementation data, which, for the majority, involved tracking the program's essential components with ratings made by independent observers or program implementers. In three of these 20 studies, fidelity was assessed indirectly. High fidelity was assumed when a significant difference was found between program participants and controls along a behavioural dimension targeted by the intervention (e.g., teacher practices or student perceptions). Although this method provides important information, it cannot verify that the behavioural changes were not due to one or more factors unrelated to the intervention. Regarding other dimensions of implementation, dosage was reported in 33 percent of the studies. Four programs assessed participant responsiveness, and two programs assessed program differentiation, i.e. the degree to which participants in each condition actually differed only with respect to whether or not they received the treatment intervention.

Durlak and Dupre's review (2008) of nearly 500 individual studies and meta-analyses has shown that only 59 studies assessed the relationship of fidelity and outcomes. Of these 59 studies, 76% report that fidelity had a significant positive association with targeted program outcomes. Durlak and Dupre (2008) identified that programs had visible positive results with at least 60% of fidelity coverage. Spoth and colleagues (2002) reported about null relations between fidelity and intervention outcomes. Among those 59 studies which examined fidelity, only 6 of them examined effects of quality of delivery on program effects. Few of the studies have found a positive relationship, for example, quality of process used to deliver family-

based programs predicted improvement in parenting skills. Quality of delivery has also been associated with decreases in adolescent substance use (Kam et al., 2003).

Meyers and colleagues (2012) report on the development and research of the Quality Implementation Tool (QIT) which derived from the Quality Implementation Framework of Meyers, Durlak and Wandersman (2012). QIT is a worksheet format that relates with action steps (Meyers and colleagues, 2012) which have to be completed in a collaborative process of Support and Delivery System members already mentioned in sections before (Wandersman et al., 2008). It covers six practical components: **develop an implementation team, foster supportive organizational climate and conditions, develop an implementation plan, receive training and technical assistance, practitioner-developer collaboration in implementation and evaluate the effectiveness of the implementation.** All of those components are followed by action steps which can be assessed in paper in three different purposes: a) to plan for quality implementation, b) real-time monitoring of implementation and c) for evaluating the extent to which the innovation was implemented with quality. Meyers and colleagues (2012) report that QIT was piloted in two different human services projects: the psychological services centre and at an individualized treatment protocol for pregnant substance-abusing women. After the tool piloting, Meyers and colleagues (2012) conclude that QIT can be used to help the Support and Delivery Systems to work more collaboratively. They have found that this tool raised awareness of steps needed for quality implementation as well as it helped in identifying capacity limitations that inhibit implementing with quality. They conclude that the Quality Implementation Tool serves as an active in-depth planning work-sheet that is easily shared within stakeholders and is organized so that the content in the tool can document progress, barriers and group decisions.

1.2. Mental health promotion and prevention in Croatia

1.2.1. Development of mental health promotion and prevention in Croatia

Mental health promotion and prevention science in Croatia has a short history and was mostly developed during the last twenty years within the University of Zagreb's Faculty of Education and Rehabilitation Sciences. Before that period, historically speaking, scientists in Croatia, especially on the Faculty of Education and Rehabilitation Sciences, were studying criminology, juvenile delinquency and behavioural disorders of children and youth mostly

from the aspect of treatment of mentioned phenomenon. The shift away from treatment towards prevention approach began during the nineties with the efforts of a group of scientists from the Faculty of Education and Rehabilitation Sciences and the project "Integral Method" (Bašić, Koller Trbović, Žižak, 1993). The project was developed for parents of kindergarten children and pre-school teachers and designed to establish a foundation of universal prevention and positive development for children and youth in Croatia. The "Integral Method" was one of the first initiatives to set the basis for social and emotional development of children in Croatia through collaboration of diverse scientists, experts and practitioners.

While reviewing the history of prevention in Croatia, it is also important to recognize the activities of the National Council for Children as well as of the Government Commission for Prevention of Behavioural Disorders of Children and Youth. These two boards gathered influential members of governmental institutions and researchers across academic disciplines who were interested in family, children and youth issues. The Government Commission for Prevention of Behavioural Disorders of Children and Youth was constituted in 1997 (first president prof.dr.sc. Josipa Bašić) and was composed of representatives from different Ministries, the State Attorney's Office and Judiciary practice, the Institute for Family, Motherhood and Youth, as well as scientists from the fields of interest. During the 2000s, the Government Commission supported the development of mental health promotion and prevention field with several important publications regarding risk and protective factors and community prevention: "Risk and protective factors in the development of behavioural disorders of children and youth" (Bašić, Janković, Eds, 2000), "Prevention of behavioural disorders of children and youth in local community" (Janković, Bašić, Eds, 2001) and "Local community: source of national strategy of prevention of behavioural disorders of children and youth" (Bašić, Janković, Eds, 2003).

Prevention science and mental health promotion in Croatia is as well strongly influenced by the public health sector which has a long and prosperous tradition, especially because of the worldly renowned School of Public Health Andrija Štampar. Many mental health promotion and prevention activities in local communities were conducted within the World Health Organization initiative of Healthy Cities linked to the same School of Public Health. The Croatian Healthy Cities Network was established in 1992, although the initiative was present from 1988 which makes it one of the oldest Healthy Cities network in Europe. The main intention of this network is to promote healthy behaviours among community

members through different initiatives aimed at health of children, youth and families, community development, healthy life choices, urban planning as well as physical health.

Systematic education of future experts in the field of prevention of behavioural problems and promotion of mental health at academic level is also significant for the development of this field in Croatia. The Faculty of Education and Rehabilitation Sciences of the University of Zagreb within the Department of Behavioural Disorders provides continuous education of future social pedagogues in the field of prevention of behavioural problems. Prevention experts at the Faculty are delivering prevention courses at the undergraduate and graduate level preparing future social pedagogues for implementation of science-based prevention practice. Based on more than a decade of investments in the field of mental health promotion and prevention, in 2007 the University of Zagreb's Faculty of Education and Rehabilitation Sciences has initiated an international doctoral programme "Prevention science: prevention of mental and behavioural disorders and promotion of mental health". The doctoral programme has been initiated by Josipa Bašić, regular professor of prevention of behavioural disorders at the Department of Behavioural Disorders of the Faculty of Education and Rehabilitation Sciences, University of Zagreb. It's a unique doctoral program in this part of Europe which is gathering prevention experts from Croatia and from abroad who are empowering future prevention scientists in Croatia with the most recent knowledge and skills in this field. In 2011, a second generation of doctoral students was enrolled into this unique doctoral programme.

Since 2000, Croatian prevention scientists established collaboration with scientists and centres for prevention science worldwide, especially with the Prevention Research Centre of Penn State University (USA) and the Prevention Research Centre of the Radboud University Nijmegen and Maastricht University (the Netherlands). They also started a collaborative relationship with prevention scientists from Scuola Universitaria Professionale della Svizzera Italiana (Switzerland) and other relevant institutions from abroad. Croatian prevention experts have attended Society for Prevention Research conferences annually and the EU SPR conference, have participated in the European IMHPA project on mental health promotion (Jané-Llopis and Anderson, 2005). Collaboration with foreign mental health promotion and prevention experts enabled knowledge and expertise exchange and it encourages international projects which has significant influence on mental health promotion and prevention science in Croatia.

It is important to emphasize that several prevention "model programs" acknowledged worldwide were adopted, implemented and researched in Croatia until this very moment. One of the model programs, the Communities that Care model has being implemented in the Croatian County of Istria since 2002 (Bašić, Ferić Šlehan, Kranželić Tavra, 2007a and 2007b; Bašić, Grozić-Živolić, Eds, 2010). The Northland project was implemented in the city of Split during 2002 within the context of the international Healthy cities network, the Life Skills Training in the city of Rijeka from 2005 and the PATHS model program on socio-emotional learning in the County of Istria, Zagreb and Rijeka since 2008 (Bašić, Grozić-Živolić, Eds, 2010).

1.2.2. Core problems of mental health promotion and prevention field in Croatia

Although a review of the history of prevention in Croatia suggests that there are a lot of initiatives going on in such a small country, it is still evident that science-based prevention practice in Croatia is still in its roots and is facing a lot of bottlenecks (Bašić, 2009).

There are several general dimensions of existing problems in this field:

- lack of science based approach to the promotion of mental health and prevention of behavioural problems,
- lack of coordination between institutions, stakeholders and activities concerned with promotion of mental health and prevention of behavioural problems and
- lack of consistent implementation of existing law regulations and policies concerning the well-being of children, youth and families into practice.

Bašić (2009) has emphasized the strong need for using scientific knowledge and systematic approach in organizing, developing, implementing and evaluating prevention interventions and initiatives in Croatia. In general, there is a lack of evidence-based programs widespread in every local community in Croatia. Local and national authorities often do not demand any evidence of quality assurance or evidence of program effectiveness and mental health promotion and prevention programs are mainly initiative of non-governmental organizations. Croatian mental health promotion and prevention programs are mostly run by local practitioners non-trained in prevention and not familiar with science-based approach to mental health promotion and prevention. Prevention programs are rarely theory-based and their outcomes are often not evaluated (Bašić, 2009; Bašić, Mihić and Novak, 2011). An additional problem is that the interventions are often not attuned to the specific needs of the population.

Also, coordination between institutions, stakeholders and activities concerned with mental health promotion and prevention is not strong enough. Croatia doesn't have an active "umbrella" institution which takes care of policies and interventions of mental health promotion and prevention (Bašić, 2009; Bašić, Mihić and Novak, 2011). The country lacks a national database for evidence-based prevention programs similar to the ones in the United States, the Netherlands, and Norway. Deficiency in national coordination of prevention in Croatia has resulted in partial prevention initiatives of some national departments such as the Ministry of Science, Education and Sports, Ministry of Health and Social Welfare or initiatives of local authorities. Various mental health promotion and prevention interventions get remarkable financial support from local and state agencies, but they are treated as single and incidental actions rather than a part of a more comprehensive strategy.

The problem of weak coordination is closely connected with non-consistent implementation of law regulations and policies concerning the well-being of children, youth and families into a practice. Even though high quality law regulations and policies exist (for example, National strategy for prevention of behavioural problems of children and youth, 2009), they are not implemented very effectively and they have a narrow reach. The reason for that is that division of tasks between various institutions in this field is not transparent enough and it often remains unknown who is responsible for what. One of the reasons for this condition is surely connected with a lack of an infrastructure for mental health promotion and prevention.

If we consider all of the mentioned problems, there are some possible approaches which could improve the state of this field in Croatia. It is evident that there is a strong need for investment into knowledge of developers and implementers of the interventions of mental health promotion and prevention. Intervention developers and implementers should be much more aware of advantages of science based practice and continuously trained for incorporation of science based principles into their practice. A systematic investment into knowledge of mental health promotion and prevention intervention developers and implementers is the first step in improving the effectiveness of this filed. At the same time, there is a strong need for promoting evaluation and carrying out evaluation studies in Croatia. Assessment of interventions' impact and effectiveness, as well as investments in implementation research will lead to the detection and dissemination of best practices on the one side and improvement of current interventions on the other side.

1.3.Research project “Preffi - Quality Assurance in the County of Istria”

1.3.1. The case of Istria

To develop mental health promotion and prevention in a country with poor national infrastructure and governance in that field, the research team from the Faculty of Education and Rehabilitation Sciences of the University of Zagreb concentrated their efforts in a geographical area which would offer the best perspective on success. Since 2002, the research team from the Faculty started with the process of developing a national laboratory for research, policy making, implementation and quality assessment in the County of Istria. It is one of the most developed and relatively more prosperous regions in Croatia. Members of the research team from the Faculty realized that conditions for investment and development were more favourable in that region and that the outcomes of learning experience can serve in future initiatives of developing mental health promotion and prevention on a nation-wide scale.

The Faculty of Education and Rehabilitation Sciences of the University of Zagreb started a long-standing collaboration with local authorities of the County of Istria in 2002. The Faculty was supported by the Istrian Department of Health and Social Care in running the before mentioned project “Communities That Care: Development of a Model for Behavioural Disorders Prevention" (Bašić, Ferić Šlehan, Kranželić Tavra, 2007a and 2007b; Bašić, Grozić-Živolić 2010). The idea of this project was to experiment with applying science-based principles of prevention in real life settings and to invest in the development of an organizational system for prevention. Special attention is given to capacity building in non-professionals dealing with children and youth, using a combination of bottom-up and top-down principles. During the CTC project, important steps were taken to improve the quality of mental health promotion and prevention practice in the County of Istria. These include: (1) assessment of readiness for mental health promotion and prevention, (2) needs assessment, (3) setting of mental health promotion and prevention priorities, (4) systematic identification of resources, (5) implementation of mental health promotion and prevention programs according to the defined needs and (6) evaluation of those programs and whole projects.

The Department of Health and Social Care (DHSC) strives to systematically develop prevention by financing programs of nongovernmental organizations provided by local practitioners. In 2002, the DHSC started to develop a systematic procedure for allocating these funds and financed 11 prevention projects. Initially, the criteria for financing were that

proposals should provide a solution to a particular problem and that the proposed services had to be broadly offered within the community (Bašić, Ferić Šlehan, Kranželić Tavra, 2007a; internal materials from Department of Health and Social Care, County of Istria, 2010). In 2004, 27 projects were financed and this expanded to 32 projects in 2006. Over time as the demands for quality, accountability and sustainability grew; this approach to funding was seen as insufficient. To improve the ability to select programs with the best potential for success, the DHSC began including other criteria to the selection process as advised by the Faculty of Education and Rehabilitation Sciences. This included: 1) clear and specific program goals, 2) firm organizational structure for program implementation, 3) partnership with other organizations in the community, and 4) the involvement of volunteers (internal materials from Department of Health and Social Care, County of Istria, 2010).

Through the Departments' collaboration with the Faculty of Education and Rehabilitation Sciences and international research centres, new and expanding research on program effectiveness and knowledge on success factors became visible. The leaders within the DHSC realized that the criteria against which they judged applications for funding needed further improvement by better attuning them to the current scientific knowledge on effect predictors. At the same time, intervention developers and deliverers showed to have been responsive to the earlier quality criteria included in the previous evaluation system for grant allocation. The research team from the Faculty of Education and Rehabilitation Sciences and the Department of Health and Social Care of the County of Istria noticed that there is a need for developing a quality assessment tool for the mental health promotion and prevention programs, a need for improvement of the mental health promotion and prevention programs' quality and a need for evaluation of mental health promotion and prevention programs' effectiveness in the County of Istria.

1.3.2. Project overview and doctoral research studies

Based on the detected needs, in 2010 the Department of Health and Social Care decided to continue the collaboration with the Faculty of Education and Rehabilitation Sciences, University of Zagreb in improving existing mental health promotion and prevention practice in the County of Istria. The research team from the Faculty of Education and Rehabilitation Sciences developed a research project named »Preffi – Quality assurance in the County of Istria« (project team: professor Josipa Basic, PhD, Miranda Novak, M.A., Josipa Mihic, M.A.) as a sub-project within the project "Communities That Care: Development of a

Model for Behavioural Disorders Prevention". The general aim of this sub-project was to decrease mental and behavioural problems of children and youth in the Istria region through the incorporation of evidence based principles into the mental health promotion and prevention practice. In order to achieve this long-term goal, aims of the project were:

1. To enhance the knowledge and capacities of NGO's leaders, program directors and deliverers, financed by the County of Istria, about the principles of science-based practice by providing them "Training for Prevention",
2. To improve the quality of written proposals of mental health promotion and prevention projects proposed by NGO's in Istria,
3. To improve the outcomes of mental health promotion and prevention projects financed by Department of Health and Social Care, County of Istria,
4. To create science-based criteria for financing mental health promotion and prevention programs in the County of Istria through incorporation of effect predictors into financing criteria.

Main assumption of the project was that the incorporation of evidence-based principles is crucial in improving the quality and effectiveness of mental health promotion and prevention practice in the County of Istria. For that reason, the project team had an intention to encourage and prepare the Department of Health and Social Care to be oriented towards implementation of evidence-based policy in their region. Evidence-based policy has been defined as an approach that helps people make well informed decisions about policies, programs and projects by putting the best available evidence form research at the hearth of policy development and implementation (Davies, 1999, according to Shaw, Green and Melvin, (Eds), 2007). This approach stands in contrast to opinion-based policy, which relies heavily on either the selective use of evidence or on the untested views of individuals or groups, often inspired by ideological standpoints, prejudices or speculative conjecture. Gray (1997, Shaw, Green and Melvin, (Eds), 2007) has suggested that there is a new dynamic to decision making in mental health promotion and other areas of public policy, whereby the speculation of opinion-based policy is being replaced by a more rigorous approach that gathers, critically appraises, and uses high-quality research evidence to inform policy-making and professional practice.

Considering the mentioned reasoning, DHSC was open to use a more comprehensive set of research-based quality criteria to guide their funding decisions which led to a subproject

titled "Preffi – Quality assurance in the County of Istria". This project included 24 mental health promotion and prevention programs proposed by NGOs and financed during 2011 and 2012 by the Department of Health and Social Care, County of Istria. Programs were assessed with an instrument called the Preffi 2.0 instrument (see Appendix 6, page 233) which assesses factors associated with the effectiveness of prevention programs based on the presence and quality of a set of research-based effect predictors (Hosman, 1994; Hosman, 2008, Peters et al., 2003; Molleman, 2005, Molleman et al., 2005). After the appraisal with Preffi, programs were divided into matched pairs, and then assigned to an intervention or control group. After the first assessment of programs with the Preffi 2.0 instrument, program leaders (i.e., managers and implementers) from the experimental group participated in the Training for Prevention intervention.

The Training for Prevention intervention was developed within the project "Preffi – Quality Assurance in the County of Istria" by Josipa Mihic and Miranda Novak, two young researchers from the Faculty of Education and Rehabilitation Sciences, University of Zagreb, conducting their doctoral research within this project. Before the initiation of this project in 2008, the authors of the Training for prevention were included in the Department of Health and Social Services' committee for the projects' appraisal in the County of Istria. The role of the committee was to evaluate the quality of the written project proposals and to assess whether proposed projects involve sufficient level of effect predictors. Within the evaluated projects, overall weaknesses and gaps were identified. During several years of that experience, Training's authors realized that all projects have similar difficulties in transferring science-based principles into a practice. It became evident that the programs' authors and deliverers come from different professional backgrounds and have poor knowledge on promotion and prevention. The weakest elements of the written project proposals were problem analysis, target group description, theory behind their programs, connection of goals, activities and expected program outcomes, as well as evaluation design. The theoretical concept of the Training is based on:

- Knowledge and research on effect predictors (Hosman, 1994; Hosman and Engels, 1999; Raphael, 1999; Kok, Van den Borne and Dolan Mullen, 1997; Tobler and Stratton, 1997; Brown and colleagues, 2000; Jane-Llopis and Barry, 2005; Stice et al, 2009),
- Theory of planned behaviour (Ajzen, 1991),
- Transtheoretical model (Prochaska, Redding and Evers, 2002)

- Intervention mapping approach (Bartholomew, Parcel, Kok and Gotlieb, 2006),
- Preffi 2.0 instrument (Peters et. al, 2003, Molleman, 2005¸ Molleman et al., 2005; Molleman et al., 2006).

Within the Training, mentioned theoretical concepts were used on three levels:

1. the transfer of knowledge about the process of change to Training's participants
2. the incorporation of effect indicators in participants' programs
3. developing the skills of Training's participants needed for initiating the process of change in their target groups

The primary aim of the Training for Prevention intervention was to enhance the knowledge, skills and capacities of target group about the principles of science-based prevention practice. This is based on the hypothesis that incorporating these principles into programming will result in better understanding of theory and logic model of the program, improvement of quality of written project proposals of mental health promotion and prevention programs. This should lead to higher implementation quality and better behavioural and mental outcomes (presented in the Model of Training in the Figure 7.).

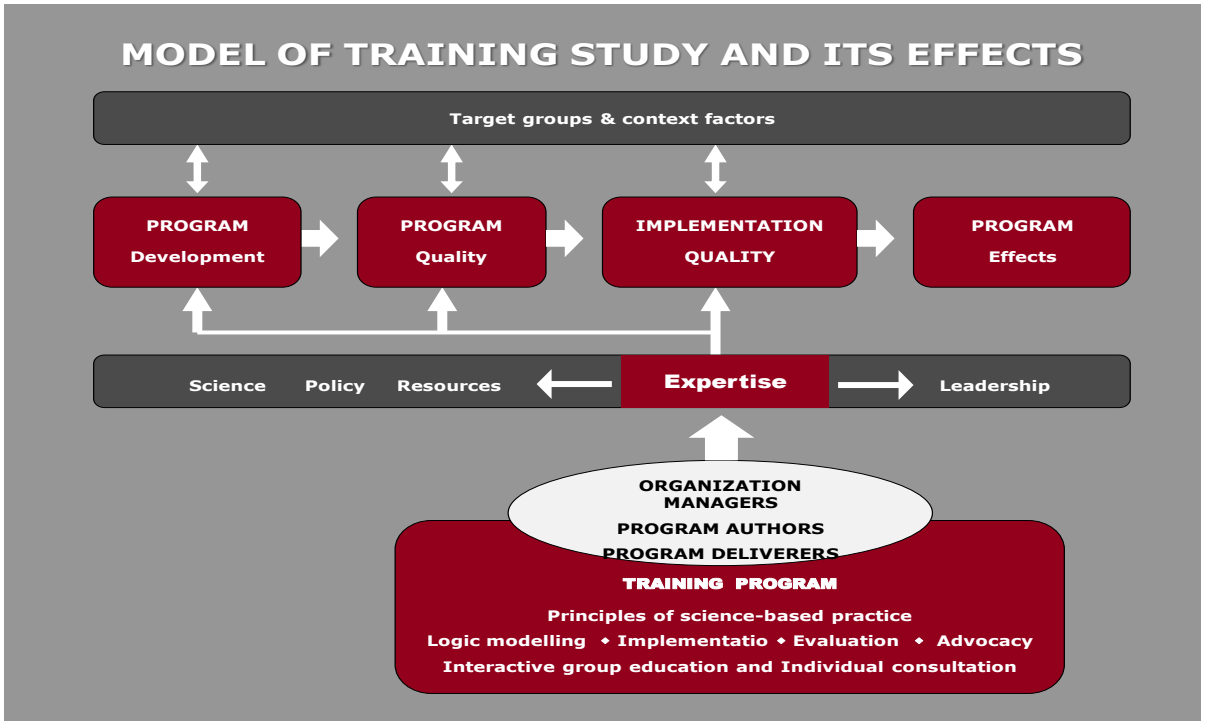


Figure 7. Model of Training for Prevention and its effects.

From the overall project "Preffi – Quality assurance in the County of Istria" activities two doctoral research studies were designed:

- *Study of effectiveness of prevention programs* (doctoral candidate: Josipa Mihić, Faculty of Education and Rehabilitation Sciences, University of Zagreb, mentor: Clemens Hosman, PhD, Radboud University of Nijmegen and Maastricht, the Netherlands) and
- *An empirical study on implementation quality in prevention programs* (doctoral candidate: Miranda Novak, Faculty of Education and Rehabilitation Sciences, University of Zagreb, mentors: Clemens Hosman, PhD, Radboud University of Nijmegen and Maastricht University, the Netherlands and Celene Domitrovih, PhD, Pennsylvania State University, USA).

Study of effectiveness of prevention programs is the doctoral study the main aim of which was to adapt the Preffi 2.0 quality assessment instrument and to assess some of its metric characteristics. Other aims of the study were to measure the Training for prevention impact on the effectiveness of programs of mental health promotion and prevention in achieving desired outcomes and to measure the Training for prevention impact on the quality of written projects' proposals. Assumption of the study was that education of NGO's leaders, program directors and deliverers about effect predictors can improve the effectiveness of their programs and quality of programs' written proposals. Also, predictive validity of the Preffi 2.0 instrument in predicting the effectiveness of programs will be assessed within this study. The results of this research study will provide insight into the quality of NGO's programs of mental health promotion and prevention involved in a study. Study results will offer suggestions for creating a science-based mental health promotion and prevention practice in the County of Istria and directions for improvement of existing criteria for assessing programs' quality.

The study *An empirical study on implementation quality in prevention programs* is presented in this monograph. The main aim of this dissertation was to monitor the overall level and variability of implementation quality of 24 mental health promotion and prevention programs. In order to achieve that, two new measures of implementation factors and two new measures of implementation quality were constructed, relying on the literature and trends in mental health promotion and prevention science. These are measures for monitoring the quality of implementation of programs from the position of organization manager, from the position of program implementers and from the position of program participants. Also, this study strives to explore the relationship among implementation factors and indicators of

implementation quality. This study attempts to answer the question whether Training for Prevention which was delivered to an experimental group improve the factors that influence implementation and the overall quality of implementation in the programs. Study results will offer contributions and suggestions for implementation research field in general, especially regarding the interventions whose effectiveness still has to be proved and will contribute to the efforts of quality improvement of mental health promotion and prevention in the County of Istria.

1.4. Conceptual framework for the research of implementation quality in the County of Istria

For the purpose of this doctoral research project and as a contribution to mental health promotion and prevention science, a conceptual model of factors affecting implementation quality has been designed based on the research literature and theory presented in the previous section. This model incorporates ecological, individual, and intervention factors described in the work of Fixsen and colleagues (2005, 2009), Domitrovich and colleagues (2008), Durlak and Dupre (2010), and Durlak (2010) that are necessary for a successful implementation process. The definition of the key outcome in the model, implementation quality, is based on the literature review presented in the sections above. As seen in Figure 8, the conceptual model includes factors at two levels which are represented on the left hand side of the figure in a series of embedded boxes. The first level reflects the capacity of an organization to support the implementation of programs. This includes providing adequate training and knowledge to staff, supporting the individuals who deliver the program, attitudes towards the program in general and monitoring the process of program implementation. The second level of predictors reflects the characteristics of the program itself. These include the skill of program implementer, implementer's attitudes towards the intervention and program standardization which in this study included how standardized the intervention was perceived to be. These factors are chosen among many which are presented in the literature review as the ones that are most often repeated. Also, the chosen factors are according to experience of researcher the ones most important for Croatian mental health promotion and prevention practice.

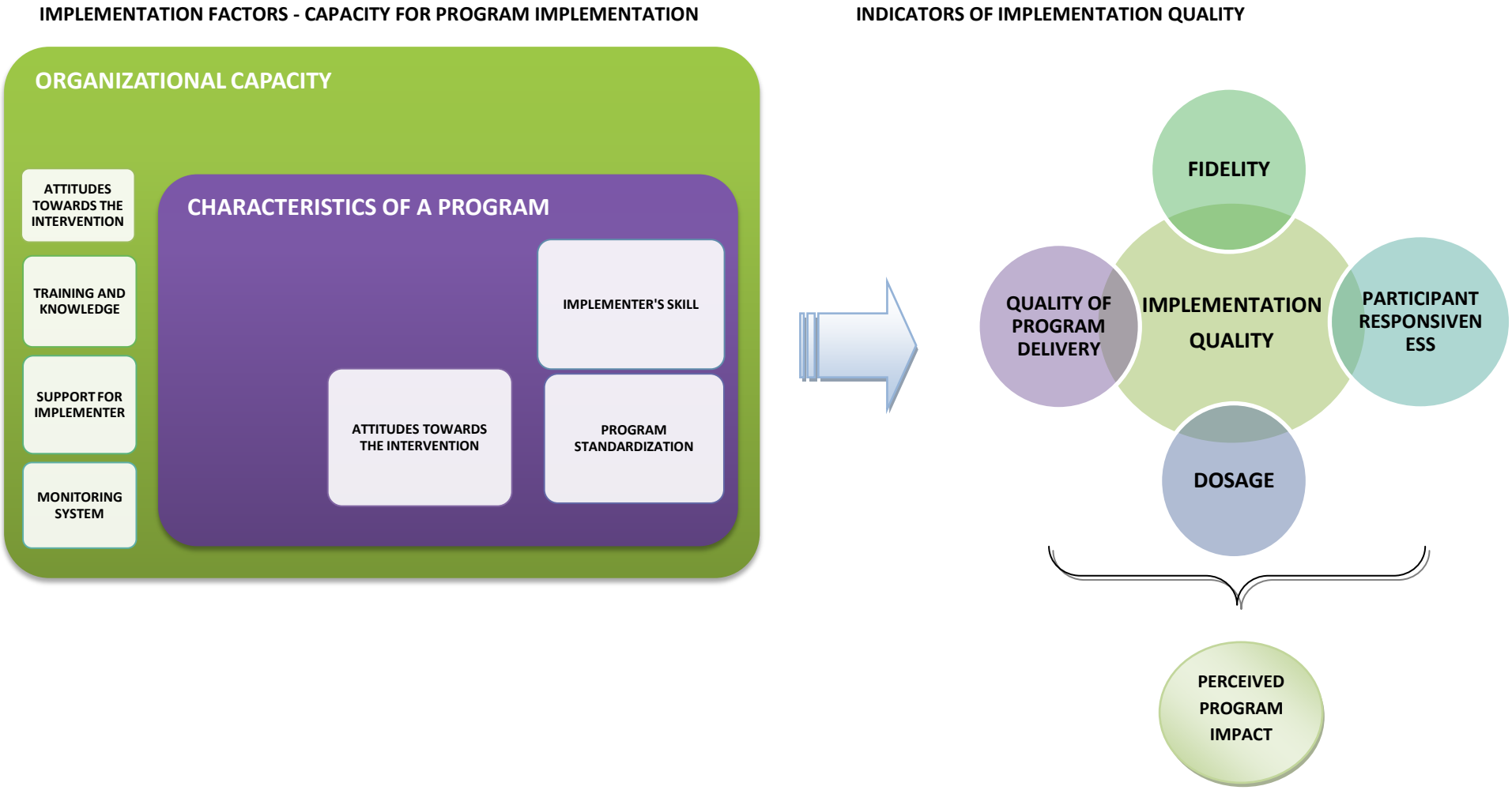
Implementation quality, the primary outcome in the conceptual model, is represented across 5 indicators on the right hand side of the figure. Four dimensions of implementation quality as defined by Durlak and Dupre (2008) were examined and assessed from the perspective of both program participants and program implementers. The four dimensions included program fidelity, quality of program delivery, dosage and participants' responsiveness. Measures of participant responsiveness included ratings of 1) the participants' response to the intervention, 2) their response to the program deliverer, and 3) their attendance in the program. In addition to the four dimensions of implementation quality, implementers' and participants' perceptions of program impact were also collected and used as indirect measure of implementation quality. This is based on the assumption that if

program implementers or participants report an impact of the intervention on the participant's behaviour, then it is highly likely that the program was implemented successfully.

Information regarding the implementation factors was gathered from two types of informants: organization managers and program implementers. Information about implementation quality was assessed through self-report measures completed by program implementers and ratings completed by program participants. This multi-measurement strategy was employed because different sources have different perceptions of the constructs being examined and because when conducting analyses examining the associations between implementation factors and outcomes, it is important to use different sources. This reduces the potential that positive associations are a function of shared measurement variance.

The factors at each level are assumed to be connected and interdependent, affecting the level of implementation quality. This doctoral research will study to what degree each of the implementation factors is related to indicators of implementation quality.

Figure 8. Conceptual model of implementation factors and implementation outcomes used in this doctoral research



2. CHAPTER TWO: AIMS AND RESEARCH QUESTIONS

As discussed in the introduction, this doctoral research examines the issue of implementing preventive interventions in community settings taking into account the characteristics of the programs and the support system surrounding them. Recent research and emerging theories of implementation quality stress the importance of researching implementation quality measured through the multiple implementation dimensions (Durlak and Dupre, 2008) as well as factors affecting implementation quality (i.e. implementation drivers, Fixsen et al., 2005, 2009). Previous studies from the implementation field mostly concentrate on the implementation of evidence-based model programs which, as their title indicates, are already proven to be effective. This restricted focus does not fit in the field of mental health promotion and prevention in Croatia since most of the programs are locally developed and have not been tested in rigorous evaluation studies.

Durlak and Dupre (2008) have defined eight different implementation aspects: fidelity, dosage, quality, participant's responsiveness, program differentiation, monitoring of control/comparison conditions, program reach and adaptation. Most of those eight implementation aspects defined by Durlak and Dupre (2008) are concerned with characteristics of effective programs which are being disseminated in different settings. There is almost no research on the implementation of local community driven programs and about the aspects of implementation defined by Durlak and Dupre (2008) which could be studied within real-life context. From this study perspective, such local initiatives should also be the focus of implementation research because they represent the reality of many prevention and health promotion practices, especially in countries where the mental health promotion and prevention field is still developing. Even if a program is not yet proven effective, research on implementation quality can inform program developers and deliverers about the aspects of program quality and possibilities of improvement. This approach to implementation quality is new in mental health promotion and prevention but reflects the real life conditions.

This doctoral research will examine implementation factors that represent capacity for program implementation and four out of eight indicators of implementation quality defined by Durlak and Dupre (2008) i.e. fidelity, dosage, quality of delivery and participant

responsiveness with addition with perceived program impact which will serve as a proximal indicator of possible program outcomes.

Previous studies have rarely included multiple informants when researching implementation quality. Earlier research has mainly reflected the perspective of implementers, rarely including the information about implementation quality from program managers or participants. This study is contributing to the field of implementation research by including the program managers', implementers' and program participants' perspectives on implementation factors and indicators of implementation quality. Results of previous research stress the importance of assessing both system support factors, characteristics of the program itself and different informants, recommending comprehensive well-thought-out theories of implementation development including macro-level and individual factors (Domitrovich et al., 2008). In addition, prior reviews and meta-analytic studies on implementation research have rarely included the analysis of relationship between implementation factors and indicators of implementation quality. Also, lots of prior studies recommend the investment in the training of managers and implementers in order to affect system and organizations which are carrying the programs. This study is also unique because of the Training for Prevention which was developed to check its impact on the implementation quality through the improvement of behaviours of managers and implementers.

This doctoral research will try to encompass all mentioned recommendations for future implementation research in the context of locally developed programs implemented in real life settings. The rigour of the study design and its power have been affected by the fact that this research is including a cohort of studies which are not yet proven effective and that there was no possibility to include a control group for each of the programs. With mentioned limitations in mind, this study wants to assess the general level of implementation quality of the program cohort in focus, with intention to stimulate the implementation research movement in mental health promotion and prevention field, especially in the European context. Additionally, this study wants to test the potential of the Training for Prevention intervention, i.e. a training about the science-based principles of mental health promotion and prevention science to improve both contextual factors that affect the implementation process and implementation quality.

The general aim of this doctoral research is to study implementation processes and their outcomes in prevention programs in Croatia. In order to achieve the above stated aim, this doctoral research was conducted through pre-research concerning the construction of implementation scales and two studies, 1) study on implementation quality and 2) study of the impact of Training for Prevention. More specifically, here are aims and research questions for both of the studies.

I. STUDY ON IMPLEMENTATION QUALITY

(1) **To identify the relationship between implementation factors and indicators of implementation quality in order to describe the implementation process in the cohort of 24 mental health promotion and prevention programs from the County of Istria.**

Study on implementation quality was conducted with four newly designed measures of implementation quality: one for program managers, two for program implementers and one for program participants. Likewise, the study on implementation quality investigates the relationship of implementation factors and indicators of implementation quality.

This goal was examined through several research questions:

1.1. What is the level and variability of implementation quality of preventive programs in the County of Istria?

1.2. Is there a difference in the self-report assessment of implementation factors and indicators of implementation quality collected from program managers, program implementers and program participants?

1.3. What is the relationship between implementation factors (i.e. program standardization, implementer's skill, attitudes towards the intervention, training and knowledge, support for implementer and monitoring system) and indicators of implementation quality (i.e. fidelity, quality of program delivery, participants' responsiveness, dosage and perceived program impact)?

II. STUDY OF THE IMPACT OF TRAINING FOR PREVENTION

(2) To assess the effects of the Training for Prevention on the implementation factors and implementation quality reported by program managers, implementers and program participants.

The Training for Prevention was designed to inform program managers and implementers about effect predictors and methods of implementation quality enhancement. Study of the impact of Training for Prevention will examine if the newly designed Training for Prevention has affected implementation factors and implementation quality in the experimental group.

This goal was examined through several research questions:

2.1. Does the Training for Prevention result in the increase of implementation factors for the experimental group of projects at post-test?

2.2. Does the Training for Prevention result in the increase of indicators of implementation quality for the experimental group of projects at post-test?

In accordance to this, it is possible to define following research tasks:

1. **To construct valid and reliable measures of implementation quality based on implementation literature and existing measures.**

2. **To explore the level and variation of implementation quality in preventive programs in Istria.**

3. **To explore the differences in perception of implementation quality between program managers, program implementers and program participants.**

4. **To explore the relationships of implementation factors and indicators of implementation quality.**

5. **To test the impact of the Training for Prevention on the level of implementation quality in experimental group by comparing the implementation quality of programs in the experimental and control conditions.**

Related to the fifth research task the following hypotheses will be tested:

- *Hypothesis 5.1:* After the Training for Prevention, program managers and implementers from the experimental group will report improved implementation factors in comparison with the control group.
- *Hypothesis 5.2:* After the Training for Prevention, program implementers and program participants in the experimental group will report improved indicators of implementation quality in comparison with the control group.

3. CHAPTER THREE: METHODS

The research within this doctoral study was conducted in three steps in order to examine and test research tasks which are proposed in chapter two. Mentioned steps were a part of a wider project explained in the introduction section, project "Communities That Care: Development of a Model for Behavioural Disorders Prevention", within which "Preffi: Assuring the Quality of Prevention Programs in the County of Istria" was conducted by researchers from the Faculty of Education and Rehabilitation Sciences, University of Zagreb, in cooperation with the Department of Health and Social Services, County of Istria. This dissertation focuses on three studies:

1) **Construction of scales for the implementation qualityassessment –** the aim of this first study was to select items for constructing the measures of implementation quality for three types of informants: organization managers, program implementers and program participants.
2) **The study on implementation quality**, which was conducted in order to construct the final version of the implementation quality measures, to test their reliability and to explore the level of implementation quality in preventive programs in Istria.
3) **The study of the impact of the Training for Prevention**, the third study was conducted during and after the "Training for Prevention" (intervention) and was designed with the purpose of testing the impact of the Training on the level of implementation quality.

This chapter is organized in the way that each of the three studies is described individually, connecting participants and the procedure introduction. Measures used are described at the end of the chapter, corroborated by metric characteristics, and followed by ethical considerations. To ensure more clarity in this methods section, it is important to stress that the overall sample is the same in all presented studies and consists of 24 programs and their managers, implementers and participants. Numbers of actual participants differ from one study to another, mainly because studies are connected with different time points and for the reason that some programs did not have two data collection. Additionally, there were a few subjects on each level (managers, implementers or participants) who did not return the questionnaires at post-test data collection. From those 24 programs, 24 managers participated in the data collection. In the mid-intervention assessment 51 implementers from 24 programs participated in the data collection while there were 55 of them in the post-test measurement.

434 intervention participants gave their self-report in the mid-intervention assessment while 744 intervention participants were covered in the post-test data collection. All samples will be described for each of the studies below in more detail.

Strategy for data analyses

Before detailed methods overview which will be shown for each of the three studies, it is necessary to mention used strategies for data analyses. All statistical procedures will be explained in detail within each chapter but this subsection aims to enhance the transparency and clarity of the whole dissertation. First of all, metric characteristics of the Preffi 2.0 instrument were conducted within the doctoral research of colleague Josipa Mihić and were taken over for methods overview. For the calculation of G factors, which stands for Cronbach's alpha, the theory of generalizability was applied, so the presented methods chapter offers the calculated G factor and the corresponding standard error of measurement (Shavelson & Webb, 1991). Details can be found in the other dissertation which was part of the same project. In order to answer first research task concerned with scale construction, reliability analyses were conducted: Cronbach's alpha was calculated, Inter-Item correlation matrix was produced and according to Item-Total statistics, items were deleted if they were not consistent with other items, did not have high part-whole correlation and if Cronbach's alpha increased when the item was deleted. Additionally, construct validity was checked with the analysis of the dimensionality of each construct within the scale, whether it was an implementation factor or an indicator of implementation. Since we wanted to include items that assess a single theoretical construct, per each of the constructs a principal component factor analysis was carried out to determine how strongly each item loads onto a single component, i.e. the first common factor. Also, test-retest correlation was calculated for all scales to check test-retest reliability. The second research task which was dealing with level and variability in implementation quality of preventive programs in Istria was met with simple descriptive analyses. The third and fourth research tasks which are connected with the differences in perception as well as with relationships of implementation factors and indicators of implementation quality were followed with the correlation analyses as well as with multiple regression procedures and hierarchical linear modelling since data was nested. The fifth research task which aimed to test the impact of the Training for Prevention on the level of implementation factors and indicators of implementation quality reported by managers, implementers and participants also employed multiple regression procedure and hierarchical linear modelling. Additional analyses of moderator effects on the impact of the

Training were also tested with HLM (Baron and Kenny, 1986). All simpler analyses were conducted with SPSS software while multiple regression and HLM were done with SAS software.

3.1. Construction of scales for implementation quality assessment

Elements and steps in scale development were: focus groups with program managers and implementers, analysis of implementation literature and available measures for implementation assessment, selection of an initial set of items which would operationalize the concept of implementation, revision of the initial set of items with support of renowned prevention researchers, and a quantitative survey research with the purpose to select the final set of items.

A focus group with program managers and implementers was held in June 2010 in Zagreb. Seven participants of the focus groups were gathered from the most eminent non-governmental organizations in Zagreb conducting mental health promotion and prevention programs for children and youth. Key questions for participants in focus group were: 1) "When you develop your own preventive interventions, which aspects are in your view the most important in order to assure the effectiveness of your program?", 2) "According to your experience in program dissemination, what are the core components of your interventions?", 3) "How do you assure program dissemination and sustainability?", 4) "What is the level of quality of prevention programs in Croatia and what steps should be done to improve them?" and 5) "What additional knowledge and skills on the principles of effectiveness would you like to receive?". Discussion during focus group has revealed that mental health promotion and prevention programs in Croatia are rarely evidence-based, their dissemination is not standardized, and they rarely have a specified set of themes and well developed materials and handbooks. Further, training of program implementers is not a crucial concern of organizations while organizations only occasionally assess the quality of program delivery, quality of program implementers' work and intervention effects on participants. Data collected from the focus group was used as an additional input to the literature overview, especially to offer insight about the current situation with preventive programs' implementation quality in Croatia.

An initial set of items for implementation measures was designed during the period from October 2010 to January 2011 based on the implementation literature review and

implementation measures used in implementation research presented in the literature. While constructing the scales for implementation quality assessment, items were generated according to theoretical definitions of constructs, both implementation factors which reflect the capacity of structure for implementation and indicators of implementation quality. Items were generated according to the conceptual model of implementation factors and indicators of implementation quality used in this doctoral research and presented in the introduction (see Figure 8, page 45). In the tables presented in Appendix 1 (see page 212) and Appendix 2 (see page 216) it is evident how theoretical concepts were described with the help of items. Six factors which reflect the capacity of structure for implementation were assessed through the self-report of organization managers and program implementers while five aspects of implementation quality collected perspectives of program implementers and program participants (see Table 3.1).

Table 3.1
Assessment of implementation factors and indicators of implementation quality through different types of informants.

IMPLEMENTATION FACTORS REFLECTING THE CAPACITY FOR PROGRAM IMPLEMENTATION	
CHARACTERISTICS OF A PROGRAM	
PROGRAM STANDARDIZATION	
Organization managers	Program implementers
PROGRAM IMPLEMENTER'S SKILL	
Organization managers	
ATTITUDES TOWARDS THE INTERVENTION	
	Program implementers
ORGANIZATIONAL CAPACITY	
ATTITUDES TOWARDS THE INTERVENTION	
Organization managers	
TRAINING AND INTERVENTION KNOWLEDGE	
Organization managers	Program implementers
SUPPORT FOR IMPLEMENTER	
Organization managers	Program implementers
MONITORING SYSTEM	
Organization managers	Program implementers
INDICATORS OF THE IMPLEMENTATION QUALITY	
FIDELITY	
Program implementers	
QUALITY OF PROGRAM DELIVERY	
Program implementers	Program participants
PARTICIPANTS' RESPONSIVENESS	
Program implementers	Program participants
DOSAGE	
	Program participants
PERCEIVED PROGRAM IMPACT	
Program implementers	Program participants

Items for organization managers and program implementers had to overlap partly because they were reporting on the same implementation concepts of their programs: program standardization, training and intervention knowledge, support for program implementers, monitoring of program deliverers' work and attitudes toward the intervention. Also, items in the scale for program implementers and program participants overlap because they were both reporting on participants' responsiveness i.e. participants' engagement and participants' acceptance of the intervention, quality of program delivery and perceived program impact. It is evident from the table A1 and A2 in Appendix 1 (see page 212) and 2 (page 216) that some items, although parallel, when answered by different informants actually represent a different concept: for example, some items from the perspective of program implementer represent his attitudes towards the intervention (implementation factor) while the same items from the perspective of program participants reflect their responsiveness to the program (indicator of implementation quality). That is the same for the perspective of program manager who is reporting on his beliefs in the skill of program implementer (implementation factor) while the same items from the perspective of program implementer reflects the quality of program delivery.

The initial pull of claims resulted in 85 items. Of the overlapping items, 40 were included both in the scale for organization managers and program implementers, and 30 items were both in the scale for program implementers and the scale for program participants. Another 15 items were added to the set of items for participants because they reported on their views about the quality of program delivery and the implementer or on the perceived impact of the intervention on their life. Regarding the fact that participants of mental health promotion and prevention programs are also children who could have difficulties in understanding the questions, a child's version of the scale of implementation quality for participants was also designed.

3.2. Study on implementation quality

3.2.1. Participants

Sample of the study of implementation quality included managers, implementers and participants from the cohort of 24 community-based mental health promotion and prevention programs in Istria, that were being provided by local organizations. The programs were drawn from the 2011 applicants for financial support from the County of Istria's Department of Health and Social Care initiative entitled "Prevention of Behavioural Disorders and

Prevention of Substance Abuse." In order to describe the level of quality of mental health promotion and prevention programs in one local area, it was very important to study a representative sample of programs. All 24 programs included in this research embody key mental health promotion and prevention activities run by various organizations in the County of Istria. Their organizations, organization managers, experts and practitioners are the region's most active stakeholders in the field of mental health promotion and prevention. Most of the mental health promotion and prevention activities in that Region are initiated by NGOs and almost all of these organizations are applying for financial support of the County of Istria's Department of Health and Social Care. Also, most of them have been financed by the Department regularly in the last 5 years. Their tradition guarantees that those programs reflect the current state of the art of mental health promotion and prevention programs in the County of Istria.

Table 3.2 describes the profile of each of the studied projects, including information of the provider and name of the project, program content and methods, type of managers, implementers and program users, number of study participants, and the prevention level of the intervention. As shown, the final sample of 24 programs includes parenting interventions, socio-emotional programs for children and youth, programs for alcohol and drug abuse prevention, mentor programs, programs for violence prevention among peers and mental health and resilience promotion programs. It is important to stress that included programs differ in their length, number of workshops with participants, number of participants, number of people delivering the program and the techniques used for program delivery. As seen in Table 3.2, data for each of the 24 included programs were collected from 24 managers and 55 program deliverers. Participant numbers for each program are shown in Table A4 in Appendix 4 (see page 225), altogether 434 program participants gave their report about the implementation quality at mid-intervention while 744 participants gave their report at post-test implementation assessment.

Table 3.2

Details on organizations and projects of mental health promotion and prevention from the County of Istria included in the research

	NAME OF THE PROJECT	PROGRAM CONTENT	PARTICIPANTS	MANAGERS AND DELIVERERS	LEVEL OF PREVENTION
	PROGRAMS PROMOTING CHILD AND ADOLESCENT WELLBEING, POSITIVE DEVELOPMENT AND SOCIO-EMOTIONAL COMPETENCIES				
1.	ZIID TEATAR Teen theatre workshop	Assertiveness training using theatre techniques 1 cycle, 49 meetings	School children, age 11-14 **12 participants**	1 manager 1 deliverer	Universal
2.	ART STUDIO Junior plus – parent-child art classes	Structured free time health promotion program using creative techniques 1 cycle, 10 workshops	School children, age 7-14 **10 participants**	1 manager 1 deliverer	Universal
3.	LABIN ZDRAVI GRAD My picture in the mirror	Health promotion program for self-confidence training 4 cycles, 4 workshops	School children age 10-11 **30 participants**	1 manager 4 deliverers	Universal
4.	FOND ZDRAVI GRAD POREČ Helping peers – volunteers of healthy city: social skills peer program	Positive development promotion program 1 cycle, 30 meetings	School children, age 11-12 **27 participants**	1 manager 1 deliverer	Universal

	NAME OF THE PROJECT	PROGRAM CONTENT	PARTICIPANTS	MANAGERS AND DELIVERERS	LEVEL OF PREVENTION
5.	FOND ZDRAVI GRAD POREČ Empowering children through dance – Dancing classrooms	Program of health promotion aimed at life skill training 1 cycle, 12 workshops	School children, age 12-13 **113 participants**	1 manager 2 deliverers	Universal
6.	DRUŠTVO PSIHOLOGA Media literacy program	Program for prevention of cyber-bullying and promotion of responsible behaviour on the Internet 5 cycles, 4 workshops	School children, age 9-11 **139 participants**	1 manager 3 deliverers	Universal
7.	OBITELJSKI CENTAR Neighbourhood circle	Structured free time health promotion program using art techniques 1 cycle, 12 workshops	School children, age 7 - 15 **19 participants**	1 manager 3 deliverers	Universal
8.	OBITELJSKI CENTAR Mentor program „Give me 5“	Mentor program promoting positive adult and child relationships *pairs seeing each other during 1 year	School children, age 7 - 15 **3 participants**	1 manager 1 deliverer	Universal/selective
9.	DND PAZIN Growing up together	Social skills peer program 2 cycles, 4 workshops	School children, age **42 participants**	1 manager 2 deliverers	Universal/selective

	NAME OF THE PROJECT	PROGRAM CONTENT	PARTICIPANTS	MANAGERS AND DELIVERERS	LEVEL OF PREVENTION
10.	OAZA: TEEN CLUB Promoting healthy lifestyles in children and youth from foster families	Structured free time health promotion program for children in foster care *meetings during one year	Children and adolescents in foster-care, age **8 participants**	1 manager 4 deliverers	Selective
11.	SUNCOKRET Easier through school	Program for academic support for children with learning difficulties 1 cycle, 5 workshops	University students age 20-23 **10 participants**	1 manager 3 deliverers	Selective/indicated
	PARENTING PROGRAMS				
12.	ODISEJA Successful parenting	Parent training program for elementary school children, age 11-13 5 cycles, each having 7 workshops	Parents **50 participants**	1 manager 10 deliverers	Universal
13.	DND PULA Quality parenting course	Parent training program, mixed age of children 2 cycles, each having 10 workshops	Parents **23 participants**	1 manager 1 deliverer	Universal

	NAME OF THE PROJECT	PROGRAM CONTENT	PARTICIPANTS	MANAGERS AND DELIVERERS	LEVEL OF PREVENTION
14.	OBITELJSKI CENTAR Parenting with a smile	Parent training program for parents of pre-school children 1 cycle, 6 workshops	Parents **8 participants**	1 manager 1 deliverer	Universal
15.	GRAD BUZET Happy parent – happy child	Parent training program, mixed age of children 1 cycle, 8 workshops	Parents **12 participants**	1 manager 1 deliverer	Universal
16.	LABIN ZDRAVI GRAD Supporting parenting	Parent training program for parents of pre-school children 1 cycle, 3 workshops	Parents **5 participants**	1 manager 2 deliverers	Universal
17.	ANTE BABIĆ UMAG Quality parenting training	Parent training program, mixed age of children 1 cycle, 10 workshops	Parents **10 participants**	1 manager 1 deliverers	Universal
18.	GRAD PAZIN Quality parenting course	Parent training program, mixed age of children 2 levels: for motivated parents and higher risk group of parents, 8 workshops	Parents **8 participants** **7 participants**	1 manager 1 deliverer in universal version 2 deliverers in selective version	Universal/selective

	NAME OF THE PROJECT	PROGRAM CONTENT	PARTICIPANTS	MANAGERS AND DELIVERERS	LEVEL OF PREVENTION
19.	ASANDO CHER Let's grow up together	Parent training for the Roma parents, of pre-school children 1 cycle, 7 workshops	Parents **13 participants**	1 manager 3 deliverers	Selective
	SUBSTANCE ABUSE PREVENTION				
20.	ZUM Supporting community in substance use prevention	Teen substance abuse prevention program 1 cycle, 6 workshops	Adolescents, age 15-17 **8 participants**	1 manager 2 deliverers	Universal
21.	INSTITUT I know who I am	Substance abuse education 1 lecture	Adolescents, age 15-16 **50 participants**	1 manager 2 deliverers	Universal
22.	Program of substance abuse prevention for teachers	Substance abuse education, teacher based 1 lecture	High-school teachers **63 participants**	1 manager 1 deliverer	Universal/Selective
23.	Program of substance abuse prevention for parents	Substance abuse education, parent-based 1 lecture	Parents of high-school children **29 participants**	1 manager 1 deliverer	Universal/Selective
24.	GRAD PAZIN Early drinking of youth and its prevention	Prevention of alcohol consumption in youth 1 cycle, 6 workshops	Adolescents, age 12-15 **45 participants**	1 manager 1 deliverer	Selective

3.2.2. Procedure

In November 2010, researchers from the Faculty of Education and Rehabilitation Sciences, University of Zagreb have in collaboration with the stakeholders from the Department of Health and Social Care, County of Istria, started the procedure of selecting a representative sample of mental health promotion and prevention programs. Written program proposals were chosen from the cohort of 2011 applicants for financial support from the Department initiative entitled "Prevention of Behavioural Disorders and Prevention of Substance Abuse" and from the pool of local organizations conducting interventions in the field of mental health promotion and prevention in Istria. The Department's application form is standardized and contains 13 sections that are completed by applicants. It includes questions about the organizations previous experience, a description of the outcomes targeted by the intervention, the community need assessment, goals and targeted results of the project, description of participants and activities, evaluation of efficiency, planned staff, partners and volunteers as well as the planned budget. Researchers from the Faculty of Education and Rehabilitation Sciences supplemented that form with a structured questionnaire about organizational issues and internal communication.

Firstly, three independent assessors from the Department of Health and Social Care assessed all received applications and decided which organizations and programs will get a financial support from the County of Istria. After the Department's assessors selected the total of 30 programs to be financed by the County of Istria, researchers from the Faculty have from that group selected 24 programs focused on mental health promotion and prevention of mental, emotional and behavioural disorders. Six of the programs that got financial support from the Department were excluded from this doctoral dissertation as they were focusing only on treatment. The final sample of programs included in this study consisted of 24 programs described in Table 3.2.

3.3. Study on the impact of the Training for Prevention

3.3.1. Participants and matching

In order to examine the impact of the Training for Prevention on the implementation quality of mental health promotion and prevention programs, two groups of participants were involved in this study. The total sample of this third study included managers, implementers

and participants of 24 community-based mental health promotion and prevention programs in Istria, which were described and presented in Table 3.2 above.

The 24 programs were divided in an experimental and control group, each containing 12 programs, which was done by the equal pairs matching method. Experimental and control group were matched based on the following criteria: type of the mental health promotion and prevention program, number of program users, number of years during which the program was financed as well as the duration of program tenure in Istria. While most of the organization managers and program deliverers from the sample are acquainted with each other and sometimes even collaborate because the County of Istria is rather small, researchers have tried to prevent that experimental and control group are locally overlapping to reduce the possibility of contamination, i.e. the risk of communication between experimental and control groups about the content of the "Training for Prevention". For example, when one organization or local community had several programs included in this study, all of these programs had to be in the same conditions, experimental or control. Researchers intended to have both groups as similar as possible, both containing similar types of programs and similar levels of program quality.

That is the reason why all 24 included programs were assessed with the Preffi 2.0 instrument which is designed to assess the extent to which mental health promotion and prevention program is likely to be effective. Both experimental and control group of programs were additionally equalized according to the average group results on the Preffi 2.0 total scores. Division of the programs in experimental or control conditions is shown in Table 3.3. The Training for Prevention intervention included managers and their program implementers. Implementation measures were administered to managers, implementers and program participants of all 24 programs included in the studied cohort.

Table 3.3

Division of 24 mental health promotion and prevention programs in control and experimental conditions and their results on the Preffi 2.0 assessment

	Experimental conditions	Control conditions
1.	ZIID TEATAR: Teen theatre workshop Assertiveness training using theatre techniques **Preffi 2.0 score: 5.33**	FOND ZDRAVI GRAD POREČ: Dancing classrooms Program of health promotion aimed at life skill training **Preffi 2.0 score: 6.90**
2.	OBITELJSKI CENTAR: "Give me five" Mentor program promoting positive adult and child relationships **Preffi 2.0 score: 6.22**	OAZA: Promoting healthy lifestyles in youth from foster families Structured free time health promotion program for children in foster care **Preffi 2.0 score: 6.29**
3.	LABIN ZDRAVI GRAD: Supporting parenting Parent training program for parents of pre-school children **Preffi 2.0 score: 4.17**	ANTE BABIĆ UMAG: Quality parenting training Parent training program for mixed age of children **Preffi 2.0 score: 4.19**
4.	DRUŠTVO PSIHOLOGA: Media literacy program Program for prevention of cyber-bullying and promotion of responsible behaviour on the Internet **Preffi 2.0 score: 7.10**	GRAD PAZIN: Early drinking of youth and its prevention Prevention of alcohol consumption in youth **Preffi 2.0 score: 5.84**
5.	SUNCOKRET: Easier through school Program for academic support for children with learning difficulties **Preffi 2.0 score: 7.40**	FOND ZDRAVI GRAD POREČ: Helping peers – volunteers of healthy city; Social skills peer program **Preffi 2.0 score: 8.56**
6.	ZAVOD ZA JAVNO ZDRAVSTVO Substance abuse education, parent-based **Preffi 2.0 score: 5.01**	INSTITUT: I know who I am Substance abuse education **Preffi 2.0 score: 5.09**
	Experimental conditions	**Control conditions**
7.	ZAVOD ZA JAVNO ZDRAVSTVO Substance abuse education, teacher-based **Preffi 2.0 score: 5.18**	ZUM: supporting community in substance-use prevention Teen substance abuse prevention program **Preffi 2.0 score: 5.41**

8.	DND PULA Parent training program **Preffi 2.0 score: 5.78**	GRAD PAZIN: Quality parenting course Parent training program for mixed age of children, universal and risk version **Preffi 2.0 score: 5.55**
9.	LABIN ZDRAVI GRAD: My picture in the mirror Health promotion program for self-confidence training **Preffi 2.0 score: 4.53**	DND Pazin: Growing up together Social skills peer program **Preffi 2.0 score: 4.69**
10.	OBITELJSKI CENTAR: Neighbourhood circle Structured free time health promotion program using art techniques **Preffi 2.0 score: 6.62**	ART STUDIO: Parent-child art classes Structured free time health promotion program using creative techniques **Preffi 2.0 score: 4.74**
11.	GRAD BUZET: Happy parent – happy child Parent training program for mixed age of children **Preffi 2.0 score: 5.50**	ODISEJA: Successful parenting Parent training program for parents of elementary school children **Preffi 2.0 score: 6.16**
12.	OBITELJSKI CENTAR: Parenting with a smile Parent training program for pre-school children **Preffi 2.0 score: 6.00**	ASANDO CHER: Let's grow up together Parent training program for parents of pre-school children **Preffi 2.0 score: 4.21**
	TOTAL PREFFI SCORE EXPERIMENTAL CONDITIONS 5.74	**TOTAL PREFFI SCORE CONTROL CONDITIONS 5.64**

3.3.2. Procedure

Firstly, during December 2010 and January 2011 all 24 included mental health promotion and prevention programs were scored with the Preffi 2.0 instrument (see Appendix 6, page 233) to get a quantitative appraisal of their program quality. This appraisal was also used as a reference point for the matching of programs in control and experimental conditions. The scoring of each program within the sample (N=24) was based on the evaluation of the written proposal that was submitted with the application for funding to the Department of Health and Social Care, County of Istria. As it was previously stated, the Department's application form is standardized and contains 13 sections that have to be completed by applicants. It includes questions about the organizations' previous experience, a description of the outcomes targeted by the intervention, the community need assessment, goals and targeted results of the project, description of participants and activities, evaluation of efficiency, planned staff, partners and volunteers as well as the planned budget. The form was supplemented with a structured questionnaire about organizational issues and internal communication provided by the researchers.

All 24 written project proposals were firstly read and assessed with the Preffi 2.0 questionnaire by 3 independent assessors (i.e. two doctoral candidates conducting the research within the project and the project leader, a senior prevention expert) based on the recommendations of Preffi authors (Molleman and colleagues, 2005). After the independent appraisals, the 3 assessors discussed the results and agreed on the general ratings on each of 8 Preffi clusters. After the total scores for each project are calculated, internal consistency scores have to be computed. The reliability of Preffi 2.0 has to be assessed by using the generalizability theory and calculating the generalizability coefficient (G) and the standard error of measurement (SEM) (Shavelson and Webb, 1991). Cronbach's alpha could not be used as a reliability estimate as both graders and items may contribute to the measurement error. While Cronbach's alpha is only applicable in situations where there is only one source of measurement error, the generalizability theory accommodates complex measurement designs with more sources of error. For this study, G and SEM were computed on different levels of aggregation: for each of the eight clusters and for Preffi total score. The conventional minimum reliability threshold for reliability coefficients such as G is 0.70, similar to the minimum reliability threshold of Cronbach's alpha. There is no generally accepted maximum value for SEM, but the convention is that the accepted value of SEM is lower than 0.26. Molleman and his colleagues (2006) found following Preffi 2.0 reliability indicators: for all

clusters together: G=0.85, SEM=0.49 and for total project score: G=0.67, SEM= 0.86. Both G factors, for the total project score and all clusters together, indicate sufficient reliability and are higher than the conventional minimum threshold.

After the calculation of the total Preffi scores which could range from 3.33 to 10, programs were divided across the control and experimental condition. Pairs of programs were created based on the following criteria: whether participants were children, teenagers or adults, the type of the program, the locality from which the program was coming, the duration of the intervention, the number of participants and the total Preffi 2.0 results. The research team has organized meetings with organization managers and program implementers from both control and experimental group of programs, explaining the aim of the study, study design, measurement dates and following steps. Participants from control conditions were told that they would receive the Training for Prevention intervention after the whole study and measurement are finished. To respect that, the Training for Prevention intervention was delivered to control group participants during April and May 2012.

Participants from the experimental conditions were asked to respect the discretion rules and secrecy about the content of the Training for prevention intervention. Organization managers and program deliverers had to sign a secrecy agreement which was attached to the financing contract from the Department of Health and Social Care. Also, they were asked to commit that at least one member of organization and program implementer will be present in all blocks of the Training for Prevention intervention. For the experimental group, a 32-hour education of Training for Prevention was organized mostly during March 2011, exact dates being 25th February 2011, 3rd March 2011, 11th March 2011 and both 17th and 18th March 2011. It was decided that the timing of the Training for Prevention intervention will be set in the first trimester of 2011 because programs differed in their starting date and length. Besides the direct involvement in the intervention, all organizations from the experimental conditions received individual consultation and feedback from researchers about the strengths and weaknesses of their programs. Individual consultation lasted 3 hours in total.

To measure the changes in implementation quality in both experimental and control cohorts of programs, scales for implementation quality were sent to the organization managers, program implementers and program participants at approximately 1/3 i.e. close to the half of each program delivery and at the very end of the program delivery. Dates of data collection are presented in Table A3 in Appendix 3 (see page 220) while number of participants for which data is collected at the two time points is presented in Table A4 in Appendix 4 (see page 225). That was done both to see the changes in implementation quality

throughout the program but also to assess the influence of the Training for Prevention on implementation quality. Ideally, it was intended that implementation quality measures will be assessed for organization managers, program deliverers and program participants at two time points for all 12 programs from the control and all 12 programs from the experimental condition. Unfortunately, both timeline assessments were not possible for participants as referrers in six programs, 3 being in the experimental and 3 being in the control condition. In the experimental condition these include: 1) "Media literacy program" because of the lack of time for two data collections; 2) "Program of substance abuse prevention for teachers" which had only one lecture and 3) "Program of substance abuse prevention for parents" which had only one lecture. In the control condition they concerned: 1) "Underage drinking prevention" because of the lack of time for two data collections, 2) "Substance abuse prevention" program that also had only one lecture, and 3) "Parenting programme V" whose participants did not return the questionnaires after the program ending. These programs will have to be excluded from the analysis because the calculation of composite result of implementation quality at both time points is not possible without the data from all referrers: managers, program deliverers and program participants.

3.4. Content of the Training for Prevention

Training for Prevention consists of 32 hours of direct interactive group education and 3 hours of additional individual consultations. Group education was based on lectures, group activities, exercises, case studies and was accompanied by continuous feedback from training deliverers. Activities within group education followed the precise structure of six main topics:

1. science-based prevention practice (4 hours),
2. logic modelling and quality (8 hours),
3. implementation (8 hours),
4. evaluation (8 hours)
5. advocacy (4 hours).

All themes were delivered within the period of one month leaving time between five group sessions during which participants could integrate the knowledge, work on assigned tasks and practice skills. Parallel with the group work and activities, training deliverers had three hours of individual consultations with program leaders, authors and program deliverers during which learning process was discussed. Individual consultation included reflection on the tasks fulfilled during group training with the feedback on the level of achieved quality. Special attention was given to the transfer of gained knowledge during the training into specifics of their program demands. Each topic covered several sub-areas which are elaborated in text which follows.

1. **Science-based prevention practice**

   Regarding the differences between participants professional background, their experience and level of education about prevention, at the beginning of the Training for Prevention participants were introduced to recent concepts of prevention science and practice. The topic Science-based prevention practice included lectures on theoretical models of prevention, prevention continuum, risk and protective factors and prevention cost-effectiveness.

2. **Logic modelling and quality**

   The topic of Logic modelling and quality focused on the transfer of knowledge about all the phases and processes needed for development of comprehensive and precise program's logic model. During this phase all participants continuously worked on the

development of the logic model of their own program. Firstly, the target group was informed about the importance of elaborate problem analysis in project development and taught how to define those problems which they want to effect with their program's activities. That was followed by the theme of need assessment during which the connection between problem analysis and need assessment was emphasised. Participants were educated about the methods of need assessment, how to use available research, resources and data and conduct need assessment for their programs. In this phase, the difference between detected problems and existing needs was emphasised and explained on several examples. According to the results of need assessment, participants were taught how to precisely describe the target group which they want to include in their intervention. After gaining that knowledge, participants were instructed how to define specific and quality project goals, based on conducted problem analysis, need assessment and available resources analysis. Description of short and long term outcomes follows defined goals and represents the projection of goals achievement.

All of the described processes in this phase of Training for Prevention focused on better understanding of the theory behind the program. Participants were educated about the principles of internal theory of change inherent to each program. Participants analysed the causal assumptions behind their programs, were directed to connect their activities with the existing theoretical models and possibilities how to detect and overcome potential barriers in the process of project development and delivery. Importance of this part of the Training for Prevention was to raise the awareness of participants about the role of all described elements in the overall program quality.

3. Implementation

During the Implementation topic, participants were educated that implementation quality is crucial for programs effectiveness and quality. This part of the training contained an overview and characteristics of effect predictors related to the implementation process which leads to better target group outcomes. For each type of prevention and promotion program, specific knowledge gathered from up-to-date research was transferred.

This phase started by emphasizing the crucial role of professional capacities of program deliverers which include their professional education, level of training and experience in similar program delivery. Also, a possibility of in-service training

organization was recommended as a method of professional capacities enhancement. This part of the training described moderators of implementation quality which included deliverer's motivation for conducting the program as well as beliefs and expectations of programs effectiveness. Training also paid special attention to development of group management personal skills which increase engagement of participants and their motivation to change. Providing a constructive and continuous feedback from deliverer to the target group has a positive effect on outcomes. Providing organizational support to program deliverer through organizing supervisions and program monitoring, regular organizational meetings, involvement of an organizational manager in program implementation process and assuring administrative conditions is added value to the implementation quality. Program activities have to follow developmental trajectories of the target group in order to address crucial developmental demands relevant for individual change. Participants were educated how to tailor their activities according to the characteristics and needs of target group, what is the optimal number of program participants, what dosage of activities is appropriate and which techniques are the most innovative and efficient for specific programs to achieve expected outcomes. Regarding outcomes enhancement, training participants were directed to encourage their target group to practice skills and generalize content learned during the program to other social environments. Standardization of program content and model of delivery contributes to implementation quality as well. Participants were encouraged to follow fixed schedule of themes and activities in their programs, to develop structured written materials and program manuals.

4. Evaluation

The introduction to the topic of evaluation emphasized the need for continuous and comprehensive assessment of program outcomes and quality of program implementation. Evaluation process was thoroughly described starting with definition, theoretical overview of qualitative and quantitative indicators of program effectiveness, research methods of data collection and data sources. Using a logic model as a starting point, participants were taught which steps they have to follow during the evaluation process. It was explained that concepts which they have to measure are defined within the logic model and program objectives that serve as indicators of desirable change. Based on program objectives, clear evaluation

questions need to be formulated in a way to be measurable. Experimental and quasi-experimental designs were also presented with emphasis on the number of participants needed for plausible conclusions about program effectiveness. Regardless of the design used in evaluation research, measures planned to be used have to be standardized, reliable and in accordance with theoretical concepts behind their programs. Participants were informed about different types of evaluation (process and outcome evaluation, implementation quality research and costs analysis) and developed evaluation plan of their program. For stakeholders who are planning evaluation, the training emphasized advantages of participative evaluation which integrates science-based principles and involves collaboration of program author, researcher and target group. Concerning the level of independence of researcher, evaluation can be internal or external. Training deliverers encouraged participants to plan external evaluation and engage research experts in order to ensure objective conclusions about program effectiveness. Training participants received information that adequate evaluation research gives them insight into successes and failures and an overview of where planning process gaps are located and where improvement is needed. Information gathered during the evaluation process are significant for program development, needed changes in program content and implementation which all leads to program sustainability in community.

5. **Advocacy**

The final phase of the Training for Prevention was focused on the role of advocacy for setting the conditions for success and quality of programs. Participants were educated that through the process of advocacy they can ensure sufficient resources for program development and implementation. Adequate funding, community support, networking and partnership are benefits of quality advocacy ensuring program sustainability. All mentioned influences the visibility of the program and organization, affects the motivation of the target group to participate in the program and gives credibility to program deliverers/authors in the process of policy development. This part of the Training explained characteristics of quality advocacy, steps of the advocacy process, starting with clear definition of advocacy goals. The process of advocacy is a continuous process and its activities start before the program and last during program implementation and after the program is finished. Participants were trained how to detect stakeholders and decision makers they want to address, adjust the message and

their interests to the interest of key people, find common language and use key moments for lobbying. Special attention was given to the methods of advocacy, especially to the usage of media for communicating the message and making more efficient impact on the decision makers. Besides lobbying, participants were directed to recognize available funding resources.

3.5. Measures

3.5.1. The Preffi 2.0 instrument

The Preffi 2.0 instrument (Molleman et al., 2005a, 2005b), an improved version of Preffi 1.0, consists of 39 quality criteria divided in 8 clusters. Each quality criteria and cluster can be scored on a scale from 0 to 3 (0-non assessable, 1-weak, 2-moderate and 3-strong). The instrument is designed to assess the extent to which a health promotion and prevention program is likely to be effective based on the presence and quality of a set of research-based "effect predictors". It is a set of criteria and guidelines for health promotion and prevention specialists that are regarded as essential for the quality and effectiveness of interventions.

- The first cluster includes 14 items and reflects the "Contextual conditions and feasibility" of the intervention being considered. It describes the quality of support and commitment of internal and external partners, capacities for the project, leadership by the project manager including expertise and characteristics of the manager.
- The second cluster includes 13 items and reflects the "Problem analysis". This cluster is representing a quality level of nature, severity and scale of the problem analysis, analysis of distribution of the problem and problem perception by stakeholders.
- The third cluster reflects the "Determinants of behaviour and environment" and consists of 13 items. It refers to the quality level of the program's theoretical model, description of contributions of determinants to the problem, amenability of factors to change and the quality of how determinants are prioritized and selected.
- The fourth cluster includes 7 items and reflects the "Target group" of the intervention. It describes a quality level of general and demographic characteristics of the target group, motivation and opportunities of the target group to change and accessibility of the target group.
- The fifth cluster concerns the "Objectives" and includes 12 items. It is assessing if project's objectives are fitting in with problem analysis, if they are specific, specified in time and measureable, but also if they are acceptable to the main stakeholders and

feasible. It also describes if objectives are considered achievable given the available resources, contextual conditions and intended period of time.

- The sixth cluster, "Intervention development", is the most comprehensive one and consists of 33 items. It reflects the rationale of the intervention strategy, previous experience with intervention, duration, intensity and timing of the intervention, fitting to the target group and to the culture, participation of the target group and usage of effective techniques. It also shows the feasibility in existing practice, characteristics of implementability of the intervention and coherence of the interventions/activities.
- The seventh cluster, "Implementation" has 14 items. It reflects the model of implementation, the fit of implementation interventions to intervention deliverers, appropriateness of the supplier for intermediating intervention deliverers, monitoring and generating feedback, and incorporation of the intervention in an existing organizational structure. The last cluster, "Evaluation" consists of 16 items. This cluster is describing the quality of clarity and agreement on the principles of evaluation between different stakeholders and the quality of process and effect evaluation. Effect evaluation refers to changes which are planned to be measured and if it is plausible that the change was caused by the intervention. The same cluster also assesses the quality level of the feedback on evaluation findings to the relevant stakeholders in a community.

Each Preffi cluster contains different number of variables which could be scored with mark 0 for non assessable, 1 for weak, 2 for moderate and 3 for strong. The final score for each cluster was calculated as the sum of the ratings per variable divided by the maximum possible score for that cluster, and multiplied with 10. The total Preffi rating for the whole project was calculated as an average score of all the cluster scores. Following that procedure, total project ratings on the Preffi 2.0 instrument could range from 3.33 to 10. For an individual project, results can be shown by separate score for each cluster and as a total Preffi score for the whole project. That enables comparison between projects, on both cluster level and the total score. In our research, the measure of agreement between the 3 assessors found in baseline Preffi assessment for total project score is G=0.79 and SEM=0.44. Results for all clusters together for baseline Preffi assessment are G=0.92 and SEM=0.28. Internal consistency scores compared with conventional minimum and Molleman original Preffi study (Molleman, 2005) are showing that results can be perceived as reliable.

3.5.2. Measures of implementation quality

To answer all research questions and follow the research tasks of this doctoral thesis, measures for the implementation quality research had to include 4 scales, each representing a different observatory perspective. After the expert revision of the initial set of generated items by two renowned prevention scientists, surveys were designed and accompanied by a 4 point Likert-type scale: (1) "I absolutely don't agree", (2) "I don't agree", (3) "I agree" and (4) "I completely agree".

- Scale for implementation factors reported by program managers: ***Implementation Factors Questionnaire for Program Managers*** had 31 items in total and has consisted from items covering six implementation factors: standardization, implementers' skills, attitudes, training, support and monitoring. Results on each subscale could range from 1 to 4. Results on this questionnaire are expressed in six subscales, each representing one implementation factor (the lowest of six α=.702 for attitudes subscale and the highest α=.870 for implementers' skills subscale).
- Scale for implementation factors reported by program implementers: ***Implementation Factors Questionnaire for Program Implementers*** has 33 items covering five implementation factors: standardization, attitudes, training, support and monitoring. Results on each subscale could range from 1 to 4. Results on this questionnaire are expressed in five mentioned subscales, each representing one implementation factor (the lowest of six α=.714 for standardization subscale while the highest is α=.808 for support subscale).
- Scale for indicators of implementation quality reported by program implementers: ***Indicators of Implementation Quality Questionnaire for Program Implementers*** has 21 items describing constructs of fidelity, quality, responsiveness and perceived program impact. Results on each subscale could range from 1 to 4. Results on this questionnaire are expressed in four mentioned subscales, each representing one indicator of implementation quality (the lowest of six α=.419 for fidelity subscale while the highest is α=.792 for perceived program impact).
- Scale for indicators of implementation quality reported by program participants had two versions: because some programs had adult participants and some had children participants. ***Indicators of Implementation Quality Questionnaire for Program Participants – adult version*** has 35 items covering dosage, quality of program

delivery, participants' responsiveness and perceived program impact. Results on this questionnaire are expressed in four mentioned subscales, each representing one indicator of implementation quality (the lowest α=.803 for quality of delivery subscale and the highest α=.893 for responsiveness subscale). Results per dosage can be expressed as a number of sessions held or as a percentage of the number of lessons which was delivered opposed to the number which was planned while other three subscales could range from 1 to 4.

Indicators of Implementation Quality Questionnaire for Program Participants – child version has 20 items covering dosage, quality of program delivery, participants' responsiveness and perceived program impact. Results on this questionnaire are expressed in four mentioned subscales, each representing one indicator of implementation quality (the lowest is α=.689 for quality of delivery subscale and the highest is α=.857 for responsiveness subscale). Results per dosage can be expressed as a number of sessions held or as a percentage of the number of lessons which was delivered opposed to the number which was planned while other three subscales could range from 1 to 4.

For all of the 24 included projects, the researcher administered questionnaires for implementation quality to the organization. Managers and program implementers of all 24 organizations filled in questionnaires individually while program participants were assessed during the program's meeting/workshop. Data assessment per implementation measure lasted about 20 minutes for organization managers and program participants and about 30 minutes for program implementers. Research was anonymous for program users but not for managers and program deliverers which are known to public, Department of Health and Social Services as well as to researchers.

3.6. Ethical considerations

Before the start of the whole study procedure, the researchers from the Faculty of Education and Rehabilitation Sciences asked the County of Istria's Department of Health and Social Services for permission to conduct the study, committing to follow all ethical principles common for research with human subjects. The whole research design was also approved by the Ethical Committee of the Faculty of Education and Rehabilitation Sciences before research was conducted.

After the selection of 24 programs, the Department of Health and Social Services organized a meeting with program managers and implementers where the purpose and methods of the study were explained in order to get their approval for inclusion in the study. The research team explained the aim of the study, study design, measurement dates and following steps. Since one of the goals of the Training of Prevention was to improve written project proposals whose quality is directly connected with the financing by the Department, during the meetings with the authorities from the County of Istria, it was arranged that the committee for the year 2012 will take into account which programs have received the Training and that it won't penalize the control group of programs. Participants from the control conditions were informed about that arrangements and told that they will receive the Training for Prevention intervention in 2012, after the whole study and measurement within that research has been completed. Therefore, Training for Prevention intervention was delivered to control group participants in April and May 2012. Also, the Department made an agreement with each of the organizations guaranteeing financing for 2011 on the condition that they continue regular collaboration with research staff.

Each of 24 included organizations signed the agreement with the Department of Health and Social Services and researchers from the Faculty of Education and Rehabilitation Sciences. After that, each of the 24 program's managers and implementers asked users of their programs to consent to participation in the study while parents of participating children had to sign consent for their participation in the research.

Since this dissertation and its results have direct implications on the mental health promotion and prevention practice and policy in the County of Istria, in the presentation of results full names of organizations will not be mentioned to avoid any negative consequences

for programs, their managers and implementers. Their names will be avoided and programs will be presented descriptively. The purpose of the findings presented in this dissertation is to improve the implementation quality of included cohort of programs, so all results will be delivered and presented to the Department of Health and Social Services and to each of the organizations.

4. CHAPTER FOUR: VALIDATION OF MEASURES FOR THE IMPLEMENTATION QUALITY RESEARCH

To answer the research questions of this doctoral thesis, the first research task **is the construction of valid and reliable measures of factors that influence prevention program implementation and indicators of implementation quality based on the implementation literature and existing measures**. As explained in the methods section, items for both types of measures were generated according to theoretical definitions of implementation and the conceptual model of implementation created for this research (see Figure 8, Introduction section). Managers were asked to report on implementation factors; implementers rated both implementation factors and indicators of implementation quality, and participants only reported on program implementation quality. Because some programs had adult participants and some had children participants, two versions of the participant measure of implementation quality were constructed. This chapter will present the validation of the constructed measures, including reliability analyses and validity of each constructed scale.

There was not enough time for preliminary research on the characteristics of the measures used in the current research study due to time limits and policy reasons connected with the Department of Health and Social Care in the County of Istria, who plans to use the results of this research in the process of making the Region's Plan for Health. As a consequence, the construction of scales was done parallel with the main research. It is important to note that in the procedure of scale construction, data collected per program managers, program implementers and program participants from first measurement was used (see Table A4 in Appendix 4, see page 225) while both reliability and construct validity were checked on post measurement data.

Plan of the analyses

Reliability analyses for each of the four scales were conducted using the same procedure. All of the included items were presented to participants with a four point Likert scale: (1) Strongly disagree, (2) Disagree, (3) Agree and (4) Strongly agree. Firstly, items were reverse coded when needed. Items which were describing one theoretical construct, i.e. an implementation factor or an indicator of implementation quality, were entered into reliability analysis: Cronbach's alpha was calculated, their Inter-Item correlation matrix was produced and according to their Item-Total statistics, items were deleted if they weren't consistent with

other items, didn't have high part-whole correlation and if Cronbach's alpha increased when the item was deleted.

Construct validity was checked with the analysis of the dimensionality of each construct within the scale, whether it was an implementation factor or an indicator of implementation. Since we wanted to include items that assess a single theoretical construct, per each of the constructs a principal component factor analysis was carried out to determine how strongly each item loads onto a single component, i.e. the first common factor. Weakly loading items were discarded from the final set of items. Since all individual theoretical constructs are supposed to describe implementation as an overarching construct incorporating all subordinate constructs, a factor analysis was conducted to check the uni-dimensionality of all implementation factors/indicators of implementation quality together.

4.1. Reliability and Construct Validity of Measures of Implementation Factors

4.1.1. Implementation Factors Questionnaire for Program Managers

The initial pool of items for program managers included 36 items that were divided across six implementation factors: 1) program standardization, 2) program implementer's skill, 3) attitudes towards the intervention, 4) training and intervention knowledge, 5) support for implementer and 6) monitoring system. Responses on the original set of items were collected from all program managers (N=24) at the first assessment time point and from 18 of 24 program managers at the post-test measurement point. After conducting reliability analyses, inter-item correlations, and the factor analysis, the final scale included 31 items. Since the scale is reflecting implementation factors that describe the capacity of a system for implementation, the scale was named ***Implementation FactorsQuestionnaire for Program Managers***. Table 4.1 below shows Cronbach's alpha's per each of the six implementation factors reported by program managers, reflecting also items included in each of the theoretical implementation factors and the percentage of variance explained by a single component calculated within principal component analysis. Cronbach's alpha and extracted components per each of the six constructs are indicating that theoretical sub-scales are both reliable and valid.

Table 4.1
Items and results of reliability analysis and construct validity analysis of Implementation Questionnaire for Program Managers

	Cronbach's Alpha
PROGRAM STANDARDIZATION	
Program deliverers are provided with specific guidelines for program delivery. Program deliverers are provided with a course of themes for this program which have to be touched on. Program deliverers are told to follow a schedule of themes. Program deliverers are expected to keep up with set of themes without making changes. Program deliverers are told to conduct program in the same way for all participants. Program deliverers are told that only small changes should be made to the program. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 58.35% OF VARIANCE**	**.849**
PROGRAM IMPLEMENTER'S SKILL	
The program deliverer is skilled at delivering this program. Program deliverer is prepared for the program sessions/ meetings/activities. Program deliverer is conducting core components of our preventive intervention. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 70.02% OF VARIANCE**	**.870**
ATTITUDES TOWARDS THE INTERVENTION	
This intervention makes a real difference in the lives of participants. I am afraid that program effects are short-lived and fade out quickly after the program ends. Our intervention meets the needs of participants in sufficient manner. This program needs to include more meetings or workshops with participants to be effective. This program needs to cover more themes to have more impact. Our program affects behaviour or attitudes of program participants. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 42.03% OF VARIANCE**	**.702**
TRAINING AND INTERVENTION KNOWLEDGE	
I as organization manager invest enough in the development of skills and knowledge program deliverer needs for program implementation. I organize in-service trainings where program deliverers practice skills needed for program implementation. I send our program deliverer to different educations and trainings which can benefit our program implementation. Program deliverer training is covering skills and knowledge needed for program implementation. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 61.91% OF VARIANCE**	**.782**
SUPPORT FOR IMPLEMENTER	
When deliverer encounters difficulties in program implementation, I am available to provide advice. I provide sufficient administrative and technical support for program deliverers throughout program implementation. I provide enough emotional support to program deliverers through different phases of	**.736**

program implementation. Program deliverer is included in supervision arranged by our organization or similar human services. When a problem in implementation arises, I as an organization manager work collaboratively with program deliverer. Program deliverers perceive me as supporting and someone he/she can rely on. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 46.54% OF VARIANCE**	
MONITORING SYSTEM	
I regularly communicate with the program deliverer regarding program implementation. I follow phases of program delivery and I know what is happening on the field. Program deliverer sends me written feedback about the program implementation regularly. I regularly hold meetings with program deliverer to talk about important steps in the process of program implementation. Our organization has a structured employee appraisal form to assess deliverers working quality. I come to the field and watch my staff delivering the program. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 48.20% OF VARIANCE**	**.762**

After the reliability analysis of time one implementation assessment of program managers has shown that items describing six constructs are reliable, reliability on the chosen items was also checked on the data collected at time two, post-test ratings of implementation factors by program managers. Post-test data has also shown that each construct has a high degree of internal consistency: standardization α=.905, program implementer's skill α=.863, attitudes towards intervention α=.688, training and intervention knowledge α=.782, support α=.865 and monitoring α=.800. Per each of the six constructs included in the *Implementation Factors Questionnaire for Program Managers* average summative scores were calculated. Those summative scores were included into the principal component factor analysis to determine their structure. Correlation matrix of six implementation factors assessed by Implementation Factors Questionnaire for program Managers is presented in Table 4.2.

Table 4.2

Correlation matrix of six implementation factors in the Implementation Factors Questionnaire for Program Managers

	Program Standardization	**Implementer skills**	**Attitudes towards intervention**	**Training and intervention knowledge**	**Support**	**Monitoring**
Program standardization	1	.038	.222	.301	.406	.552**
Program implementer's skills		1	.637**	.177	.451*	.341
Attitudes towards intervention			1	.377	.544**	.492*
Training and intervention knowledge				1	.303	.289
Support					1	.590**
Monitoring						1

***** correlation is significant at the level p<.01 * correlation is significant at the level p<.05***

Results of the principal components analysis show that the six implementation factors load onto first component, explaining together 49.43% of the variance. Loadings of implementation factors onto first component are shown in Table 4.3 below.

Table 4.3

Results of factor analysis of Implementation Factors Questionnaire for Program Managers

	First component
Program standardization	.576
Program implementer's skills	.643
Attitudes towards intervention	.800
Training and intervention knowledge	.541
Support	.810
Monitoring	.796
Variance explained	**49.43%**

Regarding the loadings onto first component and the amount of variance explained, Cronbach's alpha was calculated for six implementation factors together, being α=.757 which proves high internal consistency. When the factor structure of the scale was checked on the

data for program managers in post-test measurement, principal component analysis has again revealed that first component explains 45.59% of variance. Cronbach's alpha of α=.727 calculated for six implementation factors in the post measurement with *Implementation Factors Questionnaire for Program Managers* shows that six constructs are both reliable and reflect construct validity. Since we had a repeated measurement of all constructs in two time points, test-retest reliability was also checked. Results are presented in Table 4.4 below. Presented correlations from first and second measurement of the same constructs also show that *Implementation Factors Questionnaire for Program Managers* is consistent over time.

Table 4.4

Results of test-retest reliability between time one and post-test measurement with the Implementation Factors Questionnaire for Program Managers

	Standardization 2	**Implementer skills 2**	**Attitudes 2**	**Training 2**	**Support 2**	**Monitoring 2**
Program standardization 1	.892**					
Program implementer's skills 1		.875**				
Attitudes towards intervention 1			.816**			
Training and intervention knowledge 1				.774**		
Support 1					.614**	
Monitoring 1						.683**

***** correlation is significant at the level p<.01***

4.1.2. Implementation Factors Questionnaire for Program Implementers

Initial pull of items for program managers had altogether 37 items that were divided across five implementation factors. Program implementers reported on following implementation factors: 1) program standardization, 2) attitudes towards the intervention, 3) training and intervention knowledge, 4) support for implementer and 5) monitoring system.

That original set of items was answered by N=51 program implementers from 24 programs in the first measurement of implementation and by N=55 program implementers in the post-test measurement of implementation. Measure covering implementation factors was named ***Implementation Factors Questionnaire for Program Implementers*** (see Table 4.5).

Table 4.5 below shows the results of reliability and construct validity analysis of *Implementation Factors Questionnaire for Program Implementers*: Cronbach's alpha's per each of the five implementation factors administered to program implementers, reflecting items included in each of the theoretical implementation factor and the amount of variance explained by single component calculated within principal component analysis. Cronbach's alpha and extracted components per each of the five implementation factors constructs are indicating that theoretical constructs are both reliable and valid. The final version of ***Implementation Factors Questionnaire for Program Implementers*** has 33 items.

Table 4.5
Final set of items in the Implementation Factors Questionnaire for Program Implementers and results of reliability analysis and construct validity analysis

	Cronbach's Alpha
PROGRAM STANDARDIZATION	
I am provided with specific guidelines for program delivery. I am provided with a course of themes for this program which have to be touched on. I keep up with themes according to the schedule. I am expected to keep up with set of themes without making changes. I conduct program in the same way for all participants. If I make changes to the program, that should be only small changes. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 41.85% OF VARIANCE**	**.714**
ATTITUDES TOWARDS THE INTERVENTION	
This intervention makes a real difference in the lives of participants. I am afraid that program effects are short-lived and fade out quickly after the program ends. Our intervention meets the needs of participants in sufficient manner. This program needs to include more meetings or workshops with participants to be	**.797**

effective. This program needs to cover more themes to have more impact. I like this program very much. This program is a good model for dealing with this problem. I think there are changes that could be made to improve this program. Most of the activities in whole program were in place and adequate. This program only scratches the surface of the problem it is designed to prevent. The activities in this program are comprehensive. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 35.31% OF VARIANCE**	
TRAINING AND INTERVENTION KNOWLEDGE	
My organization invests enough in the development of skills and knowledge I need for the program implementation. My organization provides in-service trainings which give me the possibility to practice skills needed for program implementation. I feel prepared to deliver the intervention. The training I was provided gave me the skills and knowledge needed for program implementation. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 57.01% OF VARIANCE**	**.725**
SUPPORT PROVIDED BY MANAGER	
When I am insecure about the program implementation, I can consult manager of the organization. Management of my organization provides me with sufficient administrative and technical support throughout the whole program implementation. When needed in different phases of program implementation, I can get enough emotional support from my superiors. I am included in supervision of my work where I can talk about experiences and problems connected with program implementation. When a problem in implementation arises, organization manager works with me collaboratively. I perceive organization manager as a person of trust I can rely on. Organization manager possesses skills needed for quality management. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 46.63% OF VARIANCE**	**.808**
MONITORING SYSTEM	
I regularly communicate with organization manager to share the information about the program implementation. Organization manager is along with phases of program delivery and knows what is happening on the field. I regularly hold meetings with my organization manager to talk about important steps in the process of program implementation. Program manager comes to the field and watches me delivering the program. Someone in our organization observes me while conducting the program. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 49.98% OF VARIANCE**	**.727**

After the reliability analysis of time one assessment of program implementers has shown that final set of items describing five implementation factors are reliable, reliability on the chosen items was also checked on the data collected at time two, post-test assessment of program implementers. Post-test data has also shown that each implementation factor construct has a high degree of internal consistency: program standardization α=.730, attitudes towards intervention α=.835, training and intervention knowledge α=.773, support α=.872 and monitoring α=.697.

Per each of the five constructs included in the *Implementation Factors Questionnaire for Program Implementers* average summative scores were calculated. Those average summative scores were included into the principal component analysis to determine their structure. Correlation matrix presented in Table 4.6 shows high inter-correlation among implementation factors seen by program implementers, all of them being significant. Principal component analysis has shown that implementation factors seen by program implementers load on the single factor with very high loadings explaining 59.94% of variance. Results of principal component analysis and Cronbach's alpha α=.821 calculated for this five constructs proves that *Implementation Factor Questionnaire for Program Implementers* is uni-dimensional scale.

Table 4.6
Correlation matrix and loadings onto a single component of five implementation factors in the Structure for Implementation Questionnaire for Program Implementers

	Program Standardization	**Attitudes towards intervention**	**Training and intervention knowledge**	**Support**	**Monitoring**
Program standardization	1	.348**	.472**	.580**	.318*
Attitudes towards intervention		1	.750**	.468**	.309*
Training and intervention knowledge			1	.526**	.552**
Support				1	.622**

Monitoring					1
Loadings on the single component	**.691**	**.751**	**.864**	**.829**	**.723**

***** correlation is significant at the level p<.01 * correlation is significant at the level p<.05***

Principal component analysis conducted on the post-test implementer data collected by the *Implementation Factors Questionnaire for Program Implementers* has also confirmed the same structure of the scale: implementation factors have also loaded onto one component explaining 56.02% of variance. Also, Cronbach's alpha α=.802 at post-test measurement calculated for all five implementation factors reported by implementers together indicates the reliability and internal consistency of scale.

Since we had a repeated measurement of all implementation factors for program implementers in two time points, test-retest reliability was also checked. Results are presented in Table 4.7 below. Presented correlations from first and second measurement of the same constructs also show that *Implementation Factors Questionnaire for Program Implementers* is consistent over time.

Table 4.7
Results of test-retest reliability between time one and post-test measurement with the Implementation Factors Questionnaire for Program Implementers

	Standardization 2	**Attitudes 2**	**Training 2**	**Support 2**	**Monitoring 2**
Program standardization 1	.767**				
Attitudes towards intervention 1		.713**			
Training and intervention knowledge 1			.489*		
Support 1				.813**	

Monitoring 1					.729**

***** correlation is significant at the level p<.01 *correlation is significant at the level p<.05***

4.2. Reliability and Construct Validity of Measures of Implementation Quality

4.2.1. Implementation Quality Questionnaire for Program Implementers

Initial pull of items assessing indicators of implementation quality from implementers' perspective had altogether 28 items that were divided across four indicators of implementation quality. Program implementers reported on following indicators of implementation quality: 1) fidelity, 2) quality of program delivery, 3) participants' responsiveness and 4) perceived program impact. The newly constructed measure was named ***Implementation Quality Questionnaire for Program Implementers*** (see Table 4.8).

Regarding the reliability and construct validity of *Implementation Quality Questionnaire for Program Implementers*, Cronbach's alpha's per fidelity, quality of program delivery, participants' responsiveness and perceived program impact are presented in Table 4.8. It can be seen that selected items for quality of program delivery, participants' responsiveness and perceived program impact are highly internally consistent while Cronbach's alpha for fidelity is below our expectations. Internal consistency of the items under the construct of fidelity could be this low not because of the unreliability of the items but because of the characteristics of Croatian programs which are not evidence-based, which is especially affecting fidelity. Table 4.8 also reflects items included in each theoretical construct i.e. indicator of implementation and the amount of variance explained by a single component calculated within principal component analysis. The final version of *Implementation Quality Questionnaire for Program Implementers* has 21 items.

Table 4.8
Final set of items in the Implementation Quality Questionnaire for Program Implementers and results of reliability analysis and construct validity analysis

	Cronbach's alpha
FIDELITY	
I know what the core components of the program are. I deliver program activities as planned. I think that it is o.k. to leave out some activities as long as they are not core elements of the program. I need to make changes to this program to meet the needs of participants. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 38.27% OF VARIANCE**	**.419**
QUALITY OF PROGRAM'S DELIVERY	
I think that I am skilled in delivering this program. I give feedback to the participants about the way they have conducted a certain activity	**.725**

or exercise. I am prepared for the program sessions/meetings/activities. During sessions I am able to keep most participants active and engaged in the program. I assure active participation of all participants during the program (discussion, opinion expression). **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 43.19% OF VARIANCE**	
PARTICIPANTS' RESPONSIVENESS	
Participants are interested in themes presented in this program. In general, participants stay engaged during the whole meeting/workshop. If I give some homework or assignment to participants, they fulfil it. Atmosphere on the meetings/workshops/activities is positive. Participants from the group are supportive to each other. Participants are excited when going to workshop/activities of this program. When you think about all activities/workshops/meetings which were implemented until now, how many participants have attended regularly: 25%-50% 50% 50-75% >75% How many participants attend each workshop/meeting on average? Less than half about half more than half almost all **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 35.36% OF VARIANCE**	**.700**
PERCEIVED PROGRAM IMPACT	
Participants are changing behaviour in different phases of this program. This program has helped participants in their functioning. This program has helped participants to learn something important and relevant to their lives. This program has improved participant's relationships with others. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 63.39% OF VARIANCE**	**.792**

After the reliability analysis of time one assessment of indicators of implementation quality seen by program implementers has shown that quality of program delivery, participants' responsiveness and perceived program impact have satisfactory internal consistency; reliability on the chosen items was also checked on the data collected at time two. Post-test data has shown similar trends like the data in the first measurement: fidelity has Cronbach's α=.446, quality of program delivery has Cronbach's α=.794, participants' responsiveness has Cronbach's α=.790 and perceived program impact has Cronbach's α=.894.

For all indicators of implementation quality included in the *Implementation Quality Questionnaire for Program Implementers*, average summative scores were calculated. Those summative scores were included into the principal component analysis to determine their structure and to check their uni-dimensionality. Correlation matrix presented in Table 4.9 shows high inter-correlation among indicators of implementation quality seen by program

implementers. Principal component analysis has shown that indicators of implementation quality seen by program implementers have high loadings on the single factor explaining 54.54% of variance. Results of principal components analysis and Cronbach's alpha α=.708 calculated for indicators of implementation quality at time one assessment proves that *Implementation Quality Questionnaire for Program Implementers* is uni-dimensional scale.

Table 4.9
Correlation matrix of four indicators of implementation quality and loadings onto a single component in the Implementation Quality Questionnaire for Program Implementers

	FIDELITY	**QUALITY OF PROGRAM DELIVERY**	**PARTICIPANTS' RESPONSIVENESS**	**PERCEIVED PROGRAM IMPACT**
FIDELITY	1	.260	.331*	.204
QUALITY OF PROGRAM DELIVERY		1	.588**	.455**
PARTICIPANTS' RESPONSIVENESS			1	.454**
PERCEIVED PROGRAM IMPACT				1
Loadings on the single component	.539	.815	.835	.728

***** correlation is significant at the level p<.01 * correlation is significant at the level p<.05***

Principal component analysis conducted on the post-test implementer data collected by the *Implementation Quality Questionnaire for Program Implementers* has also confirmed the same structure of the scale: indicators of implementation quality have also loaded onto just one component explaining 55.69% of variance. Cronbach's alpha α=.711 calculated for post-test assessment of indicators of implementation quality confirms high internal consistency of scale. Since we had a repeated measurement of all indicators of implementation quality for program implementers in two time points, test-retest reliability was also checked. Results are presented in Table 4.10 below. Presented correlations from first and second measurement of the same constructs show that *Implementation Quality Questionnaire for Program Implementers* is consistent over time for quality of program delivery, participants'

responsiveness and perceived program impact. Test-retest results for fidelity again show that fidelity is not so reliable what is probably connected with the fact that Croatian programs are not evidence-based.

Table 4.10
Results of test-retest reliability between time one and post-test measurement with the Implementation Quality Questionnaire for Program Implementers

	FIDELITY 2	**QUALITY OF PROGRAM DELIVERY 2**	**PARTICIPANTS' RESPONSIVENESS 2**	**PERCEIVED PROGRAM IMPACT 2**
FIDELITY 1	.046			
QUALITY OF PROGRAM DELIVERY 1		.529**		
PARTICIPANTS' RESPONSIVENESS 1			.652**	
PERCEIVED PROGRAM IMPACT 1				.683**

***** correlation is significant at the level p<.01 ** correlation is significant at the level p<.05***

4.2.2. Implementation Quality Questionnaire for Program Participants

Some programs had adult participants and others had children participants so measures of indicators of implementation quality for participants had to be adjusted. Regarding the fact that participants of mental health promotion and prevention programs are also children who could have difficulties in understanding the questions, a child's version of the measure of indicators of implementation quality was also designed. Therefore, two versions of a measure were constructed (see Table A2, Appendix 2). Findings for both adult and child version are presented in the subsection below.

Initial pull of items for adult program participants had 45 items that were covering four indicators of implementation quality: dosage, quality of program's delivery, participants' responsiveness and perceived program impact. That original set of items was answered by N=137 adult program participants from 24 programs in the first measurement of implementation quality and by N=231 adult program participants in the post-test measurement of implementation.

Reliability and construct validity analysis was done for three indicators of implementation, quality of program's delivery, participants' responsiveness and perceived program impact in order to confirm the theoretical background of the scale. Dosage couldn't be tested for reliability and validity because it was represented with only one item. Table 4.11 demonstrates that after unreliable items were discarded, items under the indicators of implementation quality constructs are both reliable and show construct validity. The scale was named ***Implementation Quality Questionnaire for Program Participants – adult version*** and in its final version has 35 items (see Table 4.11). Regarding the reliability and construct validity of adult version of measure for participants, Cronbach's alpha's per quality of program delivery, participants' responsiveness and perceived program impact are presented in Table 4.11. It can be seen that selected items for quality of program delivery, participants' responsiveness and perceived program impact are highly internally consistent. Reliability and construct validity for dosage was not calculated because dosage is represented with just one item reflecting percentage of intervention delivered to the program participants. Table 4.11 also reflects items included in each theoretical construct i.e. indicator of implementation and the amount of variance explained with a single component calculated within principal component analysis.

Table 4.11
Final set of items in the Indicators of Implementation Quality Questionnaire for Program Participants and results of reliability analysis and construct validity analysis

	Cronbach's alpha
QUALITY OF PROGRAM'S DELIVERY	
Program deliverer is skilful in program implementation. Program deliverer gives us a feedback about the way we have conducted certain activity or exercise. Program deliverer seemed underprepared. Program deliverer represents activities in a highly engaging manner. I perceive the rhythm of program implementation as adequate. If needed, program deliverer repeats some program activities for participants. I like the working style of program deliverer. It is evident that program deliverer is positive towards the program and that he/she believes in its impact. If I have some questions, I can talk to program deliverer. Program deliverer is doing a good job and I trust him. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 40.59% OF VARIANCE**	**.803**
PARTICIPANTS' RESPONSIVENESS	
I am interested in themes presented in this program. In general, I stay engaged during the whole meeting/workshop. I am highly collaborative during meetings/workshops/activities. If I get homework or assignment on the meeting, I fulfil it. Atmosphere on the meetings/workshops/activities is positive. I perceive others from the group as supportive. There are activities in this program that I refuse to participate in. During the activities conduction, program deliverer assures active participation of all the participants (discussion, opinion expression). I am bored in this program. I like this program very much. This program is a good model for dealing with this problem. This program has fulfilled my expectations. I meet interesting people because of this program. I think there are changes that could be made to improve this program. This program could be more effective if it covered more themes. Most of the activities within the program were appropriate and adequate. Program activities are comprehensive. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 41.91% OF VARIANCE**	**.893**
DOSAGE	
How many workshops/meetings have been held until now?	
PERCEIVED PROGRAM IMPACT	
I was changing behaviour in different phases of this program. This program has helped me in my functioning. This program helped me to learn something important and relevant to my life. This program has improved my relationships with others. I have a feeling that I have gained after each workshop/meeting. I am afraid that program effects are short-lived and fade out quickly after the program ends.	**.875**

This program met my needs. I think about some themes of this program in my everyday life. I will change something in my behaviour in the future because of this program. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 51.12% OF VARIANCE**	

After the reliability analysis of time one assessment of indicators of implementation quality seen by adult program participants has shown that quality of program delivery, participants' responsiveness and perceived program impact have satisfactory internal consistency; reliability on the chosen items was also checked on the data collected at time two. Post-test data has shown similar trends like the data in first measurement: quality of program delivery has Cronbach's α=.787, participants' responsiveness has Cronbach's α=.761 and perceived program impact has Cronbach's α=.741.

For all indicators of implementation quality included in the *Implementation Quality Questionnaire for Program Participants* besides dosage, average summative scores were calculated. Those summative scores were included into the principal component analysis to determine their structure and to check the uni-dimensionality of total scale. Correlation matrix presented in Table 4.12 shows high inter-correlation among indicators of implementation quality seen by program participants, all of them being significant at the level $p<.01$. Principal component analysis has shown that indicators of implementation quality seen by program participants have high loadings on the single factor explaining 79.95% of variance. Total Cronbach's alpha α=.832 calculated for all indicators of implementation quality together also proves that Indicators of Implementation Quality Questionnaire for Program Participants – adult version is reliable uni-dimensional scale.

Table 4.12
Correlation matrix of three indicators of implementation quality and loadings onto a single component in the Implementation Quality Questionnaire for Program Participants – adult version

	QUALITY OF PROGRAM DELIVERY	PARTICIPANTS' RESPONSIVENESS	PERCEIVED PROGRAM IMPACT
QUALITY OF PROGRAM DELIVERY	1	.735**	.498**
PARTICIPANTS' RESPONSIVENESS		1	.722**
PERCEIVED PROGRAM IMPACT			1
Loadings on the single component	.848	.940	.841

***** correlation is significant at the level p<.01***

Principal component analysis conducted on the post-test of adult participant data collected by the *Implementation Quality Questionnaire for Program Participants - adult version* has also confirmed the same structure of the scale: indicators of implementation quality have also loaded onto just one component explaining 79.38% of joint variance. Also, total Cronbach's alpha α=.852 calculated for average summative results quality of program delivery, participants' responsiveness and perceived program impact on the post-test data shows high internal consistency of the scale.

Since we had a repeated measurement of quality of program delivery, participants' responsiveness and perceived program impact for program participants in two time points, test-retest reliability was also checked. Results are presented in Table 4.13 below. Presented correlations from first and second measurement of the same constructs also show that *Implementation Quality Questionnaire for Program Participants* is consistent over time.

Table 4.13
Results of test-retest reliability between time one and post-test measurement with the Implementation Quality Questionnaire for Program Participants – adult version

	QUALITY OF PROGRAM DELIVERY 2	PARTICIPANTS' RESPONSIVENESS 2	PERCEIVED PROGRAM IMPACT 2
QUALITY OF PROGRAM DELIVERY 1	.463*		
PARTICIPANTS' RESPONSIVENESS 1		.604**	
PERCEIVED PROGRAM IMPACT 1			.911**

***** correlation is significant at the level p<.01 * correlation is significant at the level p<.05***

Initial pull of items for children program participants had 22 items that were covering four indicators of implementation quality: quality of program's delivery, participants' responsiveness, perceived program impact and dosage. That original set of items was answered by N=297 child program participants from 24 programs in the first measurement of implementation quality and by N=513 child program participants in the post-test measurement of implementation quality. Those participants were assessed with the child version of the ***Implementation Quality Questionnaire for Program Participants – child version*** because they either were children or had difficulties understanding the language or writing (e.g. Roma parents).

Reliability and construct validity analysis of child version of *Implementation Quality Questionnaire for Program Participants* was done for three indicators of implementation, quality of program's delivery, participants' responsiveness and perceived program impact in order to confirm the theoretical background of the scale. Table 4.14 demonstrates that after the unreliable items were discarded, items under the indicators of implementation quality constructs are both reliable and show construct validity. Reliability and construct validity for dosage was not calculated because dosage is represented with just one item reflecting percentage of intervention delivered to the program participants. ***Implementation Quality Questionnaire for Program Participants – child version*** in its final version has 20 items (see Table 4.14). It can be seen that selected items for quality of program delivery, participants' responsiveness and perceived program impact are highly internally consistent and show construct validity. Table 4.14 also reflects the percentage of variance explained with individual items' loadings onto a single component calculated within principal component analysis.

Table 4.14
Final set of items in the Indicators of Implementation Quality Questionnaire for Program Participants – child version and results of reliability analysis and construct validity analysis

	Cronbach's alpha
QUALITY OF PROGRAM'S DELIVERY	
Program deliverer is doing a good job. Program deliverer talks with us kids about the way we have done a certain activity. Program deliverer presents activities cheerfully. Program deliverer repeats some program activities if I ask him to. I like program deliverer. If I have some questions, I can talk to program deliverer. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 39.94% OF VARIANCE**	**.689**
PARTICIPANTS' RESPONSIVENESS	
I find this program interesting. If I get homework or assignment on the meeting, I usually fulfil it. I have a lot of fun during this program. During the program, deliverer asks me what I think. I am bored in this program. I feel excited when going to workshop/activities of this program. I like this program. Activities in this program are good. It would be great if this program lasted longer. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 52.02% OF VARIANCE**	**.857**
DOSAGE	
How many workshops/meetings have been held until now?	
PERCEIVED PROGRAM IMPACT	
I have improved my behaviour because of this program. This program has helped me. This program taught me something important. On each program meeting/activity, I learn something new. **FIRST COMPONENT IN PRINCIPAL COMPONENT ANALYSIS EXPLAINING 58.12% OF VARIANCE**	**.749**

After the reliability analysis of time one assessment of indicators of implementation quality seen by children program participants has shown that quality of program delivery, participants' responsiveness and perceived program impact have satisfactory internal consistency; reliability on the chosen items was also checked on the data collected at time two. Post-test data has shown similar trends like the data in first measurement: quality of program delivery has Cronbach's α=.780, participants' responsiveness has Cronbach's α=.903 and perceived program impact has Cronbach's α=.844.

For all indicators of implementation quality included in the *Implementation Quality Questionnaire for Program Participants – child version* besides dosage, average summative scores were calculated. Those summative scores were included into the principal component analysis to determine their structure and to check the uni-dimensionality of total scale. Correlation matrix presented in Table 4.15 shows high inter-correlation among indicators of implementation quality seen by program participants which have fulfilled child version, all of them being significant. Principal component analysis has shown that indicators of implementation quality seen by program participants in child version of scale load very highly on the single factor explaining 73.92% of variance. Total Cronbach's alpha α=. 795 calculated per quality of delivery, participants' responsiveness and perceived program impact proves that Implementation Quality Questionnaire for Program Participants – child version is a reliable uni-dimensional scale.

Table 4.15
Correlation matrix of three indicators of implementation quality and loadings onto a single component in the Implementation Quality Questionnaire for Program Participants – child version

	QUALITY OF PROGRAM DELIVERY	**PARTICIPANTS' RESPONSIVENESS**	**PERCEIVED PROGRAM IMPACT**
QUALITY OF PROGRAM DELIVERY	1	.693**	.499**
PARTICIPANTS' RESPONSIVENESS		1	.628**
PERCEIVED PROGRAM IMPACT			1
Loadings on the single component	.852	.907	.818

***** correlation is significant at the level p<.01***

Principal component analysis conducted on the post-test child participant data collected by the *Implementation Quality Questionnaire for Program Participants – child version* has also confirmed the same structure of the scale: indicators of implementation quality have also loaded onto just one component explaining 78.15% of joint variance. Total Cronbach's alpha α=.867 calculated for post-test assessment with *Implementation Quality Questionnaire for Program Participants – child version* again confirms high degree of internal consistency.

Since we had a repeated measurement of quality of program delivery, participants' responsiveness and perceived program impact for program participants – child version of scale in two time points, test-retest reliability was also checked. Results are presented in Table

4.16 below. Presented correlations from first and second measurement of the same constructs also show that *Implementation Quality Questionnaire for Program Participants – child version* is consistent over time.

Table 4.16
Results of test-retest reliability between time one and post-test measurement with the Indicators of Implementation Quality Questionnaire for Program Participants – child version

	QUALITY OF PROGRAM DELIVERY 2	**PARTICIPANTS' RESPONSIVENESS 2**	**PERCEIVED PROGRAM IMPACT 2**
QUALITY OF PROGRAM DELIVERY 1	.490*		
PARTICIPANTS' RESPONSIVENESS 1		.890**	
PERCEIVED PROGRAM IMPACT 1			.606*

***** correlation is significant at the level p<.01 ** correlation is significant at the level p<.05***

4.3. Discussion regarding the construction and validation of measures

Going back to the first research task in this doctoral thesis, "*To construct valid and reliable measures of implementation quality based on implementation literature and existing measures",* it could be stated that results show that construction of measures for this implementation research was successful.

Measures *Implementation Factor Questionnaire for Program Managers, Implementation Factors Questionnaire for Program Implementers, Implementation Quality Questionnaire for Program Implementers* and *Implementation Quality Questionnaire for Program Participants* – adult and child version are internally consistent, load very highly on single component in the factor analyses and show good test-retest reliability. **It could be concluded that the first research task is fulfilled: reliability and construct validity analyses which have been undertaken show that all constructed measures are both reliable and demonstrate construct validity**.

The first research task of this dissertation was met with construction of valid and reliable measures for implementation factors and indicators of implementation quality per three different types of informants. As explained before, items for implementation quality scales were generated according to theoretical definitions of implementation factors and indicator of implementation quality, based on the conceptual model of implementation created for this research (see Figure 8, Introduction section). Managers were giving reports on implementation factors; implementers rated both implementation factors and indicators of implementation quality while participants were asked to report on indicators of implementation quality. Since all constructed measures have good metric characteristics, theoretical background of the conceptual model was confirmed.

5. CHAPTER FIVE: LEVEL OF IMPLEMENTATION QUALITY IN PREVENTIVE PROGRAMS IN ISTRIA

This chapter aims to follow the research tasks two, three and four presented in the Aims and research questions section:

2. **To explore the level and variation of implementation quality in preventive programs in Istria.**

3. **To explore the differences in perception of implementation quality between program managers, program implementers and program participants.**

4. **To explore the relationships of implementation factors and indicators of implementation quality.**

These research tasks aim to describe the implementation quality and the implementation process in the cohort of 24 mental health promotion and prevention programs from the County of Istria. As it was explained in the Methods section, to describe the implementation of those programs, data on implementation factors and indicators of implementation quality was gathered from program managers, program implementers and program participants approximately after delivery of one third of the program. There were five programs for which data from their managers, implementers and participants was not included in the mid-intervention analysis. Three different programs such as *Substance abuse prevention for parents, Substance abuse prevention for teachers,* and *Substance abuse prevention* only had one session and had data collection only at the end of program delivery. Two programs, *Media literacy* (four sessions) and *Underage drinking prevention* (five sessions) could not organize collection of data in mid-delivery, so they are also not represented in this chapter. Out of 24 total programs, there were 19 managers reporting on implementation factors at mid-intervention point. Since some programs had several implementers, 50 implementers from 19 programs had given their ratings of implementation factors and implementation quality at mid-intervention point. At the mid-intervention, 454 participants took part in the research (see Table A4 in Appendix 4 section, see page 225).

Analysis plan

This chapter will show descriptive results based on all data collected from the mid assessments conducted after approximately first third/half of the program implementation period. Descriptive results will serve as an orientation for exploration of the level and variation of implementation quality. In the next section, through the correlation between two different sources of data on either implementation factors or indicators of implementation quality, associations of different perceptions will be tested. The presented correlations between the perceptions of the two types of informants are Spearman's rho since Shapiro-Wilk's test of distribution normality has shown that some variables are not distributed normally (see Table A5 in Appendix 5, page 230). Regarding the research task which is concerned with the relationship of implementation factors and indicators of implementation quality, besides the simple correlation, in order to explore the association, it was tested if factors could be seen as predictors of implementation quality. Since implementers and participants who were the source of data are not independent but rather nested within programs, the association of implementation factors and indicators of implementation quality was analysed with hierarchical linear modelling (HLM) within SAS software. HLM is a form of analysis which is used to analyse data when participants are nested within units and thereby violate assumptions of independence. HLM simultaneously investigates relationships within and between hierarchical levels of grouped data, thereby making it more efficient at accounting for variance among variables at different levels than other existing analyses (Singer, 1998). Part of the chapter which is examining that association of implementation factors and indicators of implementation quality will show the hierarchical linear models. At the end of the chapter the presented results will be discussed and conclusions about the level of implementation quality will be presented.

5.1. Level and variation of implementation quality in preventive programs in Istria

The implementation factors, which reflect the structural capacity for implementation, were assessed by the self-report of program managers (*Implementation Factors Questionnaire for Program Managers)* and program implementers (*Implementation Factors Questionnaire for Program Implementers).* Indicators of implementation quality were assessed through self-reports of program implementers (*Implementation Quality Questionnaire for Program Implementers*) and self-reports of program participants (*Implementation Quality*

Questionnaire for Program Participants – adult and child version). All results for implementation factors and indicators of implementation quality besides dosage could range from one to four. Since this dissertation represents the first research on implementation quality in Croatia and no standards exist yet, the theoretical median of 2.5 was chosen as a cut-off point, meaning that results below that value are considered to indicate insufficient structural capacity for implementation, i.e. a low level of implementation quality.

Since there is an ethical issue regarding the usage of full organization and program names, when results are presented, programs will be referred to in a simple description of program content in order to protect the identity of organizations and to avoid negative implications of the research. After this dissertation is defended, all organizations will receive feedback with individual results and suggestions.

5.1.1. Level of implementation factors seen by program managers and program implementers: findings

Table 5.1 shows the averaged results for the cohort of programs included in this research, means and standard deviations for each implementation factor, specified for each of the two groups of informants.

Table 5.1
Descriptive statistics per implementation factors averaged across 24 programs, seen from the perspective of program managers and program implementers

	Program standardization		**Program implementer's skill**		**Attitudes towards intervention**		**Training and intervention knowledge**		**Support for program implementer**		**Monitoring system**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Program managers	2.55	0.65	3.42	0.48	2.75	0.40	2.63	0.50	3.24	0.51	2.64	0.43
Program implementers	2.87	0.58			2.93	0.38	2.88	0.64	3.41	0.51	2.98	0.60
Correlations	**.228**		**-**		**.597****		**.358***		**.524****		**.472****	

***** correlation is significant at the level p<.01; * correlation is significant at the level p<.05***

Program managers report the lowest results for three out of six implementation factors: program standardization (M=2.55, SD=0.65), training and intervention knowledge (M=2.63, SD=0.50) and monitoring system (M=2.64, SD=0.43). The highest results for implementation

factors reported by program managers are found for the program implementer's skill (M=3.42, SD=0.58) and support for program implementers (M=3.24, SD=0.51).

Program implementers consistently report slightly higher levels of all implementation factors than program managers. In their view, program standardization is the lowest estimated implementation factor (M=2.87, SD=0.58) with ratings of training and intervention knowledge which are really close (M=2.88, SD=0.64) while support for program implementers is given the highest evaluation (M=3.41, SD=0.51).

Since Table 5.1 shows the averaged results for each implementation factor, the same results are also specified per each of the 19 out of 24 programs separately (see Table 5.2). Managers' scores are represented as they were collected while score for implementers was averaged among the number of implementers which are delivering each program. Levels of assessed implementation factors for each of the 19 programs are graphically presented in Figures 10 – 15. Unfortunately, program implementers of the *Mentor programme* did not send their reports (program 2), so their results cannot be compared with program managers.

Table 5.2

Mean results per each of the implementation factors reported by program managers and program implementers of 24 programs included in the sample.

Pr	Standardization M	Standardization PI	Implementer skills M	Attitudes M	Attitudes PI	Training M	Training PI	Support M	Support PI	Monitoring M	Monitoring PI
1	1.5	1.5	4	3.67	3.27	3.5	3.25	4	4	2.83	3.2
2	2.33		3.67	3		2.25		3		1.83	
3	2.67	2.08	3	2.17	2.59	2.5	2.5	2.83	2.36	2.67	2
4											
5	1	2.83	4	3	3.07	2.25	2.83	2.17	3.37	2.33	2.9
6											
7											
8	1.33	3	3.67	2.5	3.18	2.25	3	3.5	3.71	2.67	2.8
9	2.83	2.83	3	2.67	2.64	2.5	2.83	2.5	3.1	2.83	3.45
10	3.33	3.21	4	3.17	3.26	2.25	3.16	3.67	3.65	3	2.87
11	2.17	2.58	3.33	2.17	2.73	2.25	2.5	2.83	2.93	2	2.7
12	2.83	2.83	4	3	3.27	4	3.5	4	4	3.17	3.8
13	2.67	3.23	3	2.33	2.79	2	2.8	3.67	3.6	2.83	3.28
14	3.17	2.83	4	3.33	3.45	1	3.25	4	3.14	4	3.4
15	2.67	3.08	3.33	2.67	2.64	2.75	3.17	3	3.57	2.5	3.1
16	2.67	2.33	3	2.17	2.52	2.5	2.06	2.83	2.86	2.67	2.4
17											
18	2.67	2.61	3	2.5	2.76	2.75	2.42	2.83	3	2.33	2.47
19											
20	3.25	3.33	4	3.67	3.09	4	3.5	4	4	3.75	3.5
21	1.8	2.97	3.67	2.83	3.18	1.5	4	2.83	3.75	2.5	3.5
22	2.33	2.94	3	2.67	2.88	1.5	2.82	3.5	3.79	2.17	3.43
23	1.5	1.83	2.33	2.83	3.09	2.75	2.75	2.83	2.57	1.83	2
24	1.83	3.5	3.67	2.83	2.64	1.75	2.5	3.17	3	2.17	2.8

The comparison of results on implementation factors shown in Table 5.2 and Figure 9, allow us to make multiple types of comparisons. First, we may compare the concordance of reports of program managers and program implementers per each assessed implementation factor as well as across all factors. Concordance is viewed as a degree of agreement between the views on implementation factors collected from managers and implementers where variation in assessments is not larger than 1.0. For example, programs nine (*Parenting programme III)*, sixteen (*Self-confidence training*), eighteen (*Parenting programme VI*) and twenty (*MH promotion through volunteerism*) have very high concordance of reports by managers and implementers. Second, the data show which implementation factors got the highest scores per each of 24 programs in the sample. Third, Figure 9 offers a comparison of results on implementation factors and gives an overall impression on the differences in structural capacity between the 24 programs. For example, when structural capacity represented by implementation factors is analysed together for managers' and implementers' ratings, programs three (*Parenting programme 1)*, sixteen (*Self-confidence training*) and eighteen (*Parenting programme VI*) have results below three. Also, program twenty three (*Parenting programme VII*)is close to that value. The highest results i.e. the highest system capacity, support and readiness for implementation can be seen for one (*MH promotion through the theatre)*, twelve (*Creative free time programme I.*) and twenty (*MH promotion through volunteerism*). In the text bellow, results per each of the implementation factors will be reviewed separately.

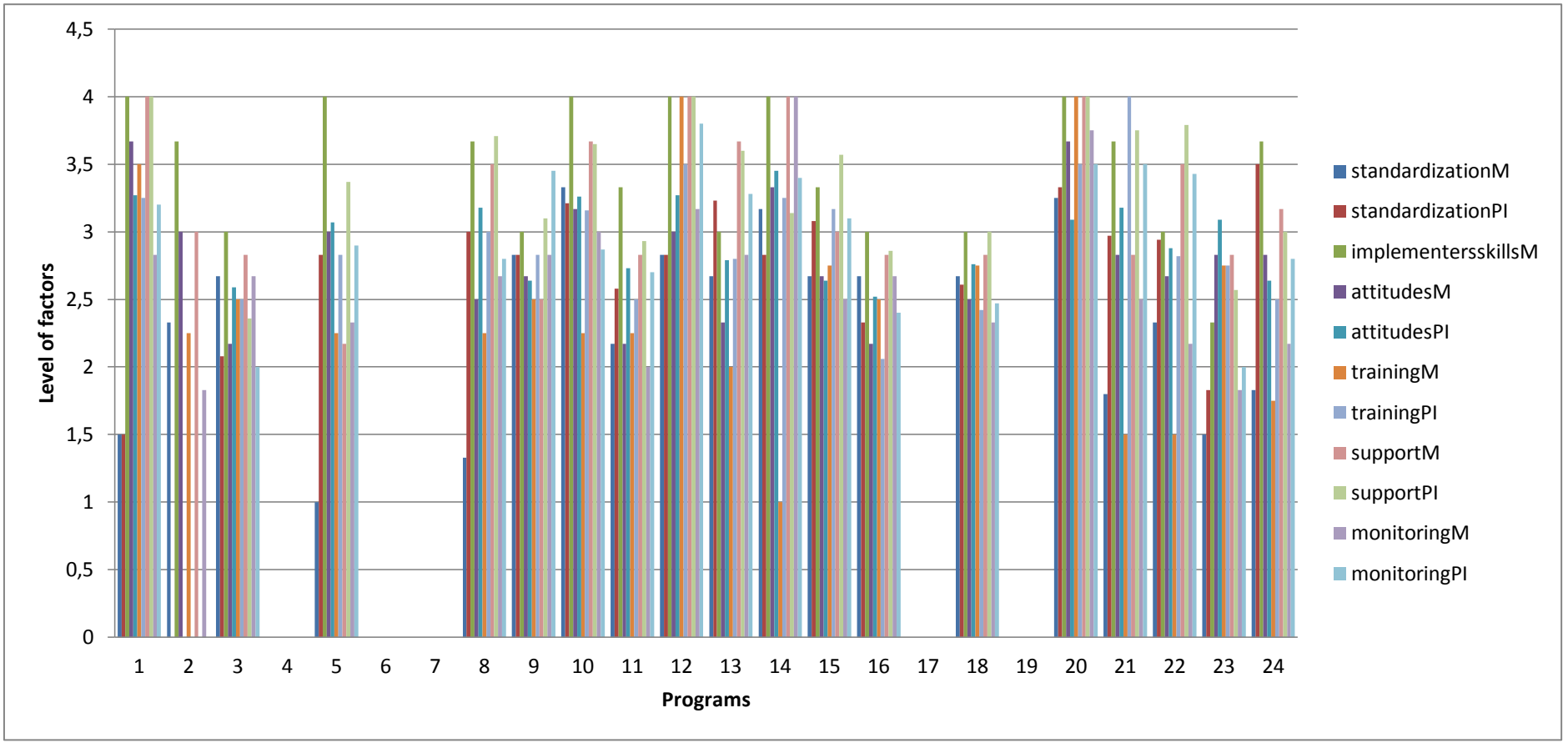


Figure 9. Comparison of reports of program managers and program implementers on implementation factors for all programs in the sample with mid-intervention assessment.

Level of program standardization

As Table 5.2shows, the results for standardization reported by program managers range from 1 for program number five (*Training for the group leaders*) to 3.20 for program twenty (*MH promotion through volunteerism).*The scores of program implementers range from 1.5 (program number one: *MH promotion through the theatre*) to 3.50 for program twenty four (*Parenting programme VIII)*. Levels of standardization, as assessed by program managers and program implementers, can be better compared when presented graphically (see Figure 10). Program implementers mostly report higher levels of program standardization than program managers do. This was the case for 12 of the 19 programs. Higher levels of program standardization have been reported by program managers only for program number three (*Parenting programme I*), ten (*Parenting programme IV*), fourteen (*Parenting programme V*) and sixteen (*Self-confidence training*). Correlation between reports of program managers and program implementers on the level of program standardization is r_s=.228 (see Table 5.1) which is positive but not significant.

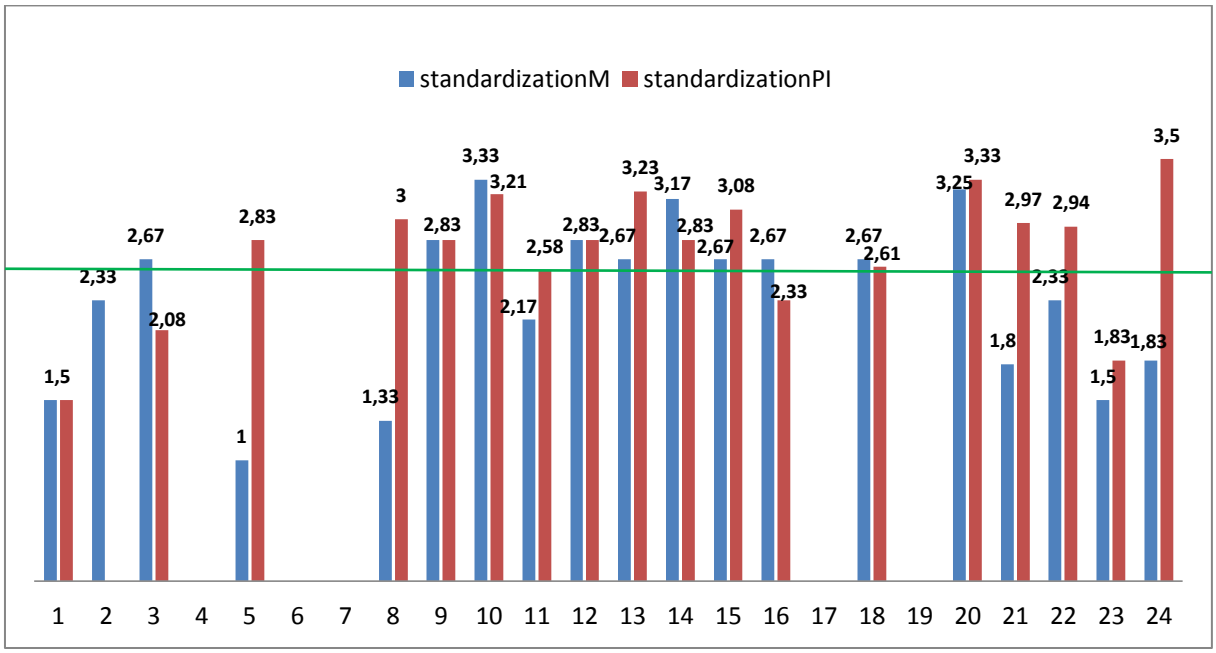


Figure 10.Level of program standardization seen by program managers and program implementers for all programs in the sample with mid-intervention assessment.

Implementers' skills

Implementer' skills were only reported by program managers and results are evidently high and show low variability (see Figure 11). They range from 2.33 (program number twenty three: *Parenting programme VII*) to maximum of 4 (program number one, five, ten, fourteen, twenty). It can be stated that program managers perceive their program implementers as skilled and competent. Managers clearly consider them as well prepared, skilful and engaged in implementing the core components of the program. Nevertheless, these scores also show room for improvement in a range of programs.

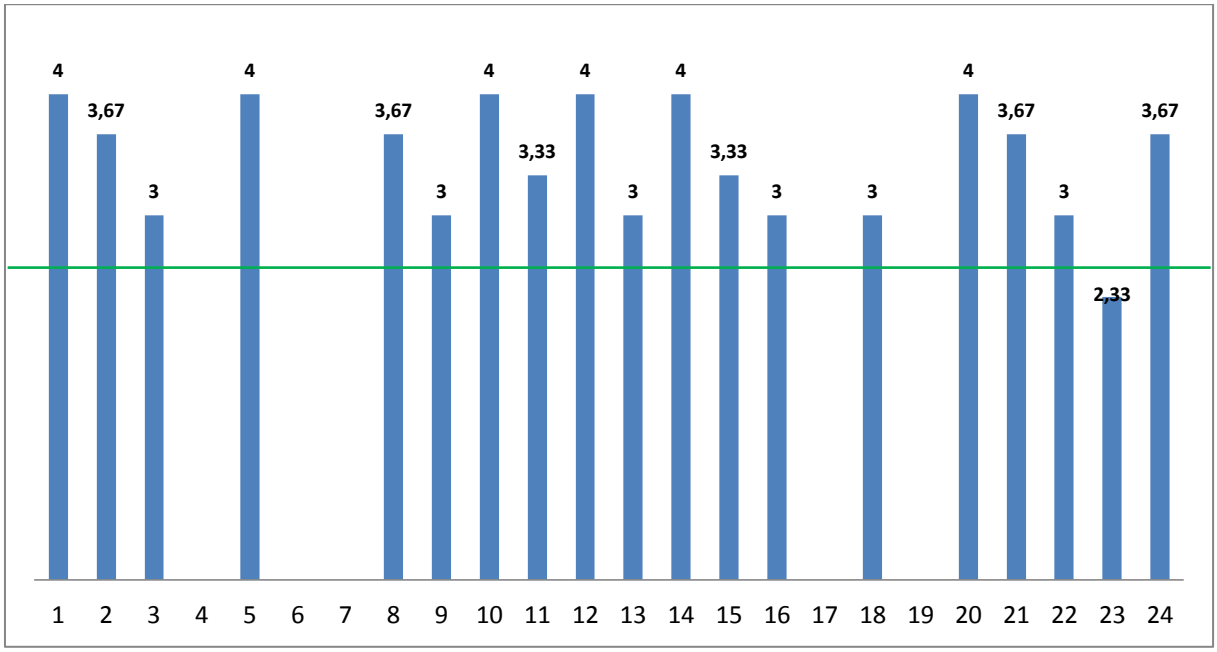


Figure 11. Level of implementers' skills reported by program managers for all programs in the sample with mid-intervention assessment.

Attitudes towards intervention programs

Both Table 5.2 and Figure 12 show that in general both program managers and program implementers have positive attitudes towards their intervention programs, their beliefs that the intervention causes differences in the lives of participants, meets their needs and is a good model for dealing with the issue in question. Several programs have low managers reports regarding attitudes: program three (*Parenting programme I*, manager's M=2.17; implementers' M=2.59), eleven (*Substance abuse prevention in the community*, manager's M=2.17, implementer's M=2.73) and sixteen (*Self-confidence training*, manager's M=2.17 and implementer's M=2.52). The attitude scores range to the maximum of 3.67 for programs one (*MH promotion through the theatre*) and twenty (*MH promotion through volunteerism*). The highest attitude score among program implementers, M=3.45, was found for program fourteen (*Parenting programme V*). The correlation between reports from program managers and program implementers of r_s=.597, p<.01, (see Table 5.1) shows that there is a high positive association in the attitudes towards programs between the two different perspectives i.e. program managers and program implementers have similar attitudes.

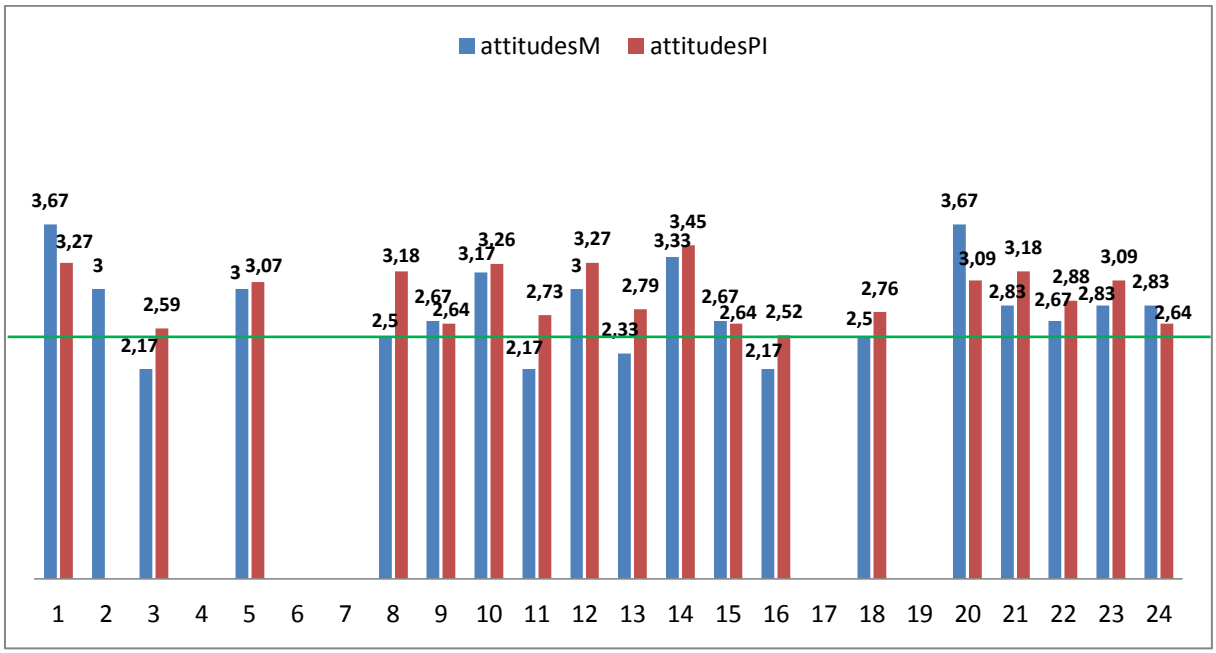


Figure 12. Level of attitudes towards the intervention reported by program managers and program implementers for all programs in the sample with mid-intervention assessment.

Level of training and intervention knowledge

Reports on how program managers and program implementers assess the available training level in their mental health promotion or prevention program are shown in Table 5.2 and Figure 13. Program managers' report on training level ranges from the lowest possible score of 1 (program fourteen, *Parenting programme V*) to the maximum score of 4 (for programs twelve: *Creative free time programme I* and twenty: *MH promotion through volunteerism*). There exists a large variability between the levels of training and intervention knowledge reported by program managers. There are a lot of programs whose managers report about the level of training and intervention knowledge below the reference score of 2.5. Managers from 10 out of overall 19 programs critically view the level of training and intervention knowledge, stating that they as program managers do not invest enough in the in-service trainings and overall development of skills and knowledge of those who implement the intervention in practice.

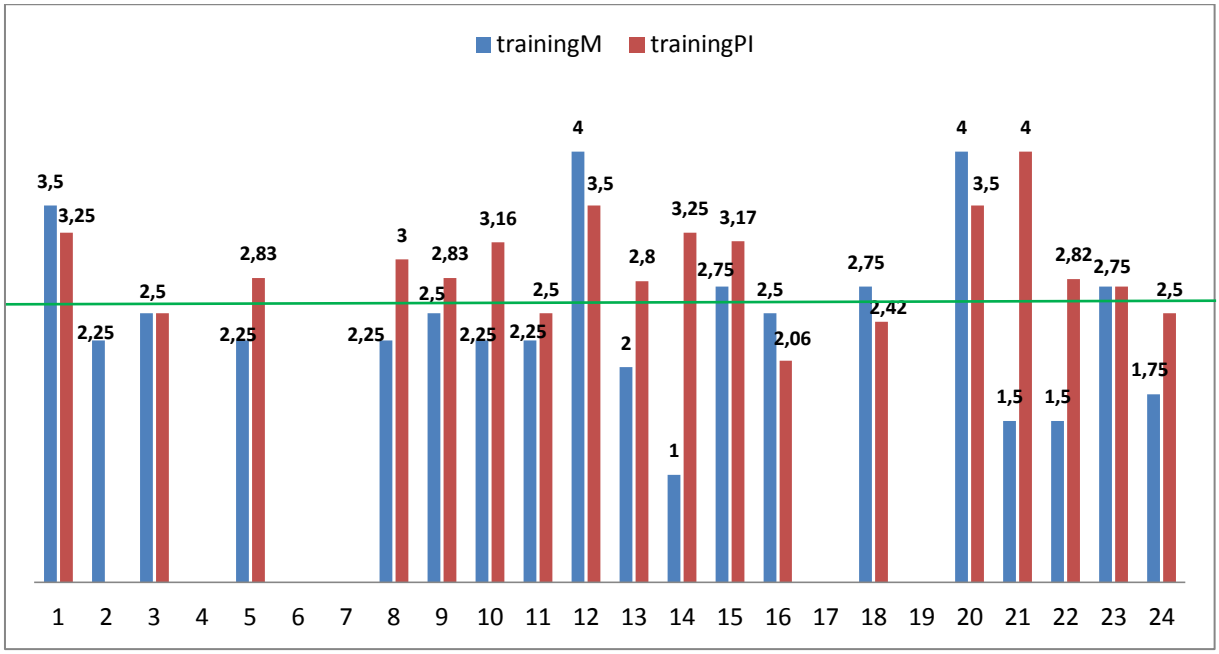


Figure 13. Level of training and intervention knowledge reported by program managers and program implementers for all programs in the sample with mid-intervention assessment.

The assessments of program implementers on the level of training range from 2.06 (for program sixteen: *Self-confidence training*) to 3.45 (for program fourteen: *Parenting programme V*). The variation in their view on training and intervention knowledge is less than that of the program manager's, nevertheless also the assessments of program implementers reflect the existence of different levels of training and knowledge in the cohort of programs. Overall, they present a more positive picture about the training levels in their programs in

comparison to the assessment of their managers. Only program implementers from program sixteen (*Self-confidence training*, implementer's M=2.06) gave a sore below the reference value of 2.5. Correlation of r_s=.358* between reports of program managers and program implementers also shows that there is small positive concordance between the views of program managers and program implementers on the level of training and intervention knowledge. The largest discrepancy in views is seen for programs fourteen (*Parenting programme V*) and twenty one (*MH promotion through dance*). For those two programs, managers report about minimum investments in the training and knowledge of program implementers while program implementers of the same program consider the level of their training and knowledge as sufficient.

Level of support for implementers

Regarding the perceived level of support that program managers and organizations give to program implementers, program managers and program implementers seem to have concordant views (see Figure 14). The correlation of r_s=.524** between the assessments of program managers and program implementers shows high positive association.

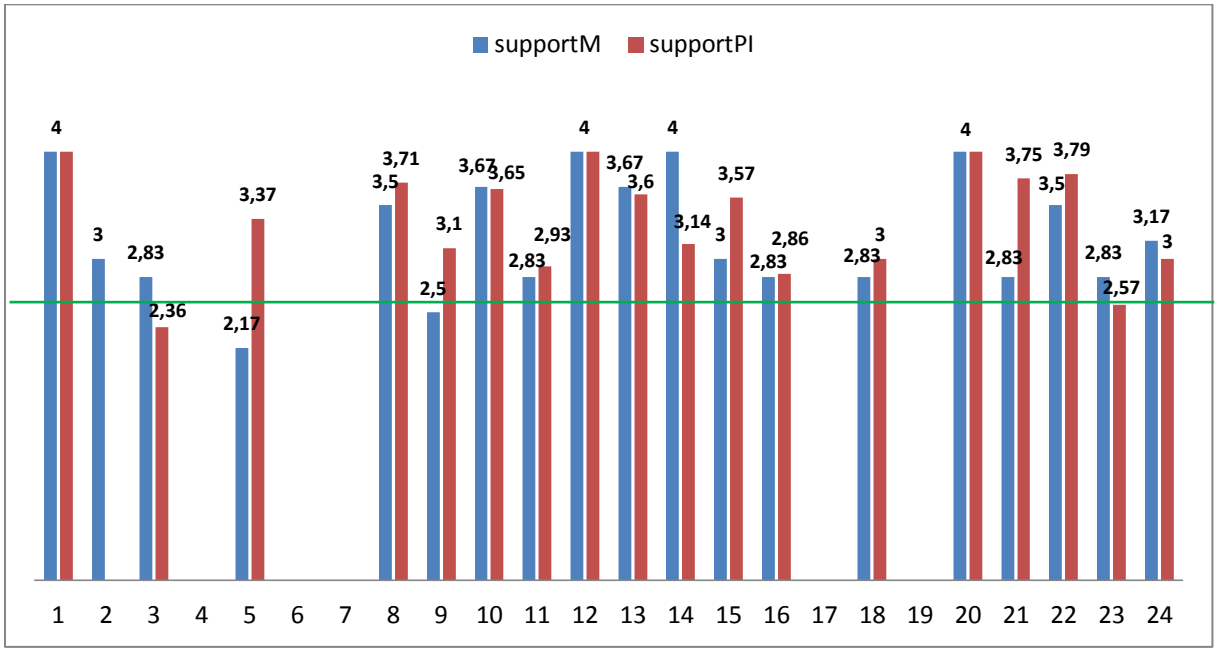


Figure 14.Level of the perceived support for program implementer reported by the program manager and program implementers for all programs in the sample with mid-intervention assessment.

Perceived support for program implementers reported by program managers ranges from 2.17 for program five (*Training for the group leaders*) to the maximum of 4 for programs one (*MH promotion through the theatre*), twelve (*Creative free time programme I.*),

fourteen (*Parenting programme V*) and twenty (*MH promotion through volunteerism*). From the program implementers perspective, the lowest perceived support of 2.36 is received in program three (*Parenting programme* 1) while some program implementers report the maximum possible support score (program one: *MH promotion through the theatre*, program twelve: *Creative free time programme I*, and twenty: *MH promotion through volunteerism*). For some programs, especially program five, there seems to be an evident discrepancy between the views of managers and implementers about the provided support for implementers, being more negatively evaluated by program managers.

Level of monitoring

Reports of program managers and program implementers about the level of monitoring, presented in Table 5.2 and Figure 15, also show significant association of reports from two groups of informants (r_s=.472, p<.001). It seems that program managers and program implementers mainly agree on the level of monitoring in their program. In the assessment of program managers monitoring ranges from 1.83 for programs two (*Mentor programme*) and twenty three (*Parenting programme VII*) to 4 for program fourteen (*Parenting programme V*).Program implementers' reports on monitoring range from 2 for programs three (*Parenting programme I*) and twenty three (*Parenting programme VII*) to 3.80 for program twelve (*Creative free time programme I*). It seems that overall program implementers report a higher level of program monitoring than program managers do, but this could be the result of differences in their perspectives and roles.

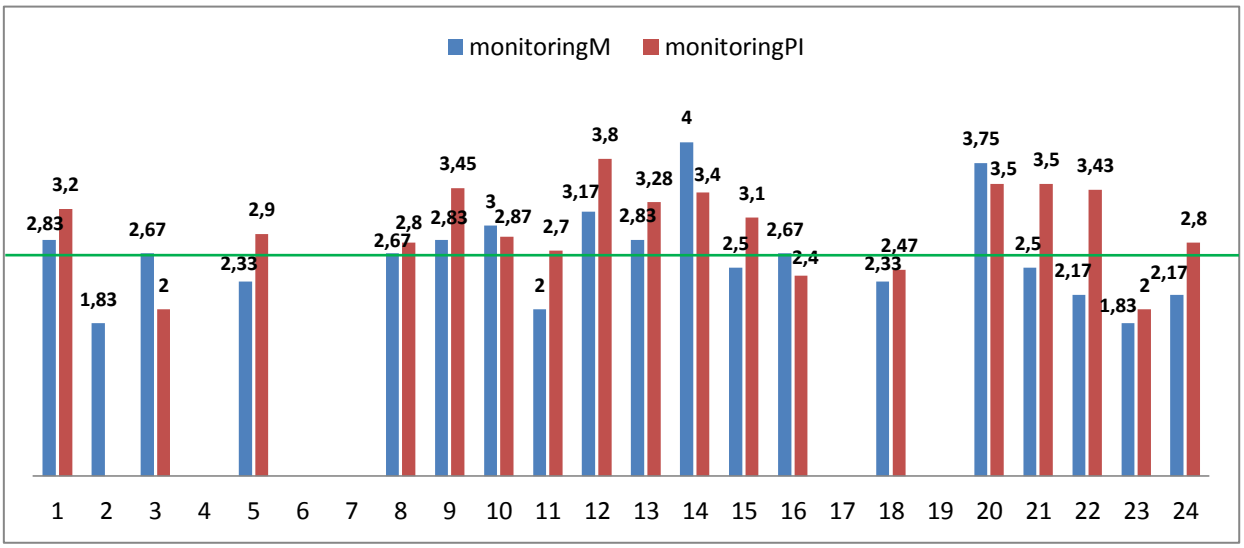


Figure 15.Level of monitoring reported by the program managers and program implementers for all programs in the sample with mid-intervention assessment.

5.1.2. Level of implementation quality seen by program implementers and program participants: findings

Group means and standard deviations per each indicator of implementation quality are shown in Table 5.3. Quality indicators include program fidelity, program quality, participant responsiveness, dosage, and perceived program impact.

Table 5.3

Descriptive statistics per indicators of implementation quality seen from the perspective of program implementers and program participants

	Fidelity		**Quality of delivery**		**Participants' responsiveness**		**Perceived program impact**		**Dosage**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Program implementers	3.11	0.97	3.68	0.35	3.55	0.25	3.46	0.37	-	-
Program participants			3.65	0.26	3.46	0.52	3.20	0.59	4.17	2.87
Correlations		**-**		**.480***		**.474***		**.314**		**-**

**** correlation is significant at the level p<.05***

Dosage was assessed with only one question per program participant. Participants were asked to report about the number of workshops/meetings held until the moment of the first implementation assessment. According to that, the dosage score does not follow the Likert scale from 1 to 4 but represents the average number of sessions held. Data for fidelity was collected only from program implementers.

As is shown in Table 5.3, program implementers give the lowest scores to fidelity (M=3.11, SD=0.97) which is followed by perceived program impact (M=3.46, SD=0.37), participants' responsiveness (M=3.55, SD=0.25) and quality of program delivery which has highest results (M=3.68, SD=0.35). The evaluations by program participants are similar: the lowest score is given to perceived program impact (M=3.20, SD=0.59), followed by participants' responsiveness (M=3.46, SD=0.52). The quality of program delivery is highest valued among indicators of implementation quality seen by participants (M=3.65, SD=0.26). In general, we may conclude that on all quality indicators scales, the average implementation quality was positively evaluated, both by program implementers and participants. The average number of sessions/meetings with participants until the first implementation quality assessment was M=4.17 (SD=2.87).

As can be seen from Table 5.3, correlations between the views of the two types of informants, program implementers and program participants are positive, although small. The lowest correlation is found for perceived program impact (r_s=.314). A significant positive correlation between the reports of participants and implementers ratings is found for responsiveness (r_s=.474, p<.05) and quality of program delivery (r_s=.480, p<.05) which tells us that program implementers and participants tend to agree in their views on the reactions of participants.

Since Table 5.4 shows the overall group results for fidelity, quality of program delivery, participants' responsiveness, perceived program impact, and dosage, in Table 5.4 the individual results per each of the 19 programs are shown. Unfortunately, program implementers did not send their reports for the *Mentor programme* (program number two) so their results cannot be compared with reports of program participants. Levels of assessed indicators of implementation quality from program implementers' and participants' perspective are also presented in Figure 16and Figures from 17 to 21. As can be seen from Table 5.4 and Figure 16, most of the programs have really homogenized indicators of implementation quality, no matter who the source of data is. Looking at individual indicators, there is a very small number of programs which have values below 3.

Table 5.4

Descriptive statistics per indicators of implementation quality for all 24 programs included in the sample.

Prog	Fidelity PI	Quality of delivery PI	Quality of delivery PART	Responsiveness PI	Responsiveness PART	Program impact PI	Program impact PART	Dosage (f)
1	3	3.67	3.74	3.75	3.66	4	3.34	14
2			3.88		3.66		3.58	7
3	3	3.25	3.8	3.63	3.61	3.13	3.29	1
4								
5	3	3.56	3.74	3.54	3.39	3.22	2.93	1
6								
7								
8	3.5	3.83	3.83	3.75	3.66	4	3.52	5
9	2.5	3.06	3.4	2.73	3.15	3	3.27	3
10	3.2	3.55	3.81	3.44	3.7	3.25	3.3	4
11	3.13	3.92	3.84	3.56	3.63	3	3.42	3
12	4	3.67	3.77	3.88	3.66	3.5	3.3	3
13	3.15	3.3	3.33	3.3	3.14	3.4	3	
14	4	4	3.8	3.5	3.65	4	3.62	5
15	3.25	3.25	3.6	3.5	3.2	3.5	3.12	2
16	2.81	3.63	3.69	3.44	3.57	3	3.32	4
17								
18	3.08	3.5	3.68	3.17	3.51	3.08	3.16	3
19								
20	2.75	3.83	3.6	3.29	3.49	3	2.84	6
21	3.63	3.92	3.62	3.75	3.25	3.83	3.11	4
22	3.29	3.64	3.84	3.71	3.72	3	3.26	3
23	2.25	3.83	3.64	3.88	3.51	4	3.28	3
24	2.5	3	3.78	3	3.64	3	3.24	3

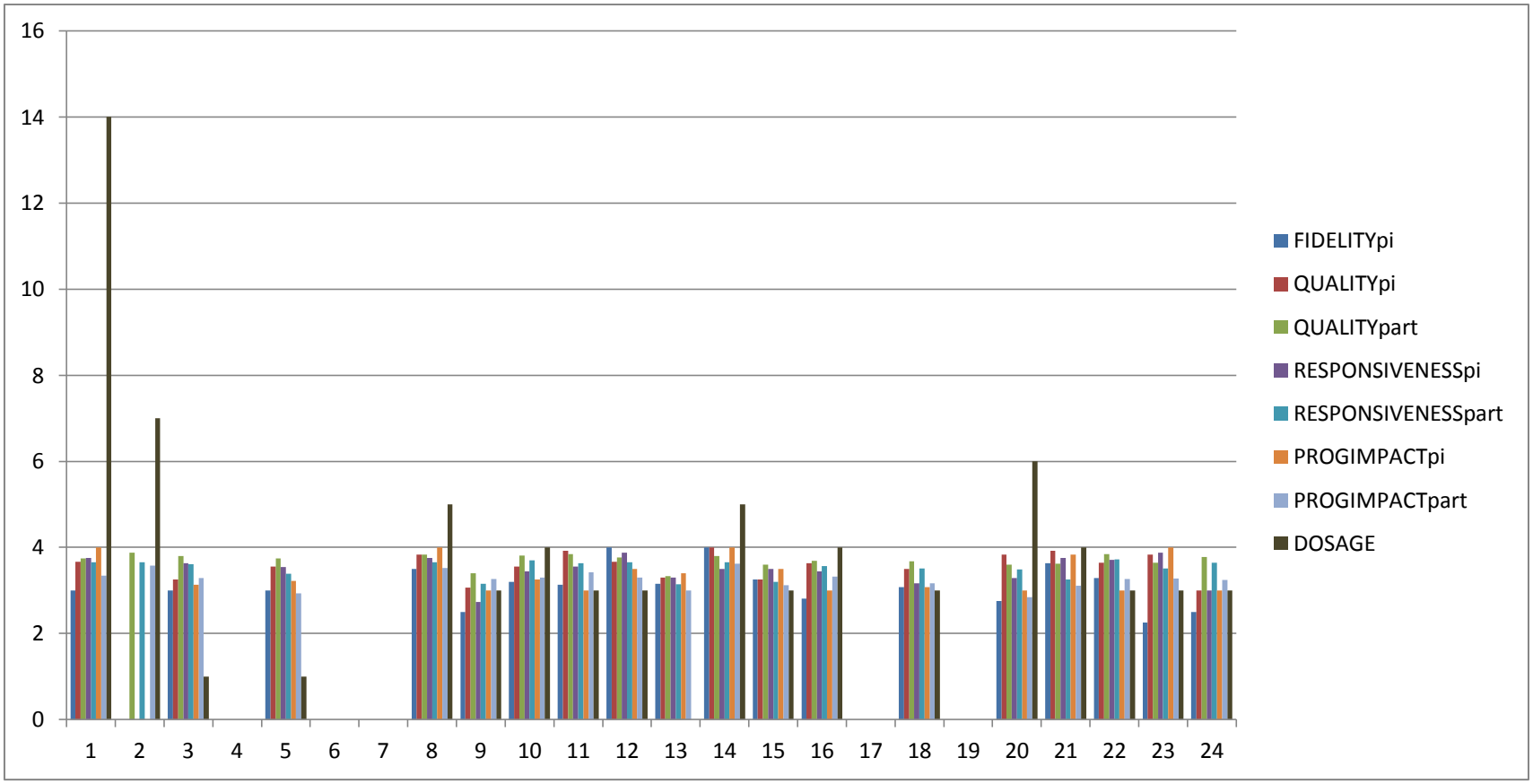


Figure 16.Comparison of reports of program implementers and program participants on indicators of implementation quality for all programs in the sample with mid-intervention assessment.

Note: all indicators of implementation quality besides dosage range from 1 to 4 while dosage is expressed in the number of sessions at mid-implementation assessment.

Fidelity

Table 5.4 and Figure 17 show that results for fidelity range from 2.25 to a maximum of 4. Lowest results for fidelity are found for the program twenty three (*Parenting programme VII*) while programs nine (*Parenting programme III*) and twenty four (*Parenting programme VIII*) both have results 2.5 (see Figure 17). If we take result of 2.5 as a cut-off point where everything below that value would count for low fidelity, it could be said that those three programs have low fidelity, meaning that program changes and tailoring program to participants needs is in those programs more dominant than conducting the program as it was planned. Maximum results for fidelity from the program implementer's perspective are seen for programs twelve (*Creative free time programme I*) and fourteen (*Parenting programme V*) which both have fidelity of 4 while program twenty one (*MH promotion through dance*) has a result of 3.63.

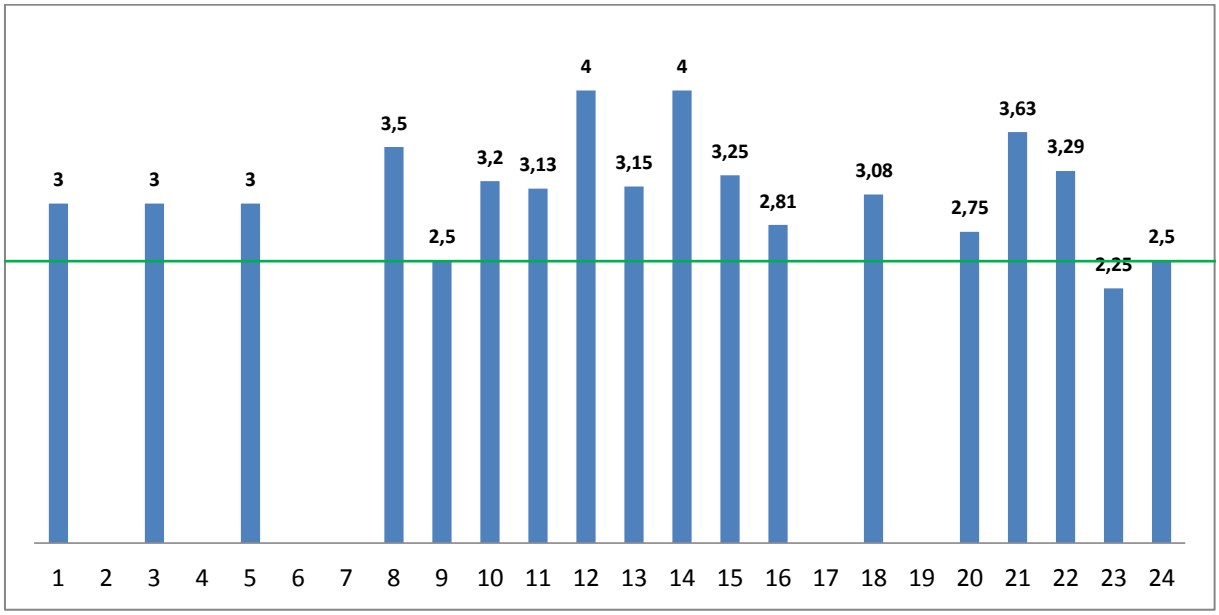


Figure 17. Level of fidelity seen from program implementers' perspective for all programs in the sample with mid-intervention assessment.

Quality of program delivery

Table 5.4 and Figure 18 show results per quality of program delivery both from the program implementers' and program participants' perspective. Several programs have a result around 3, which is the lowest reported result for quality of program delivery (program number nine, *Parenting programme III*, thirteen, *Free time for children in foster care*, and twenty

four, *Parenting programme VIII*). Highest results for quality of program delivery from implementer's standpoint are found for programs eleven (*Substance abuse prevention in the community*) and twenty one (*MH promotion through dance*) which both have a result of 3.92 while programs eight (*Parenting programme II*)and twenty three (*Parenting programme VII*) have a result of 3.83.

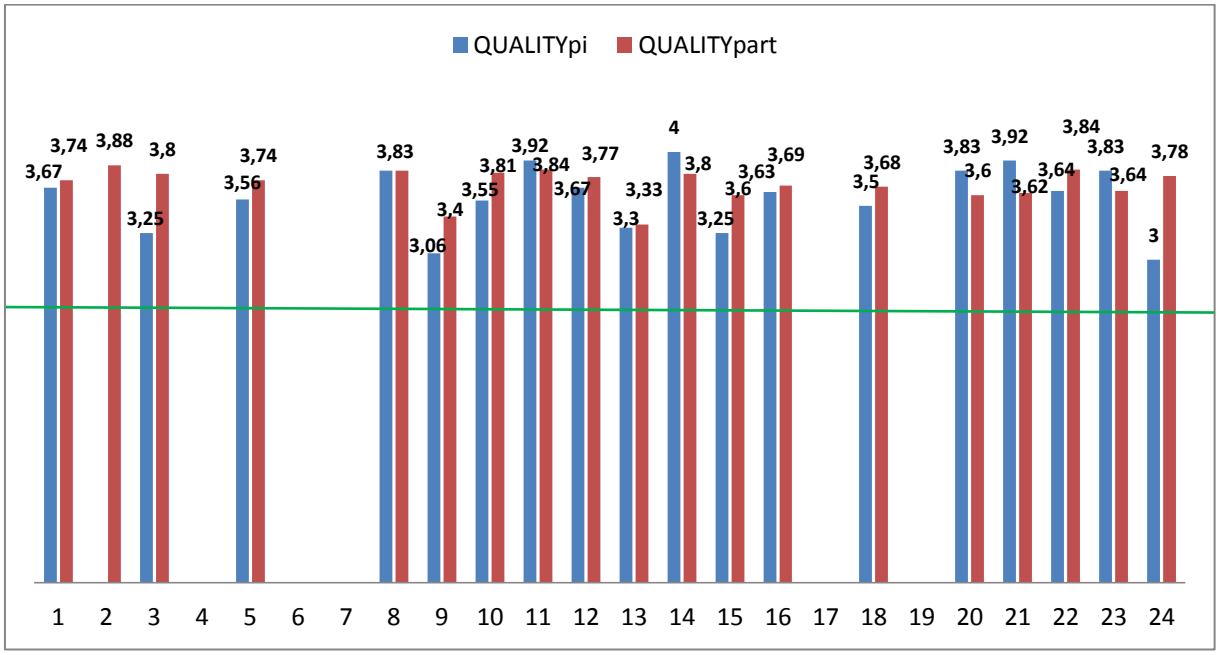


Figure 18. Level of quality of program delivery seen from program implementers' and participants' perspective for all programs in the sample with mid-intervention assessment.

Results for the quality of program delivery from participants' perspective are overall really high for all of the 19 assessed programs in mid-intervention. It seems that participants have given high reports for all indicators of implementation quality, especially for quality of program delivery and participants' responsiveness (see Table 5.4, Figure 18 and Figure 19). Average results for the quality of program delivery from participants' perspective are all above 3.40 (for program nine, *Parenting programme III*) while several programs have results above 3.80 (for example program two, *Mentor programme;* eight, *Parenting programme II;* eleven, *Substance abuse prevention in the community;* fourteen, *Parenting programme V;* twenty two, *Creative free time programme II*). Variability of reports of program participants on the quality of program delivery is quite small. Overall, participants seem to perceive the quality of program delivery as really good, they report on a high level of skill of program implementer, they like the style of implementer's work and the methods the implementer uses

in program delivery. Regarding the concordance in views of participants and implementers on quality of delivery, correlation of r_s=.480, p<.05 was found which indicates that participants and implementers have similar views on quality of delivery.

Participants' responsiveness

Results for the level of responsiveness from implementers' perspective range from 2.73 for program nine (*Parenting programme III*) to 3.88 for programs twelve (*Creative free time programme I*) and twenty three (*Parenting programme VII*). Participants report on their high responsiveness: from the participants' perspective, several programs have a result of 3.66 (program one, *MH promotion through the theatre;* program two, *Mentor programme;* eight, *Parenting programme II*; twelve, *Creative free time programme I*). Overall, it seems that both program implementers and program participants mostly have similar views on responsiveness (r_s=.474, p<.05) but participants give a bit higher reports (see Figure 19). Implementers see participants as engaged, interested during activities, active and report about good atmosphere. The same is also true for program participants.

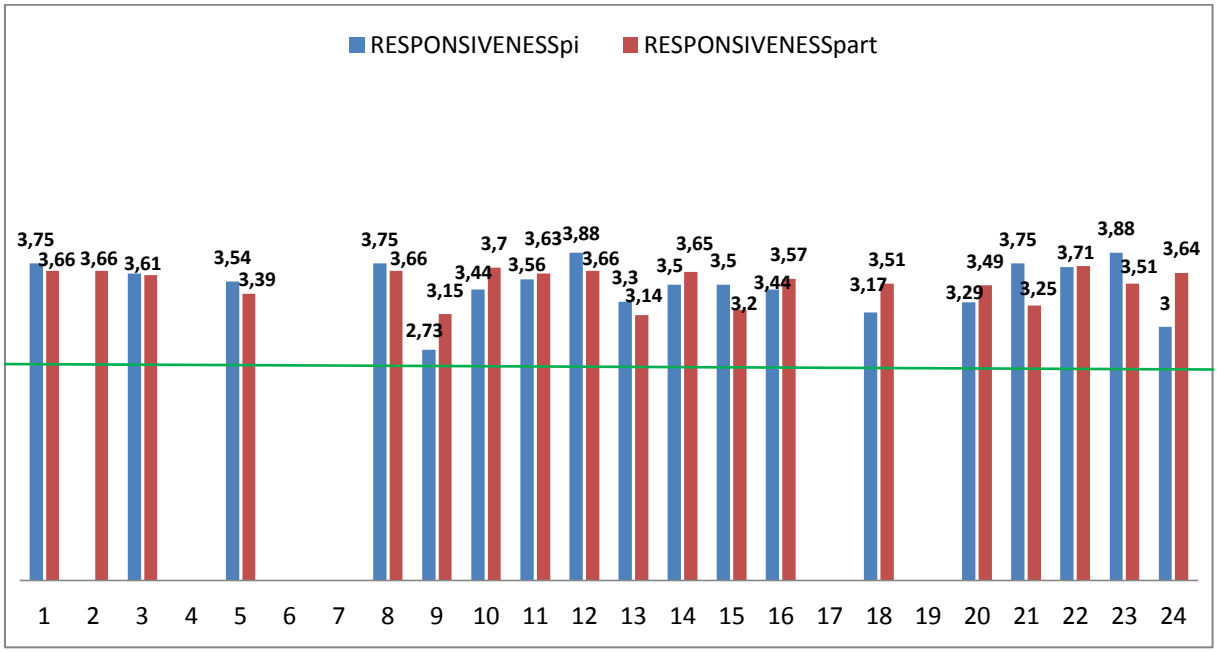


Figure 19. Level of responsiveness seen by program implementers and program participants for all programs in the sample with mid-intervention assessment.

The fact that both from program implementers' and program participants' perspective there is not a result below the reference point of 2.5 represents a high level of delivery quality. Only in few cases two types of informants report on responsiveness differently, for example

in the case of program twenty four (*Parenting programme VIII*) where implementers were stricter than participants. Those discrepancies probably depend on the kind of the program, and on the way participants were included: whether participation was voluntary, forced or whether the program was obligatory for them.

Perceived program impact

Results for program impact from program implementers' perspective have relatively small range from the minimum of 3 for several programs (nine, eleven, twenty, twenty two, twenty four) to the maximum of 4 for several programs (one, eight, fourteen, seventeen and twenty three). It seems that implementers are quite subjective and that they overestimate the impact of their intervention.

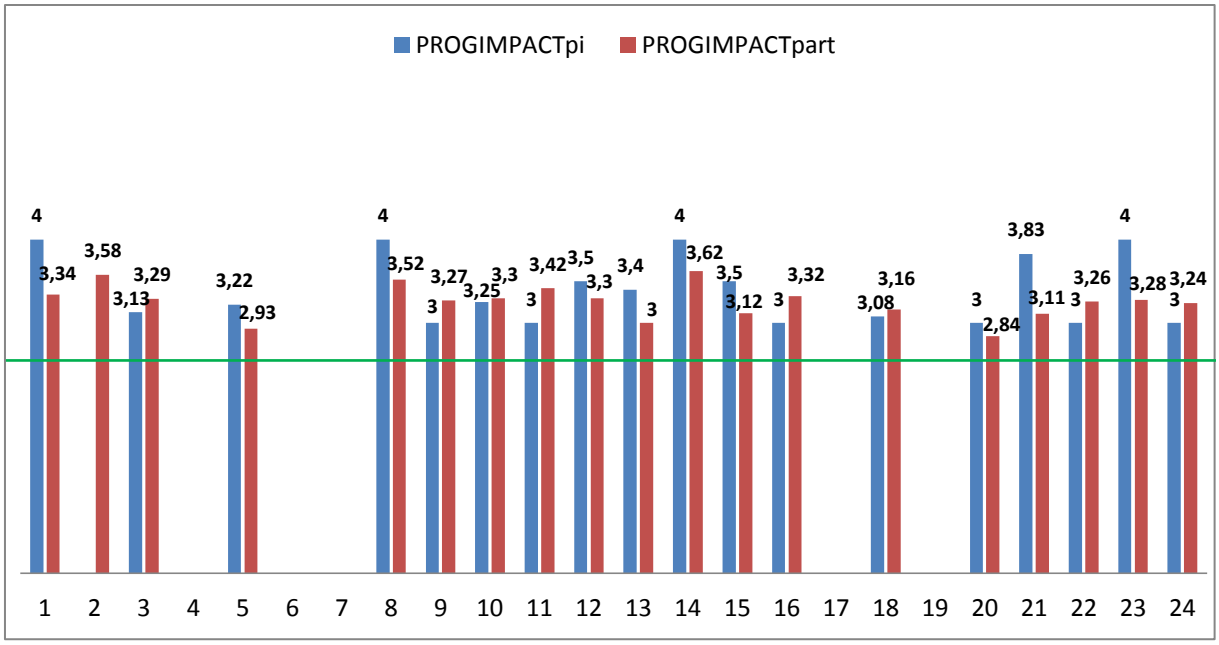


Figure 20. Level of program impact reported by program implementers and program participants for all programs in the sample with mid-intervention assessment.

Participants mainly report that they perceive programs as having a moderate impact on their lives. There are several programs where participants see really high effects of the intervention like program fourteen (*Parenting programme V*), program two (*Mentor programme*, M=3.58) and program eight (*Parenting programme II*, M=3.52). It seems that both program implementers and program participants in average give high reports about program impact. Analysing Figure 20, it seems that there are more programs that have lower

participants' ratings on the perceived program what also confirms the size of correlation which is relatively small. Participants from almost half of the programs in the sample report on lower program impact than program implementers do. Correlation of r_s =.314 between the evaluations of program implementers and program participants also represents the discrepancy in two perspectives.

Dosage

Results gathered for dosage presented in Table 5.4 and Figure 21 present differences in the number of sessions/meetings with participants between programs. It is important to stress that most programs had measured the quality of implementation at one third or one half of program delivery, so this level of dosage represents only the situation at first assessment. Since programs in the sample are different and cover various themes, that was somewhat expected. Figure 21 shows that program one (*MH promotion through the theatre*) had 14 sessions with participants at the first measurement of implementation quality and convincingly has highest number of sessions with participants, i.e. highest dosage.

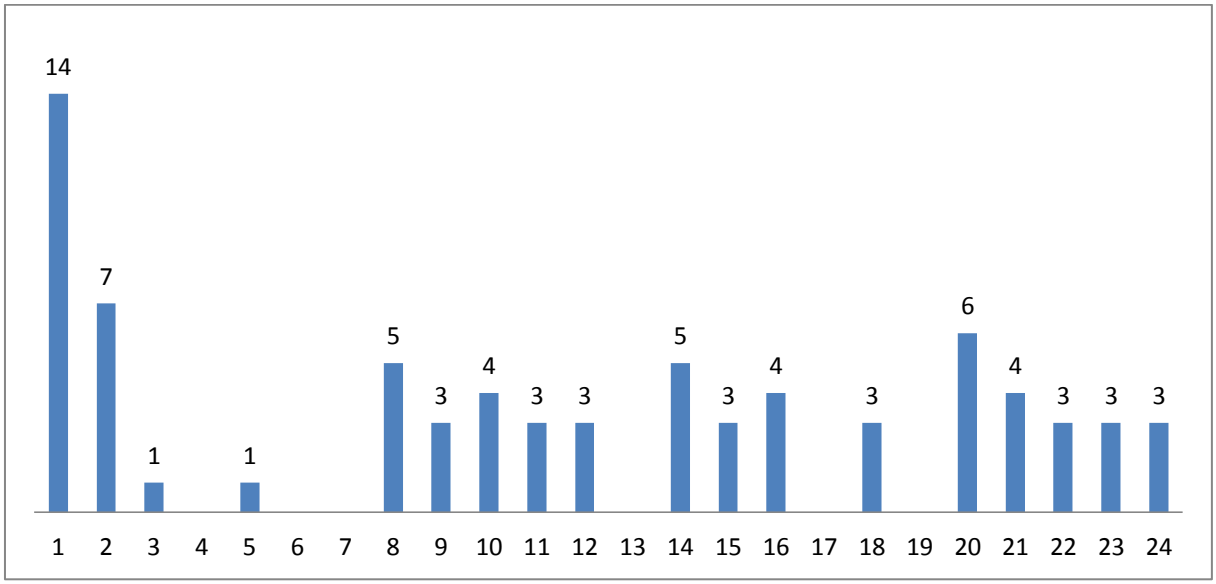


Figure 21. Dosage measured in the number of workshops/sessions held with participants in time of first implementation assessment.

Programs number three (*Parenting programme I*) and five (*Training for the group leaders*) had only one session with participants in the time of first implementation assessment. At that first implementation assessment, majority of programs had three meetings with participants while some had four or five.

5.2. Relationship of implementation factors and indicators of implementation quality: findings

In order to respond to the research task number four, i.e. *to test the relationship between implementation factors and indicators of implementation quality*, analyses have included correlation analysis and hierarchical linear modelling. Correlation analyses were conducted to see the association of individual variables while hierarchical linear modelling employed PROC MIXED procedure suitable for nested data and testing the hypothesis if output variable is predicted with independent variables. Since the conceptual model of implementation factors and indicators of implementation quality (see Figure 8 in Introduction section) used in this dissertation is assuming that capacity of structure for implementation and indicators of implementation quality are interrelated, the association of those two concepts was tested.

5.2.1. Correlation of implementation factors and indicators of implementation quality: findings

Correlations between implementation factors and indicators of implementation quality were calculated for different informants: managers and implementers, managers and participants as well as implementers and participants. It has to be stated that the significance of correlations depends on the size of the sample, meaning that correlations among managers and implementers will have different power regarding the size of sample. Firstly, capacity for implementation structure represented through implementation factors from the perspective of program managers' was analysed in regard to the indicators of implementation quality seen from program implementers' perspective. As is shown in Table 5.5, it was found that program implementers' skill and support are in consistent positive association with all indicators of implementation quality collected from implementers. Interestingly, program standardization and monitoring system are partly positively and partly negatively associated with indicators of implementation quality. Managers' report on the level of training is negatively correlated with all implementers' indicators of implementation quality.

Correlation of program standardization reported by program manager and responsiveness reported by program implementer is r_s=-.479, $p<.01$, which is the only significant although negative correlation. It seems that higher report from program manager

about the level of program standardization is followed with lower report on participants' responsiveness from the perspective of program implementer. Highest positive correlation of managers' ratings has been found for feedback on support to implementers and implementers ratings on fidelity (r_s=.417). Some of the moderate correlations have not been found significant because of the sample size, but since we explore these associations for the first time, it is interesting to see that managers' attitudes towards intervention are positively associated with implementers' ratings of perceived program impact (r_s=.347). Also, managers report on program implementers' skill is positively associated both with fidelity (r_s=.358) and quality of delivery (r_s=.327). It seems that managers views on support for program implementer are also associated with the perceived program impact collected from implementers' perspective (r_s=.325).

Table 5.5

Correlations of implementation factors seen from the perspective of program manager with indicators of implementation quality seen from the program implementers' perspective.

Managers / Implementers	Program standardization	Program implementers' skill	Attitudes towards intervention	Training and intervention knowledge	Support for program implementer	Monitoring system
Fidelity	.139	.358	.061	-.350	.417	.290
Quality	-.138	.327	.290	-.111	.286	.101
Responsiveness	-.479*	.100	.105	-.089	.130	-.163
Perceived program impact	-.270	.265	.347	-.022	.325	-.212

**** correlation is significant at the level p<.05, ** correlation is significant at the level p<.01***

Table 5.6 shows the correlations of program manager's reports on implementation factors and indicators of implementation quality seen from the perspective of program participants. It has to be stressed that significance of correlations is for those analyses visible already at lower levels of correlation because of the bigger sample size. It was found that correlation coefficients for participants' reports on dosage are only negatively associated with program standardization, while for all other managers' implementation factors and dosage positive association was found. The highest one is the correlation between dosage and

attitudes towards the intervention (r_s=.523, p<.001), managers' support for program implementer represents a moderate positive correlation (r_s=.357, p<.001). Correlations of managers' factors with participants' ratings of quality of delivery and perceived program impact are close to zero and negligibly small i.e. show pretty low associations with all implementation factors reported by managers. Responsiveness is in consistently positive association with all implementers' ratings of implementation factors, highest being for support for program implementer (r_s=.245, p<.001). Generally, there are a lot of correlation coefficients close to zero which indicates that reports from program managers and report of program participants are not associated.

Table 5.6

Correlations of implementation factors seen from the perspective of program manager with indicators of implementation quality seen from the program participants' perspective.

Managers / Participants	**Program standardization**	**Program implementer's skill**	**Attitudes towards intervention**	**Training and intervention knowledge**	**Support for program implementer**	**Monitoring system**
Dosage	-.152**	.351**	.523**	.171**	.357**	.214**
Quality	.021	.063	.001	-.072	.130**	.020
Responsiveness	.131**	.073	.034	.038	.245**	.104*
Perceived program impact	-.015	-.065	.114	-.051	.037	.004

**** correlation is significant at the level p<.05, ** correlation is significant at the level p<.01***

The last correlation analyses are those for implementers' ratings of implementation factors and the indicators of implementation quality seen by program participants (see Table 5.7). It was found that consistently all correlation coefficients are either low or close to zero, apart from for the dosage which shows some moderate association with implementers' ratings of implementation factors (for example, dosage and attitudes, r_s =.365, p<.001 and dosage and support, r_s =.377, p<.001). Results indicate that there is no association or that the association is small; or they are found to be negative what implies that the higher the level of implementation factors reported by implementers, the lower the reports of participants on the indicators of implementation quality.

Table 5.7

Correlations of implementation factors seen from the perspective of program implementers with indicators of implementation quality seen from the program participants' perspective.

Implementers / Participants	Program standardization	Attitudes towards intervention	Training and intervention knowledge	Support for program implementer	Monitoring system
Dosage	-.336**	.365**	.262	.377**	.171**
Quality	-.048	.116*	-.110*	.019	-.123**
Responsiveness	-.039	.104*	-.203**	.018	-.176**
Perceived program impact	-.069	.016	-.099*	-.104*	-.141**

***** correlation is significant at the level p<.01 * correlation is significant at the level p<.05***

5.2.2. Hierarchical linear models of association between implementation factors and indicators of implementation quality: findings

Our aim was to test if implementation factors reported by managers and implementers could serve as predictor variables which could estimate the criterion variables – indicators of implementation quality. Since we had two sources of implementation factors, managers and implementers, and two sources of implementation quality, implementers and participants, models will be presented per each criterion variable. The first stage of HLM analyses was to firstly check which of the implementation factors predict which indicators of implementation quality. Secondly, the idea was to include all significant predictors in one general model to make conclusions about the associations.

Managers' report on implementation factors predicting implementers' ratings on implementation quality

The multilevel models predicting implementers' indicators of implementation quality are summarized in Table 5.8. All conducted models show that only implementers' report on fidelity is predicted by implementation factors reported by program managers, i.e. managers'

report on implementers' skills (B=0.43, p<.04) and managers' report on support (B=0.35, p<.05) predict fidelity reported by implementers.

Table 5.8
Implementation quality reported by implementers predicted by the implementation factors collected from managers

	Standardization	Implementers' skills	Attitudes	Training	Support	Monitoring
Fidelity						
B	0.11	**0.43**	0.12	-0.12	**0.35**	0.34
SE	0.15	**0.19**	0.24	0.14	**0.16**	0.19
p	0.72	**0.04***	0.62	0.39	**0.05***	0.09
Quality of delivery						
B	-0.08	0.15	0.15	-0.03	0.11	0.03
SE	0.10	0.14	0.16	0.10	0.12	0.15
p	0.42	0.30	0.36	0.74	0.37	0.86
Participant responsiveness						
B	-0.16	0.09	0.07	0.01	0.13	-0.11
SE	0.10	0.16	0.17	0.17	0.13	0.15
p	0.15	0.55	0.69	0.93	0.34	0.48
Perceived program impact						
B	-0.17	0.17	0.30	-0.02	0.20	0.15
SE	0.13	0.18	0.18	0.12	0.15	0.18
p	0.18	0.33	0.13	0.88	0.21	0.40

**** correlation is significant at the level p<.05***

Managers' report on implementation factors predicting participants' ratings on implementation quality

The multilevel models predicting participants' indicators of implementation quality are summarized in Table 5.9. Series of multilevel models were estimated in which participants' ratings of indicators of implementation quality served as criterion variables. Results show that managers' implementation factors emerged as a statistically significant predictor for only two participants' indicators of implementation quality: dosage (predicted by managers' attitudes B=-0.42, p<.04) and participants' responsiveness (predicted by managers' support, B=0.17, p<.05).

Table 5.9
Implementation quality reported by participants predicted by the implementation factors collected from managers

	Standardization	Implementers skills	Attitudes	Training	Support	Monitoring
Dosage						
B	-0.04	-0.26	**-0.42**	-0.23	-0.17	-0.29
SE	0.25	0.21	**0.19**	0.26	0.23	0.21
p	0.87	0.24	**0.04***	0.38	0.47	0.18
Quality of delivery						
B	-0.03	0.08	0.01	-0.02	0.06	-0.01
SE	0.04	0.06	0.07	0.04	0.06	0.06
p	0.52	0.21	0.84	0.62	0.27	0.94
Participant responsiveness						
B	0.01	0.10	0.08	0.01	**0.17**	0.02
SE	0.07	0.10	0.11	0.06	**0.08**	0.09
p	0.94	0.32	0.50	0.86	**0.05***	0.87
Perceived program impact						
B	-0.02	-0.03	-0.11	-0.08	0.04	-0.04
SE	0.07	0.10	0.10	0.06	0.09	0.09
p	0.71	0.79	0.30	0.16	0.65	0.67

**** correlation is significant at the level p<.05***

Implementers' report on implementation factors predicting participants' ratings on implementation quality

The multilevel models predicting participants' indicators of implementation quality with reports of program implementers are summarized in Table 5.10. Series of multilevel models were estimated in which participants' ratings of indicators of implementation quality served as criterion variables. Results show that implementers' implementation factors emerged as a statistically marginally significant predictor for only one participants' indicator of implementation quality: dosage is predicted by implementers' report on attitudes (B=-0.42, p<.07) in a way that higher report of implementer on attitudes, smaller number of sessions in the program.

Table 5.10

Implementation quality reported by participants predicted by the implementation factors collected from implementers

	Standardization	Attitudes	Training	Support	Monitoring
Dosage					
B	-0.05	**-0.42**	-0.39	-0.28	-0.23
SE	0.20	**0.22**	0.31	0.19	0.21
p	0.81	**0.07***	0.22	0.17	0.31
Quality of delivery					
B	-0.03	0.15	-0.02	0.01	-0.05
SE	0.07	0.11	0.07	0.07	0.07
p	0.64	0.19	0.67	0.93	0.43
Participants' responsiveness					
B	-0.08	0.24	-0.09	0.01	-0.08
SE	0.10	0.17	0.10	0.10	0.10
p	0.44	0.19	0.40	0.96	0.43
Perceived program impact					
B	-0.13	0.09	-0.11	-0.12	-0.11
SE	0.09	0.17	0.09	0.10	0.09
p	0.18	0.60	0.24	0.22	0.25

**** correlation is close to the significance level of p<.05***

5.3. Discussion of results describing the level of implementation quality

This chapter aimed to explore the level of implementation quality in preventive programs in Istria, differences in the perception of implementation quality by three types of informants as well as relationship of implementation factors and indicators of implementation quality. Since this dissertation represents the first research on the implementation quality in Croatia, hypothesis for research questions were not established beforehand but were set as explorative research questions. In order to respond to them, results from the mid assessment of implementation quality were used, mainly being at one third of program delivery. Firstly, it is important to stress that implementation quality in this dissertation is defined as a result of implementation factors which describe the capacity of structure for implementation and indicators of implementation quality which are both known and researched in the tradition of implementation science (see Conceptual model, Figure 8). Information on implementation factors was gathered from program managers and program implementers who are familiar with the organizational capacity and program context. Information on indicators on implementation quality was collected from program implementers and program participants since they both witnessed program delivery.

In order to answer the research task two: ***To explore the level and variation of implementation quality in preventive programs in Istria***, descriptive statistics were shown to describe the level of implementation factors and indicators for implementation quality. Since different informants gave their reports on the level of implementation factors (managers and implementers) and indicators of implementation quality (implementers and participants), results are presented for all informants answering the research task three: ***To explore the differences in perception of implementation quality within program managers, program implementers and program participants***. Results could range from one to four for all researched dimensions besides dosage. In order to describe the level and variation in implementation quality in preventive programs in Istria, it was decided that the result of 2.5 would serve as a reference point. Results below that value for both implementation factors and indicators of implementation quality are considered as low.

Views on implementation factors

Results for program managers' perspective have revealed that managers feel that they provide high level of support for program implementer, assuring technical conditions and

emotional support (M=3.24, SD=0.51), and that they perceive implementers' as skilful and well prepared (M=3.42, SD=0.48). From the program managers' point of view, it could be stated that program standardization, organization of training and investments in intervention knowledge are low and they do not represent satisfactory conditions for quality implementation process. Average value of manager's reports for monitoring system are also really close to the cut-off point of 2.5 so it can be concluded that managers are not monitoring program delivery and don't communicate with implementer regularly.

Average results for implementation factors collected from program implementers show a bit different and more positive perception than it was the case for program managers. Program implementers also report that program standardization is the lowest of all implementation factors (M=2.87, SD=0.58) but for them average results go beyond the cut-off point. Program implementers report similarly on monitoring system (M=2.98, SD=0.60), training and intervention knowledge (M=2.88, SD=0.64) as well as for attitudes towards the intervention (M=2.93, SD=0.38). The highest implementation factor from the perspective of program implementer is support they receive from the manager and their organization (M=3.41, SD=0.51).

If we analyse programs individually per each implementation factor and compare the perception of program manager and program implementers, it could be stated that program implementers tend to report on higher **standardization** than program managers. Higher results for standardization from program implementers standpoint is found in the case of programs three (*Parenting programme 1),* ten (*Parenting programme IV),* fourteen (*Parenting programme V)* and sixteen (*Self-confidence training).* Low but positive correlation r_s=.228 for the two standpoints on standardization also shows that there is a discrepancy in their reports. Individual results per program **implementer's skills** are mostly showing a high level of implementer's skills in every program, with little variability in results. Only program twenty three (*Parenting programme VII*) with the result of 2.33 indicates that program manager feels that implementer should be more skilful and prepared. Individual results per **attitudes towards the intervention** for every program show generally high levels both for program managers and program implementers with high level of concordance in perceptions (r_s=.597, p<.001). Only managers from programs three (*Parenting programme 1*) and sixteen (*Self-confidence training*) report relatively low attitudes towards the intervention, while implementers of program sixteen also state low level of their attitudes. That would indicate that managers of programs three, six and sixteen (where implementers also think the same) do

not appreciate their intervention so much and that they feel that there are some improvements that have to be made in intervention.

Individual results per programs concerning **training and intervention knowledge** show big differences between programs in the sample. Those differences in this implementation factor are probably connected with differences in the status of organization, financial and human resources and capacities, but nevertheless indicate an area for investment. Program implementers from 10 out of 19 programs in the sample have reported about the level of the training and intervention knowledge below 2.5 which clearly indicates that this should be changed in the future from organizational but also from the county and policy level. Although more than half of program managers report that they do not organize in-service trainings or invest in the education of their implementers, only one program, program sixteen (*Self-confidence training*) has received the report below the reference point. Although managers and implementers answered the same questions, it seems that program implementers gave more positive answers regarding their level of training. Correlation of r_s=.358, p<.05 between reports of program managers and implementers about the training and intervention knowledge also shows that their views on the same matter are different and somewhat direct to the conclusion that implementers wanted to show themselves in a more positive manner.

Individual results per **program support for program implementer** and **monitoring** for each individual program show really high concordance among perceptions of managers and implementers. Correlation of results from managers' and implementers' perspective for support (r_s =.524**) and monitoring (r_s=.472**) also back up that conclusion. Lowest found support reported by program managers is found for program five (*Training for the group leaders*) while from the program implementers' perspective, lowest perceived support is received in program three (*Parenting programme 1*). Low monitoring from program managers' standpoint is found for programs two (*Mentor programme*) and twenty three (*Parenting programme VII*) while implementers report that the lowest monitoring is found in program three (*Parenting programme 1*).

Views on implementation quality

Regarding the indicators of implementation quality, program implementers rated fidelity, quality of delivery, participants' responsiveness and perceived program impact while participants reported about quality of delivery, participants' responsiveness, perceived

program impact and dosage. Average overall results per indicators of implementation quality including fidelity, quality, perceived participants responsiveness and program impact are above the value of 3, both seen from program implementer's and program participant's perspective. While the average **fidelity** result which was assessed only from the program implementers' self-report for all programs in the sample is 3.11, SD=0.97, there is one program whose individual result was below the reference point of 2.5 and two programs that had the result 2.5. Accordingly, we could conclude that programs twenty three (*Parenting programme VII*), nine (*Parenting programme III*) and twenty four (*Parenting programme VIII*) have low fidelity.

Average group results for **quality of program delivery** show that both program implementers (M=68, SD=0.35) and participants (M=3.65, SD=0..26) generally report on high quality of delivery that would indicate that the skill with which lessons were delivered, integration of concepts into program activities and deliverer's responses to participants are high quality. Moderate correlation of results from implementers' and participants' perspective for quality (r_s=.480, p<.05) also advocates that conclusion. Looking from implementers perspective, several programs have result around 3 which is lowest results for quality of program delivery (program number nine, *Parenting programme III*, thirteen, *Free time for children in foster care*, and twenty four, *Parenting programme VIII*). Results for the quality of program delivery from participants' perspective are overall really high for all of the 19 assessed programs in mid-intervention. As it was stated in the result section, average results for the quality of program delivery from participants' perspective are all above 3.40 meaning that variability of participants' reports is quite small.

Group results for **participants' responsiveness** show that both informants rate the response of participants in a similar manner (r_s=.474, p<.05), indicating that programs in our cohort stimulate the interest of participants who are active, engaged and attentive. Overall, it seems that implementers see participants as engaged, interested during activities, active and report about good atmosphere. The same is also true for program participants.

Group results for **perceived program impact** are again high for group level, both from the perspective of program implementers (M=3.46, SD=0.37) and program participants (M=3.20, SD=0.59). Although reports are high, association of two sources of data about impact shows that there is a discrepancy in their views (r_s=.314). Analysing the individual results, it seems that there are more programs that have lower participants' ratings on the

perceived program impact i.e. participants from almost half of the programs in the sample report on lower program impact than program implementers do.

Results gathered for **dosage** presented present number of sessions at one third or one half of program delivery (M=4.17, SD=2.87). Since programs in the sample are different and cover various themes, number of lessons varies. Highest dosage was found for program one (*MH promotion through the theatre*) which had 14 sessions with participants while programs three (*Parenting programme 1*)and five (*Training for the group leaders)* had only one lesson. Generally, going back to mental health promotion and prevention literature and recommendations for the number of meetings with participants (Bartholomew et al., 2006; Derzon et al., 2005; Hawkins et al., 2002; O'Connel et al., (Eds), 2009) it can be stated that programs in our cohort have low dosage.

On the general level, average results per indicators of implementation quality would support the conclusion that all indicators of implementation quality are high. At the same time, low variability in results could also mean that both program implementers and program participants gave overly positive feedback. Generally, when analysing the results, it can be concluded that program managers have given the lowest reports; program implementers gave a bit higher reports than program managers while program participants gave really high reports with low variability. It seems that those findings for implementation factors are more realistic while relatively low variability in the values of indicators of implementation quality could indicate socially desirable answers.

Relations between implementation factors and implementation quality

Research task number four:***To explore the relationships of implementation factors and indicators of implementation quality***was responded with correlation and hierarchical linear modelling analyses. Results of the association between implementation factors have revealed that implementation factors and indicators of implementation quality are not as inter-related as expected. Even though there was not a hypothesis about the level of relationship, literature indicates that implementation factors predict the level of implementation quality (Kam et al., 2003; Roberts-Gray et al., 2007; Mihalic et al., 2008). At the same time, it has to be stated that our association could be due to methodological limitations and sample size. For the **association of managers' ratings of implementation factors and implementers' ratings of implementation quality**, HLM shows significant associations only for implementers' report on implementation fidelity with managers' report on implementers'

skills (B=0.43, p<.04) and with managers' report on support (B=0.35, p<.05). **Regarding the association of managers' implementation factors and participants' indicators of implementation quality**, results show that managers' ratings emerged as a statistically significant predictor for two participants' indicators of implementation quality: dosage (negatively predicted by managers' attitudes B=-0.42, p<.04) and participants' responsiveness (predicted by managers' support, B=0.17, p<.05). **When inspecting relationship of implementers' implementation factors and participants' indicator of implementation quality**, statistically marginally significant predictor for only one participants' indicator of implementation quality – dosage, was implementers' report on attitudes (B=-0.42, p<.07).

It could be stated that our design doesn't allow us to draw conclusions on the nature of relationship between implementation factors and indicators of implementation quality. It seems that in this study, role and standpoint of the manager is a better predictor of implementation outcomes as seen by participants in comparison to implementers view on implementation factors as predictors. These indications should be further researched in the future since our study did not show outcomes that are in accordance with results of previous studies, especially when it comes to contextual factors and leadership (Riley et al., 2001; Gingiss et al., 2006; Dariotis et al., 2008; Wandersman et al., 2009). **There is a possibility that the link between implementation factors and implementation quality is indirect, mediated with some other variables which were not included in this study or cannot be confirmed because of the small study power and number of participants.**

From a practical standpoint, the value of these findings for mental health promotion and prevention practice in the County of Istria can be found in some recommendations that can be drawn from the results on implementation factors. It seems that overall, programs three (*Parenting programme 1)* and sixteen (*Self-confidence training*) which are both from the same organization, need serious investments in the capacity of structure for implementation. Overall, all organizations would benefit from investments in the level of standardization of program. That result was somewhat expected, while based on the experiences of researcher with these studied practices, most programs included in the whole sample do not have clear guidelines for program delivery, sometimes there is not even an exact set of themes that has to be covered, and sometimes changes are made in programs without a sufficient scientific base. Also, according to results on implementation factors, the level of training and intervention knowledge is another important area for future investments. Both the organizations conducting interventions and the administration of the Department of Health and Social

Services in the County of Istria should invest more in the training of implementers, organize more in-service trainings and assure the translation of up-to-date skills and knowledge necessary for program delivery. That is really important because program implementers often have different professional backgrounds, level of experience and skills. For some of the implementers, delivering interventions in the field of mental health promotion and prevention is not a permanent position but part-time job which is not controlled and standardized, which can affect the quality of implementation in general. Since those organizations and implementers are stakeholders which carry out county's policy, the County of Istria should also organize trainings which would contribute to the level of knowledge and professional readiness of those people because they are working with children, youth and families.

These recommendations have to be taken as a direction since our research does not offer strong evidence for the predictive values of implementation quality measures. From our point, lack of evidence of association of implementation factors and indicators of implementation quality is connected with the sample size which reported on implementation factors. Also, there is potential influence of other contextual or program factors which are confounded with association and were not researched in this study. Since presented data stands for only nineteen managers and fifty implementers, we would perceive these directions as valuable contribution. With these study limitations in mind, recommendations could still be valuable to the County of Istria.

6. CHAPTER SIX: IMPACT OF THE TRAINING FOR PREVENTION ON THE IMPLEMENTATION QUALITY

This chapter is targeted at the research goal connected with the study on the impact of the Training for Prevention on implementation factors and level of implementation quality. In order to fulfil the fifth research task which is connected with this goal, this chapter will test the following hypotheses:

- ***Hypothesis 5.1:*** **After the Training for Prevention, program managers and implementers from the experimental group will report improved implementation factors in comparison with the control group.**
- ***Hypothesis 5.2:*** **After the Training for Prevention, program implementers and program participants in the experimental group will report improved implementation quality in comparison with the control group.**

In previous chapters, the Training for Prevention was described as a training designed to develop knowledge, skills and capacities of organization managers and program implementers in the experimental group. The general aim of the Training for Prevention was to affect a variety of factors connected with the support system and capacities of the organization and involved professionals, with the assumption that this in turn, would improve the quality with which the programs were conducted by implementers. Given the goals of the Training for Prevention, the effect of the intervention was examined on measures of implementation factors as well as implementation quality. As shown in Table 6.1, information on implementation factors was collected from program managers and implementers while implementers and participants have reported on indicators of implementation quality. Managers and implementers are those that are familiar with characteristics of a program and with organizational capacity, while implementers and participants are those that can reflect implementation quality. This multi-measurement strategy was employed because different sources have different perceptions of the constructs being examined. As reported in chapter 4, all scales of implementation factors and implementation quality show good reliability, except for the scale on the implementers' view on fidelity. For this reason, results for fidelity are presented since this is the first implementation study in Croatia, but with strong limitations in mind.

As discussed in chapter 3, to measure the changes in implementation factors and implementation quality in both experimental and control cohorts of programs, measures of implementation factors and quality were sent to the organization managers, program deliverers and program participants to complete after approximately 1/3 or half of each program had been delivered and again at the very end of program delivery. The first intervention assessment is not a traditional pre-test measurement since program delivery had to be underway in order to assess implementation quality. Since the programs included in this study varied in length, the timing of the mid-intervention assessment was different for each program depending on its length.

Table 6.1
Sources of data for implementation factors and indicators of implementation quality

	MANAGERS	**IMPLEMENTERS**
IMPLEMENTATION FACTORS	Standardization Implementers' skills Attitudes Training Support Monitoring	Standardization Attitudes Training Support Monitoring
	IMPLEMENTERS	**PARTICIPANTS**
INDICATORS OF IMPLEMENTATION QUALITY	 Fidelity Quality of delivery Responsiveness Perceived program impact	Dosage Quality of delivery Responsiveness Perceived program impact

Plan of Analyses

Analysis Sample Description

As already explained in the methods chapter, the plan for the current study was to assess implementation factors and indicators of implementation quality for each program at two time points, after the first third of program delivery (mid-intervention) and after each program ended (post-intervention). There were five programs for which data from their managers, implementers and participants was not included in the mid-intervention analysis. Three programs like *Substance abuse prevention for parents*, *Substance abuse prevention for teachers,* and *Substance abuse prevention* only had one session and had data collection at the end of program delivery. Two programs, *Media literacy* (four sessions) and *Underage drinking prevention* (five sessions) could not organize the collection of data in mid-delivery.

Both of those two took place in schools, within regular school days and it was difficult to organize children to stay some extra time for the data collection.

Out of 24 programs, 19 managers gave their report during the first measurement time point. Five missing managers were from programs mentioned above. At post-intervention 22 of the total 24 managers had returned the questionnaires (missing managers at post-test are from organizations eighteen, *Parenting programme VI,* and twenty three, *Parenting programme VII)*. Talking about implementers, there was more than one implementer per some programs. Apart from the already explained five programs which did not have data collection at mid-intervention, for program two, *Mentor programme,* the implementer did not deliver its questionnaires for both first and second assessment. 51 implementers completed questionnaires at the mid-implementation assessment (20 from experimental conditions and 31 from control condition). In the post-assessment of implementation, 55 implementers filled in the questionnaires, 33 from experimental conditions and 22 from control conditions (for detailed explanation of sample see Table A4, Appendix 4, see page 225). Data on implementation factors and indicators of implementation quality gathered from managers and implementers matched two time points. Ratings of implementation quality were collected from 410 participants from 19 programs at the mid-assessment point. As was mentioned above, ratings were not collected twice in five programs which were short and lasted only one or just few sessions since program delivery had to be in progress before the first implementation assessment. At post-test intervention, number of participants is higher than at mid-intervention because of that fact i.e. data was collected from 744 participants at post-test. Unfortunately, participant ratings were not linked and paired from the first to the second time point because a range of participants forgot their codes when completing the forms.

Description of Statistical Approach

Since managers' data was at the program level (i.e., there was only one manager per program), it was analysed using multiple regression procedure with SAS software. Since implementers and participants in both experimental and control conditions were not independent but rather nested within programs, the impact of the Training for Prevention was analysed with hierarchical linear modelling (HLM) within SAS software. A random identification number was assigned to each program and this was used as a clustering variable. HLM is a form of analysis which is used to analyse data when participants are nested within units and thereby violate assumptions of independence that are required for traditional

multivariate approaches. HLM simultaneously investigates relationships within and between hierarchical levels of grouped data, thereby making it more efficient at accounting for variance among variables at different levels (Singer, 1998). HLM takes into account how similar implementers within the same program or participants within the same program are to one another. It estimates residual variance at the level of the individual implementer or participant and at the level of the program. It then adjusts the degrees of freedom on which the estimate of the intervention effect is based to account for the fact that it was specific programs, not individual implementers or participants, which were randomly assigned to the experimental or control conditions (Ferron et al., 2004; Woltman, Feldstein, McKay, Rocchi, 2012).

This chapter will present regression and HLM models on the impact of the Training for Prevention on implementation factors and indicators of implementation quality, as reported by managers, implementers and participants within each program. It is important to stress that all of the analyses were conducted on raw data. Results of HLM analyses are unstandardized parameter estimates which are coefficients which represent the amount of change in implementation variables due to Training. Differences between intervention and control groups are also presented as effect sizes, which represents those differences in a common metric as a proportion of the pooled standard deviation i.e. expressed in the measurement unit of a dependent variable by dividing it by the common standard deviation (Cohen, 1988; Cohen, 1992; Brace, Kemp and Snelgar, 2006). Given the small sample size in the current study, effect sizes are presented even when the mean differences between the groups were not statistically significant. According to Cohen (1988; 1992) effect sizes between .10 and .30 are considered small, between .30 and .50 are considered medium and effect sizes above .80 are considered large. Effect sizes were calculated so that positive numbers reflect differences that favour the intervention group while negative numbers favour the control group.

Results presented in this chapter for managers, implementers and participants as a source, reflect the impact of the Training for Prevention. In order to test the hypotheses 5.1 and 5.2, three immanent questions were answered with the analyses:

1) Is there an intervention effect of Training for Prevention at the end of the intervention period?

2) For those longer programs that had two implementation assessments, is there an intervention effect at mid-intervention?
3) For those longer programs that had two implementation assessments, did the effect of the Training emerge later in the process of program delivery (i.e., was there increasing differentiation between the intervention and control groups after mid-intervention).

For all three types of informants (managers, implementers, participants) and for both sets of outcomes (implementation factors and indicators of implementation quality), three blocks of analyses were conducted in order to test the hypotheses and answer the questions about the Trainings' impact. The findings from these sets of analyses will be presented in the same manner throughout the chapter: a) analysis of post-test data since at post-test we have gathered the maximum of subjects in the sample; b) analysis of mid-intervention assessment because of subset of programs at mid-intervention (smaller number of programs than at post-test) and c) for only that programs which were long enough to have two assessments, analysis of post-test data where data from the mid-intervention assessment was controlled. Because individual participants could not be tracked from the mid-intervention assessment to the post-intervention assessment, program averages of the participants' individual ratings, rather than the individual ratings themselves, had to be used for the mid-intervention control.

This third group of analyses are unlike most assessments which examine change from baseline to post-intervention. These analyses are only looking at what happens between mid-intervention and post-intervention. If all the effects of Training for Prevention are evident early on, there will be no residual treatment effects, after the mid-intervention assessment is controlled. However, if the effects of Training for Prevention take time to consolidate, there might still be residual treatment effects after the mid-intervention assessment is controlled.

To test whether the Training for Prevention intervention was effective across multiple different program types, moderator analyses also will be conducted, for implementation outcomes for managers, implementers and participants (Baron and Kenny, 1986). Tested moderator variables are length of a program, activity level of a manager, type of the program (mental health promotion program, parent program, substance abuse prevention program) and type of the participants (children, youth, adults). The findings from these analyses are presented at the end of this chapter.

6.1. Impact of Training for Prevention on Implementation Factors

In order to test the hypothesis5.1: *After the Training for Prevention, program managers and implementers from the experimental group will report improved implementation factors in comparison with the control group,* descriptive statistics and hierarchical models will be presented for all examined implementation factors, separately for managers and separately for implementers as sources of information.

6.1.1. Group Comparisons on Manager Ratings of Implementation Factors: Findings

Intervention Effects at Post-Test

Means and standard deviations for the managers' ratings on the post-test implementation assessment (N=22 managers from 22 programs) are presented separately by condition. Since managers' data was on the individual level, analyses employed a series of multiple regression models to test the impact of Training for Prevention for each factor rated by managers: standardization, report on implementer's skills, attitudes towards the program, training, support and monitoring. Table 6.2 shows mean differences in the managers' implementation factors reported at post-test. Results at post-test revealed only one positive effect size that favoured the intervention group which was for ratings of implementers' skills (d=.35). All other implementation factors had a negative intervention effect meaning that the control group of managers gave higher ratings compared to the intervention group. Multiple regression has showed that differences in post-test means were significant for standardization (B=-0.84, p<.001), training (B=-0.56, p<.05) and monitoring (B=-0.79, p<.01).

Table 6.2
Levels of implementation factors reported by managers and the effect of the Training for Prevention at post-intervention

	POST INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Standardization					
M	2.89	2.05	**-0.84**	**<.001****	-1.35
SD	0.43	0.81	**(0.28)**		
Implementers' skills					
M	3.61	3.76	0.15	0.42	0.35
SD	0.47	0.39	(0.18)		
Attitudes					
M	2.97	2.74	-0.23	0.21	-0.55
SD	0.37	0.46	(0.18)		

Training					
M	2.95	2.39	**-0.56**	**.05***	-0.92
SD	0.72	0.50	**(0.27)**		
Support					
M	3.47	3.25	-0.21	.36	-0.42
SD	0.63	0.40	(0.22)		
Monitoring					
M	3.14	2.35	**-0.79**	**.01****	-1.59
SD	0.70	0..29	**(0.22)**		
N	**11**	**11**			

Intervention Effects at Mid-Intervention

Table 6.3 shows mean differences in managers' implementation factors reported at mid-intervention for the 19 programs that collected data at this time point. Descriptive statistics at mid-intervention show that the control group of managers did view all implementation factors more positively than the intervention group of managers. Results of multiple regressions have shown that difference in mid-intervention means is significant for standardization (B=-0.83, p<.04) and monitoring (B=-0.54, p<.04). Found effect sizes at mid-intervention are negative and small, except for variables standardization and monitoring which had larger negative effect size differences between the groups.

Table 6.3
Levels of implementation factors reported by managers and the effect of the Training for Prevention at mid-intervention

	MID INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Standardization					
M	2.73	1.90	**-0.83**	**.04***	-1.52
SD	0.48	0.61	**(0.25)**		
Implementers' skills M	3.53	3.37	-0.16	.49	-0.32
SD	0.45	0.56	(0.23)		
Attitudes					
M	2.83	2.75	-0.07	.73	-0.17
SD	0.46	0.46	(0.21)		
Training					
M	2.50	2.36	-0.14	.71	-0.18
SD	0.96	0.57	(0.37)		

Support					
M	3.33	3.09	-0.24	.36	-0.43
SD	0.59	0.53	(0.26)		
Monitoring					
M	2.89	2.35	**-0.54**	**.04***	-1.08
SD	0.62	0.38	**(0.24)**		
N	**10**	**9**			

Intervention Effects after Mid-Evaluation

In order to determine if the effect of Training for Prevention emerged later in the implementation process, after the mid-intervention assessment point, mean levels of implementation factors reported by managers at post-test were compared controlling for ratings made at the mid-intervention point. Those analyses were done for managers which have had paired data in both time points, N=17 managers from seventeen programs. Table 6.4 shows the effect of Training when controlling for time one. Parameter estimates presented in Table 6.4 indicate that there are no significant differences present between control and experimental group of managers in either of managers' implementation factors. Effect sizes for implementation factors when mid-intervention assessment is controlled have shown small positive effects for implementers' skills (d=0.30) and support (d=0.09) while others are negative.

Table 6.4

Levels of implementation factors reported by managers at post-test and the effect of the Training for Prevention when mid-intervention mean differences are controlled

	CONTROL		**INTERVENTION**		**INTERVENTION EFFECT**		
	MID	**POST**	**MID**	**POST**	**BETA**	**P**	**EFFECT SIZE**
Standardization							
M	2.75	2.81	1.96	1.83	-0.33	.13	-0.34
SD	0.50	0.32	0.63	0.76	(0.20)		
Implementers' skills M	3.59	3.63	3.50	3.67	0.12	.22	0.30
SD	0.43	0.45	0.44	0.44	(0.09)		
Attitudes							
M	2.87	2.98	2.75	2.73	-0.15	.16	-0.27
SD	0.48	0.36	0.49	0.54	(0.10)		
Training							
M	2.47	2.84	2.31	2.47	-0.21	.35	-0.21
SD	1.01	0.71	0.59	0.56	(0.22)		

Support							
M	3.39	3.48	3.13	3.27	-0.01	.98	0.09
SD	0.59	0.64	0.55	0.47	(0.17)		
Monitoring							
M	2.95	3.19	2.42	2.39	-0.42	.10	-0.55
SD	0.63	0.72	0.35	0.27	(0.24)		
N	**9**	**9**	**8**	**8**			

6.1.2. Group Comparisons on Manager Ratings of Implementation Factors: Discussion

Findings of multiple regression and effect sizes calculation suggest that Training for Prevention effect was mostly negative. This pattern of findings was already evident at the mid-intervention assessment when managers' from control conditions gave higher self-report on all researched implementation variables (effect sizes ranged from .17 to 1.52). Multiple regression models at mid-intervention assessment have revealed statistically significant negative Training for Prevention effect for standardization (B=-0.83, p<.04) and monitoring (B=-0.54, p<.04). The same pattern of findings remained at post-test data with the exception of manager ratings of implementers' skills which were rated higher by managers of intervention programs compared to control (d=.35). Multiple regression at post-test showed significant negative effects for standardization (B=-0.84, p<.001), training (B=-0.56, p<.05) and monitoring (B=-0.79, p<.01). The effect scores based on the HLM analyses that compared the growth in manager ratings of implementer skills between mid-intervention and post-intervention time points were positive (d=.30) suggesting that the benefit of the intervention on this factor emerged across the second half of the intervention but did not become statistically significant until post-test.

Immediately at mid-intervention assessment, it was found that managers from control conditions give higher self-report on all researched implementation variables. The same pattern of findings remained at post-test data with the exception of manager ratings of implementers' skills. It is possible that these results are the result of insufficient knowledge and acknowledgment of science-based implementation principles. Since Training for Prevention took place before the first assessment for every program from intervention conditions, and educated all included to be critical towards their programs and project proposals, one interpretation of these findings is that participation in the Training for Prevention offered new information to managers which led them to become more sensitive

towards their implementation, develop higher standards and higher awareness of implementation quality. Managers from the intervention group may have become more self-critical of their programs and may have started to realize how their programs could be better in regard to implementation factors. It is possible that managers from control conditions perceived implementation factors level more unrealistically and that they do not know what is important and maybe they see themselves, their implementers and organizations in whole, in overly positive light.

Analysing change from mid-intervention to post-intervention revealed that the effects of Training for Prevention on standardization and monitoring were evident early on, at mid-intervention. Those effects were still evident at post-intervention; however, they did not show significant change between mid-intervention and post-intervention. In contrast, the pattern is somewhat different for the effects on Training. In this case, the effects of Training for Prevention took time to consolidate. They were not evident at mid-intervention, but they had grown large enough by post-intervention to be statistically significant. Interestingly, though, when the growth that occurred between mid-intervention and post-intervention was tested in isolation, it was not statistically significant. Thus, it appears that non-significant growth in two separate periods (prior to mid-intervention and from mid-intervention to post-intervention) was significant when combined.

Regarding these findings, **we can conclude that Training for Prevention did not have a significant positive impact on the experimental group of managers** i.e. **implementation factors reported by** managers were not higher in the experimental group. The training might have had the effect on managers to use more critical standards when evaluating the factors for implementation in their own organization.

6.1.3. Group Comparisons on Implementer Ratings of Implementation Factors: Findings

Intervention Effects at Post-Test

Table 6.5 shows means and standard deviations for implementers' ratings of implementation factors on the post-test implementation assessment (N=55 implementers from 22 programs), separately for the implementers in control and experimental conditions. Analyses employed a series of hierarchical linear models to test the impact of Training for Prevention for each factor rated by implementers: standardization, attitudes towards the program, training, support and monitoring. Although implementers from control conditions generally give higher self-reports on implementation factors (except for support) than implementers from experimental conditions, HLM has not confirmed those mean differences at implementers' post-assessment as significant. Found effect sizes are low to moderate negative for all implementers' implementation factors except for support which has a positive but really low effect size.

Table 6.5
Levels of implementation factors reported by implementers and the effect of the Training for Prevention at post-intervention

	POST INTERVENTION		INTERVENTION EFFECT		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Standardization					
M	3.10	2.85	-0.34	0.15	-0.43
SD	0.54	0.59	(0.23)		
Attitudes					
M	3.04	2.83	-0.13	0.31	-0.59
SD	0.41	0.29	(0.13)		
Training					
M	3.07	2.85	-0.22	0.21	-0.36
SD	0.66	0.59	(0.17)		
Support					
M	3.44	3.49	0.07	0.77	0.09
SD	0.47	0.69	(0.25)		
Monitoring					
M	2.97	2.80	-0.38	0.14	-0.29
SD	0.57	0.60	(0.25)		
N	**33**	**22**			

Intervention Effects at Mid-Evaluation

When going back to the subset of 18 programs whose implementers gave their reports on implementation factors at mid-assessment, Table 6.6 shows mean differences in implementers' reports on implementation factors at mid-intervention. Results at mid-intervention show that the control group of implementers did view all implementation factors more positively than the intervention group of implementers. HLM has shown that difference in mid-intervention means is significant for implementers' ratings of standardization (B=-0.42, p<.05). Found effect sizes for levels of implementation factors at mid-intervention are all negative, from low to moderate values.

Table 6.6
Levels of implementation factors reported by implementers and the effect of the Training for Prevention at mid-intervention

	MID INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Standardization					
M	3.04	2.62	**-0.42**	**0.05***	-0.76
SD	0.47	0.64	**(0.19)**		
Attitudes					
M	2.99	2.83	-0.09	0.54	-0.43
SD	0.40	0.33	(0.14)		
Training					
M	3.03	2.65	-0.38	0.11	-0.64
SD	0.66	0.53	(0.22)		
Support					
M	3.49	3.30	-0.25	0.30	-0.35
SD	0.49	0.59	(0.23)		
Monitoring					
M	3.06	2.85	-0.41	0.10	-0.34
SD	0.58	0.64	(.24)		
N	**31**	**20**			

Intervention Effects after Mid-Assessment

In order to isolate the time where the effect of Training for Prevention took place, mean levels of implementation factors reported by implementers at post-test were controlled for differences at mid-intervention. Those analyses were done for those implementers which have had paired data in both time points, N=45 from seventeen programs. Table 6.7 shows the

effect of Training on implementers' ratings of implementation factors when controlling for ratings at mid-intervention. Parameter estimates presented in Table 6.7 indicate that there are no significant differences between the control and experimental group of implementers for either of the implementation factors. Results off implementers' effect sizes for implementation factors when mid-intervention assessment is controlled have shown low effect sizes whereas standardization, attitudes and monitoring have a negative intervention effect while training and support have positive effect sizes.

Table 6.7
Levels of implementation factors reported by implementers and the effect of the Training for Prevention when time one differences are controlled

	CONTROL		INTERVENTION		INTERVENTION EFFECT		
	MID	**POST**	**MID**	**POST**	**BETA**	**P**	**EFFECT SIZE**
Standardization							
M	3.05	3.10	2.65	2.69	-0.10	.56	-0.02
SD	0.47	0.53	0.67	0.61	(0.18)		
Attitudes							
M	3.02	3.05	2.91	2.87	-0.09	.49	-0.19
SD	0.39	0.43	0.31	0.36	(0.12)		
Training							
M	3.06	3.06	2.80	2.97	0.04	.83	0.28
SD	0.66	0.69	0.51	0.57	(0.19)		
Support							
M	3.49	3.47	3.42	3.44	0.07	.68	0.07
SD	0.45	0.46	0.62	0.79	(0.17)		
Monitoring							
M	3.09	2.99	2.99	2.77	-0.16	.49	-0.19
SD	0.57	0.56	0.68	0.70	(0.22)		
N	**30**	**15**	**30**	**15**			

6.1.4. Group Comparisons on Program Implementer Ratings of Implementation Factors: Discussion

Analyses shown in previous section were conducted in order to test the impact of Training for Prevention on implementation factors rated by implementers' in experimental conditions. Hypothesis 5.1 stated that program implementers from the experimental group will report improved implementation factors after the Training for Prevention. Mid-

intervention HLM on data from 51 implementers in 18 programs has revealed that the control group of implementers did view all implementation factors more positively than the intervention group of implementers. Mid-intervention HLM has shown that the mean difference at that time point is significant for implementers' ratings of standardization, meaning that implementers from control conditions have rated standardization of their programs more positively than implementers from the intervention group ($B=-0.42$, $p<.05$). When inspecting data where reports are matched during mid and post-test assessment, 45 implementers from 17 programs were included in the HLM analysis when controlling for the mean differences for the mid-implementation. Hierarchical models shown in Table 6.7 did not reveal any effects of the Training for Prevention on implementers in experimental conditions, which suggests that there were no changes in the time from mid to post-test assessment. Interestingly, calculated effect sizes for training and support are small but positive in the period from mid to post assessment.

Since 55 implementers from 22 programs have returned filled in questionnaires at post-test, two programs from experimental group were missing (program two, *Mentor programme*, and program sixteen, *Self-confidence training*). Although implementers from control conditions generally give higher self-reports on implementation factors than implementers from experimental conditions (except for support), post-test HLM analyses have not confirmed those mean differences as significant. At post-test, found effect sizes are low to moderate negative for all implementers' implementation factors except support which has a positive but really low effect size. It seems that this negative intervention effect on standardization is present at mid-intervention and then loses its significance at post-test assessment while positive effects for support and training take time to consolidate from mid to post-assessment and are almost lost at post-test.

Regarding these findings, **it can be concluded that the Training for Prevention did not have a significant impact on the implementation factors reported by the implementers**. Part of the hypothesis 5.1 related to implementation factors should be rejected: when compared with the control group, implementers from the experimental group are not reporting on improved implementation factors. Immediately at mid assessment of implementation factors, the control group of implementers has given higher self-reports than implementers from experimental conditions. These findings are again similar to the findings at managers' level: it might be that people from programs which were in experimental conditions and may have given more realistic reports, might have been more attuned to

important issues concerning the science-base of mental health promotion and prevention and might have viewed their programs more critically, as the result of becoming aware of important implementation issues through Training received before the mid-assessment. Even though Training for Prevention did not cause significant effects on the implementers' factors from experimental conditions, it might have caused awareness and changed perspective, which could be perceived as positive outcome, despite the lower ratings of implementation quality.

6.2. Impact of Training for Prevention on Implementation Quality

As it was already explained at the beginning of this chapter, implementation quality was assessed through the ratings of program implementers and program participants. Program implementers reported about fidelity, quality of delivery, responsiveness and perceived program impact while participants were asked about dosage, quality of delivery, responsiveness and perceived program impact. In order to test hypothesis5.2: *After the Training for Prevention, program implementers and program participants in the experimental group will report improved indicators of implementation quality in comparison with the control group*, descriptive statistics will be presented for all indicators of implementation quality followed with HLM models.

6.2.1. Group Comparisons for Implementers' Ratings of Implementation Quality: Findings

Intervention Effects at Post-Test

Means and standard deviations for indicators of implementation quality collected from 55 implementers from 22 programs at post-intervention are presented separately for implementers in the control and experimental conditions (see Table 6.8). Descriptive statistics in Table 6.8 shows that at mid-assessment point, implementers from control conditions generally give a bit higher self-reports than implementers from intervention conditions on all indicators of implementation quality except for responsiveness which has higher reports in the intervention group (only positive effect size d=.34). HLM at post-intervention data has not revealed statistical differences in reports of control and experimental group of implementers for either indicator of implementation quality.

Table 6.8
Levels of implementer's ratings on indicators of implementation quality and the effect of the Training for Prevention at post-intervention

	POST INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Fidelity					
M	3.29	3.22	-0.02	0.85	-0.19
SD	0.37	0.36	(0.13)		

Quality					
M	3.55	3.49	-0.03	0.82	-0.16
SD	0.40	0.36	(0.14)		
Responsiveness					
M	3.36	3.49	0.12	0.45	0.34
SD	0.41	0.36	(0.15)		
Perceived program impact					
M	3.34	3.25	0.02	0.93	-0.19
SD	0.48	0.48	(0.19)		
N	**33**	**22**			

Intervention Effects at Mid-Assessment

Going back to those implementers who gave their reports at mid-implementation assessment (i.e., to the subset of programs which were long enough to have two assessments), HLM was conducted to check if there is an effect of Training present at that point. Table 6.9 shows the mean differences of indicators of implementation quality reported by implementers. It can be seen that HLM did not find significant differences for any of the indicators; all parameter estimates are not significant. Calculated effect sizes show a positive effect size for quality and responsiveness at mid-intervention implementation assessment: effect size for quality is rather small (d=.18) while the one for responsiveness (d=.67) could be seen as moderate. Analyses have also shown small and negative effect size for fidelity and perceived program impact.

Table 6.9

Levels of implementer's ratings on indicators of implementation quality and the effect of the Training for Prevention at mid-intervention

	MID INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Fidelity					
M	3.18	3.02	-0.30	0.15	-0.36
SD	0.37	0.43	(0.19)		
Quality					
M	3.51	3.57	0.03	0.84	0.18
SD	0.39	0.29	(0.14)		

Responsiveness					
M	3.37	3.60	0.22	0.14	0.67
SD	0.41	0.28	(0.14)		
Perceived program impact					
M	3.29	3.19	-0.02	0.93	-0.24
SD	0.39	0.44	(0.17)		
N	**31**	**20**			

Intervention Effects Mid-Assessment

When we wanted to answer the question what has happened between mid and post-implementation assessment for those 17 programs which had two assessments, hierarchical linear models were calculated per each of the indicator of implementation quality with the control of mid-assessment differences. Table 6.10 shows that there are no significant differences within control and experimental conditions for either of the indicators of implementation quality. Effect sizes indicate that there has been a small negative intervention effect on quality and responsiveness while fidelity and program impact have been affected positively in that period from mid to post-implementation assessment.

Table 6.10
Levels of implementer's ratings on indicators of implementation quality and the effect of the Training for Prevention at post-intervention when mid-intervention differences are controlled

	CONTROL		**INTERVENTION**		**INTERVENTION EFFECT**		
	MID	**POST**	**MID**	**POST**	**BETA (SE)**	**P**	**EFFECT SIZE**
Fidelity							
M	3.18	3.29	3.07	3.27	0.02	.89	0.23
SD	0.38	0.37	0.45	0.35	(0.16)		
Quality							
M	3.53	3.58	3.56	3.53	-0.01	.95	-0.21
SD	0.38	0.40	0.33	0.39	(0.13)		
Responsiveness							
M	3.37	3.38	3.63	3.56	0.02	.85	-0.22
SD	0.41	0.42	0.29	0.32	(0.12)		

Perceived program impact							
M	3.31	3.36	3.26	3.37	0.10	.53	0.13
SD	0.39	0.47	0.46	0.53	(0.15)		
N	**30**	**15**	**30**	**15**			

6.2.2. Group Comparisons on Program Implementer Ratings of Indicators of Implementation Quality: Discussion

Analyses shown in the previous section were conducted in order to test the impact of Training for Prevention on implementers' ratings of indicators of implementation quality in experimental conditions. Hypothesis 5.2 stated that program implementers from the experimental group will report improved indicators of implementation quality after the Training for Prevention.

For the mid-assessment, 51 implementers from 18 programs gave their reports. When those data were analysed, HLM did not show any significant Training for Prevention effects. At the same time, effect sizes show a positive intervention effect for quality of delivery and responsiveness while fidelity and program impact have small negative effects. When analysing the effects of the Training from mid to post-test assessment, multilevel models shown in Table 6.10did not reveal any effects of the Training for Prevention, but it seems that in this period intervention has had positive effects on fidelity and perceived program impact and small negative effects on quality of delivery and participants' responsiveness. Post-test HLM analyses did not reveal any statistically significant differences between intervention and control conditions although it has to be stated that effect sizes have shown a positive effect for fidelity (d=.34). Interestingly, that effect was present immediately at mid-assessment, has diminished from mid to post-test assessment and is found at post-test. When looking at effects sizes, it seems that effects of the Training were taking different turns on various indicators of implementation quality.

Regarding these findings, respecting the significance indicators, **it has to be concluded that Training for Prevention did not have a significant impact on the level of implementation quality reported by implementers, although there are interesting patterns shown within the effect sizes. Therefore, the first part of hypothesis 5.2 which is connected with implementers has to be rejected. When compared with the control**

group, implementers from the experimental group are not reporting on improved indicators of implementation quality.

6.2.3. Group Comparisons for Program Participant Ratings of Implementation Quality: Findings

Intervention Effects at Post-Test

Changes in implementation indicators reported by participants were examined using the hierarchical linear modelling. Hierarchical linear models were used with the purpose to analyse if Training for Prevention had an effect on dosage, quality of delivery, participants' responsiveness and perceived program impact. Means and standard deviations for indicators of implementation quality collected from 744 participants from 22 programs at post-intervention are presented separately for participants in the control and experimental conditions (see Table 6.11). It has to be stated that dosage is reported in percentage implementation of the total planned program. It is obvious that all average results for indicators of implementation quality, for dosage, quality of delivery, participants' responsiveness ad perceived program impact reported by participants in the experimental group are higher than in the control group of participants. HLM at post-intervention data for participants' ratings has revealed statistical differences in reports of control and experimental group of participants for quality of delivery (B=0.17, p<.02) and responsiveness (B=0.26, p<.04). All effect sizes for participants' indicators of implementation quality are positive and moderate.

Table 6.11
Levels of participant's ratings on indicators of implementation quality and the effect of the Training for Prevention at post-intervention

	POST INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA**	**P**	**EFFECT SIZE**
Dosage					
M	92.33	99.45	6.96	.13	0.71
SD	18.00	2.13	(4.41)		
Quality					
M	3.43	3.64	**0.17**	**.02***	0.49
SD	0.52	0.34	**(0.07)**		

Responsiveness					
M	3.09	3.50	**0.26**	**.04***	0.76
SD	0.67	0.41	**(0.12)**		
Perceived program impact M	2.94	3.23	0.19	.14	0.43
SD	0.74	0.62	(0.12)		
N	**391**	**353**			

Intervention Effects at Mid-Assessment

Table 6.12 shows the results of N=434 participants (316 from control conditions and 118 from experimental conditions) from seventeen programs at mid-intervention. As seen from parameter estimates and p-values, mean differences between participants' ratings of dosage, quality, responsiveness and perceived program impact at mid-assessment are not statistically significant. Also, it could be stated that at mid-intervention, dosage is lower in the experimental group of programs while mean differences between experimental and control group found for responsiveness are marginally significant (B=0.18, p<.09).

Table 6.12

Levels of participant's ratings on indicators of implementation quality and the effect of the Training for Prevention at mid-intervention

	MID INTERVENTION		**INTERVENTION EFFECT**		
	CONT	**INT**	**BETA (SE)**	**P**	**EFFECT SIZE**
Dosage					
M	38.53	36.54	-3.07	0.56	-0.22
SD	8.34	10.14	(5.10)		
Quality					
M	3.65	3.77	0.10	0.15	0.35
SD	0.37	0.31	(0.07)		
Responsiveness					
M	3.39	3.63	0.18	0.09	0.52
SD	0.56	0.37	(0.10)		
Perceived program impact M	3.16	3.32	0.13	0.21	0.27
SD	0.63	0.54	(0.10)		
N	**316**	**118**			

Intervention Effects after Mid-Assessment

In order to narrow down the period of time where the effect of Training for Prevention took place, mean levels of indicators of implementation quality reported by participants at post-test were analysed, controlling for differences at mid-intervention assessment. Those analyses were done for participants who have had data collection in both time points, participants from eighteen programs. Because participants were not matched through the two assessments, per each of the participants' ratings of implementation quality average program results from time one were used. Table 6.13 shows the effect of Training on participants' ratings of implementation quality when controlling for mid-intervention.

Table 6.13

Levels of participant's ratings on indicators of implementation quality and the effect of the Training for Prevention at post-test assessment with the control of mid-intervention data

	CONTROL		**INTERVENTION**		**INTERVENTION EFFECT**		
	MID	**POST**	**MID**	**POST**	**BETA (SE)**	**P**	**EFFECT SIZE**
Dosage							
M	38.53	97.09	36.54	98.43	3.20	.16	0.46
SD	8.34	7.89	10.14	3.38	(2.17)		
Quality							
M	3.65	3.46	3.77	3.61	0.10	.31	0.08
SD	0.37	0.55	0.31	0.33	(0.10)		
Responsiveness							
M	3.39	3.20	3.63	3.41	0.03	.77	-0.06
SD	0.56	0.67	0.37	0.39	(0.11)		
Perceived program impact							
M	3.16	3.06	3.32	3.19	0.06	.68	-0.05
SD	0.63	0.76	0.54	0.55	(0.14)		
N	**308**	**296**	**112**	**123**			

Parameter estimates presented in Table 6.13 for the time from mid to post-test assessment indicate that there are no significant differences present between the control and experimental group of participants in either of the indicators of implementation quality.

Results of effect sizes for implementation quality at post-test with the control of mid-assessment differences show moderate intervention effect for dosage (d=0.46), really low to zero effect for quality (d=0.08) and low negative to zero effects for responsiveness (d=-0.06) and perceived program impact (d=-0.05).

6.2.4. Group Comparisons on Participant Ratings of Indicators of Implementation Quality: Discussion

As it was stated above, to test hypothesis 5.2 concerned with changes in the participants' ratings of indicators of implementation quality, hierarchical linear modelling was used. When mid-intervention assessment of programs which had two data collections was analysed, there are no significant differences present between the control and experimental group of participants in either of the indicators of implementation quality. Although differences were not significant, the intervention group of participants has rated all indicators of implementation quality except dosage higher than the control group of participants; while mean differences found for responsiveness are marginally significant. When only the period from mid to post-assessment was analysed, no effects of the Training were found. At the post-test of full-sample, hierarchical linear modelling has found effects of Training for Prevention on two out of four indicators of implementation quality. To be more specific, Training for Prevention significantly affected participants' reports of quality of delivery (B=0.17, p=.02) and participants' responsiveness (B= 0.26, p=.04).

Regarding the effect-sizes, at mid-intervention assessment, found effects were small to medium while effect size for dosage was negative. At post-test, effect sizes for all indicators of implementation quality were medium to high indicating that it took time for them to consolidate. Going back to hypothesis 5.2, **it can be stated that part of the hypothesis connected with changes in indicators of implementation quality on the level of participants is confirmed: at post-test participants from experimental conditions report on higher indicators of implementation quality than participants in control conditions**. It is important to stress that participants did not know if their managers or implementers participated in the Training for Prevention. As a result, the participant ratings are completely unbiased.

6.3. Moderators of the Training for Prevention impact: findings and discussion

Finally, to examine whether the effects of Training for Prevention were robust across different conditions, tests of moderation were conducted for all sources of data, for managers, implementers and participants (Baron and Kenny, 1986). In particular, the level of involvement of program managers, the length of programs, the type of participants, and the type of programs were assessed as potential moderators. **Activity of a manager** represents the actual and active involvement of managers in the program delivery. Since some managers were not even familiar with the program content, some did not participate in the decision making process or just gave their name and formal support, the activity of managers was coded zero for those managers that were not active and one for those that were really important for delivery. **Regarding the length of programs**, a dummy variable was made where zero value represented short programs which classified programs which had less than six sessions with participants. Length of program was included as a moderator since there is a great difference between programs in our cohort regarding the number of sessions. Those programs that were longer than six sessions were coded with value one. **Type of participants'** variable is connected with the fact whether program participants were children, teenagers or adults. This moderator was included since some literature shows that the age of the participants is also relevant to reactions of implementers or that different age groups react differently to the program which is delivered. **Type of the program** variable is referring to the general type of programs included in the whole study. According to the content and goals, all 24 programs were divided into three groups: mental health promotion programs, parental programs and programs preventing substance abuse. We wanted to check if a difference in program type plays a role in the Training for Prevention effectiveness.

Results of hierarchical models which have tested moderators' impact and interaction with the Training are shown in Table 6.14 below (showing only significant moderator models for clarity). As explained, moderator models were tested for managers, implementers and participants. Results have shown that all of the mentioned variables did not have any effect on the impact of the Training either for implementation factors reported by managers and implementers or for indicators of implementation quality reported by implementers. For participants' ratings of implementation quality, we have found that type of the program and type of the participants were not significant moderators of Training's effect. The set of presented models in Table 6.14 examined the direct relationships between the activity of a

manager and length of a program, Training for Prevention effect and the dependent variables of indicators of implementation quality reported by participants (dosage, participants' responsiveness, quality and perceived program impact).

Table 6.14
Models testing effects of activity of a manager and the length of a program with the Training for Prevention on the indicators of implementation quality reported by participants

DOSAGE	**Effects of Training for Prevention**	**Effects of moderator**	**Effects of interaction**
Activity of a manager	B=11.57, SE=6.00, p=.07	B=8.45, SE=6.25, p=.19	B=-9.86, SE=8.82, p=.28
Length of a program	**B=0.52, SE=0.14, p=.001****	**B=1.44, SE=0.96, p=.14**	**B=-0.46, SE=0.19, p=.02***
QUALITY	**Effects of Training for Prevention**	**Effects of moderator**	**Effects of interaction**
Activity of a manager	**B=0.23, SE=0.08, p<.01****	**B=0.46, SE=0.18, p=.02***	**B=-0.09, SE=0.12, p=.45**
Length of a program	**B=0.30, SE=0.09, p<.01****	**B=0.20, SE=0.08, p=.03***	**B=-0.18, SE=0.12, p=.15**
PARTICIPANTS' RESPONSIVENESS	**Effects of Training for Prevention**	**Effects of moderator**	**Effects of interaction**
Activity of a manager	**B=0.39, SE=0.16, p=.03***	**B=0.32, SE=0.16, p=.06**	**B=-0.22, SE=0.23, p=.53**
Length of a program	**B=0.68, SE=0.15, p<.001*****	**B=0.56, SE=0.14, p<.001*****	**B=-0.60, SE=0.20, p<.001****
PERCEIVED PROGRAM IMPACT	**Effects of Training for Prevention**	**Effects of moderator**	**Effects of interaction**
Activity of a manager	B=0.28, SE=0.17, p=.11	B=0.26, SE=0.17, p=.06	B=-0.15, SE=0.24, p=.53
Length of a program	**B=0.53, SE=0.14, p<.001****	**B=0.57, SE=0.13, p<.001*****	**B=-0.46, SE=0.19, p=.03***

Significant moderator models were found for the **length of program and its combinations with all indicators of implementation quality** - dosage, quality of program delivery, participants' responsiveness and program impact. Interaction of length of program with Training for Prevention's impact is significant and presented in Figures 22, 23, 24 and 25 below.

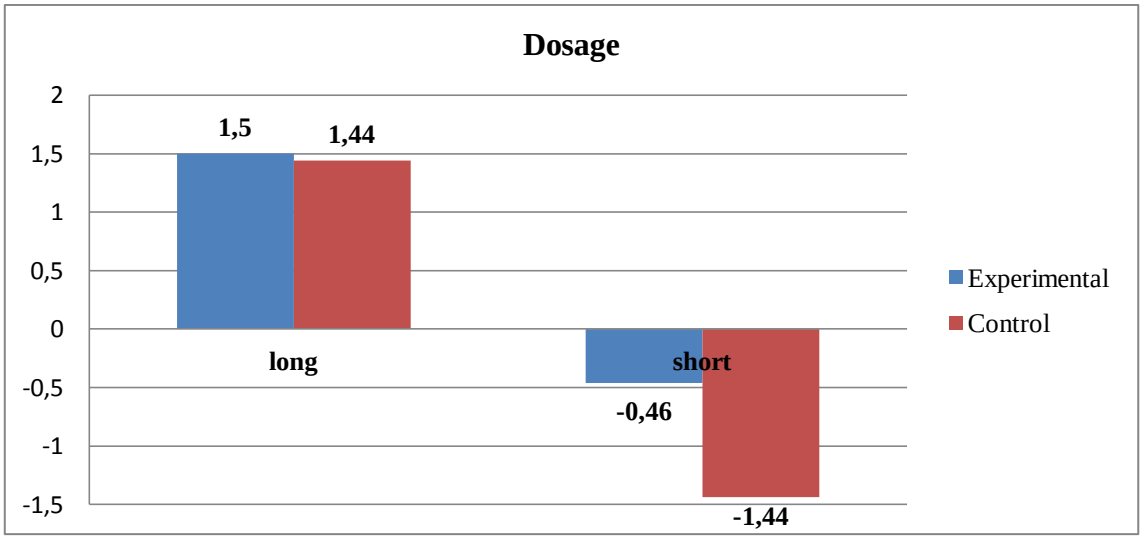


Figure 22. Effects of Training for Prevention and length of the program on the participants report on dosage.

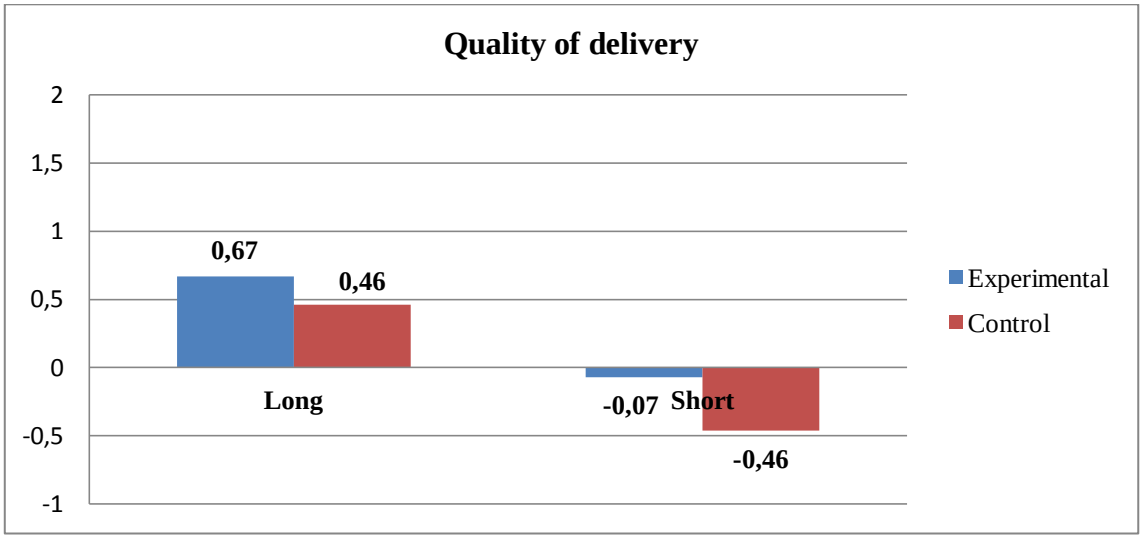


Figure 23. Effects of Training for Prevention and length of the program on the participants report on quality of delivery.

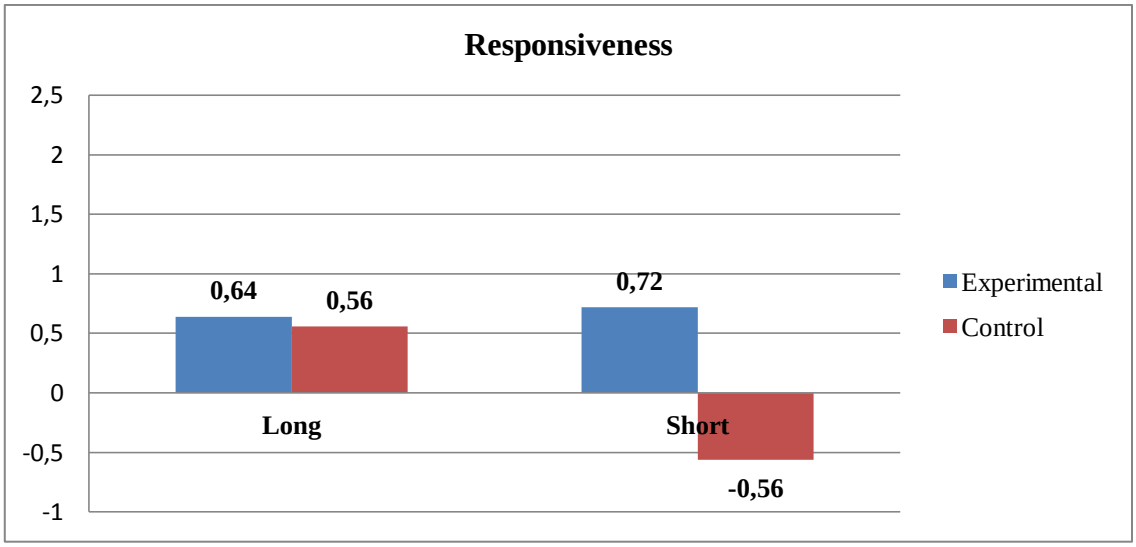


Figure 24. Effects of Training for Prevention and length of the program on the participants report on responsiveness.

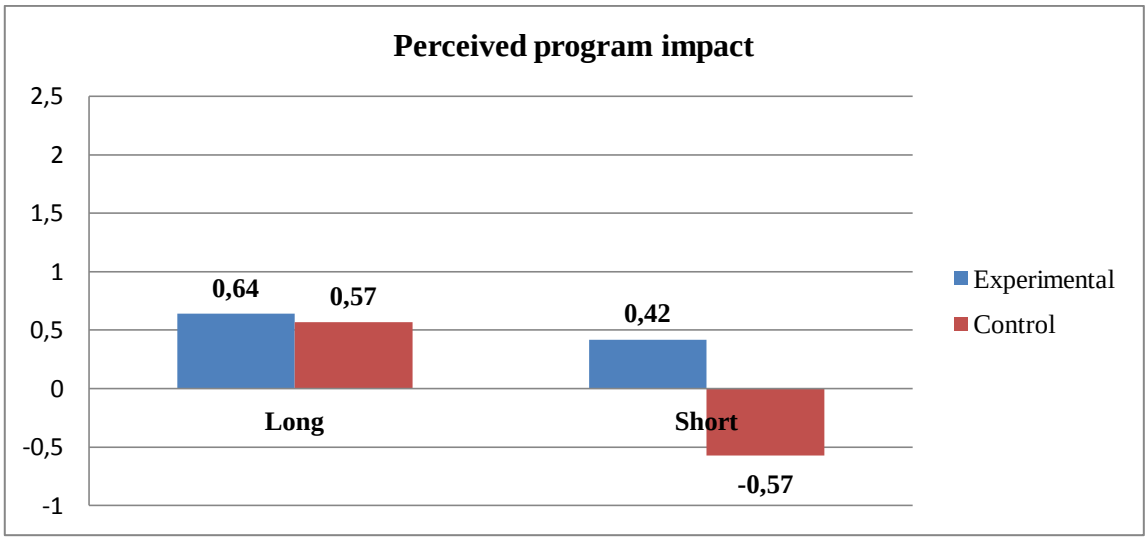


Figure 25. Effects of Training for Prevention and length of the program on the participants report about perceived program impact.

Regarding the significance of Training for Prevention's impact, moderator effects and interaction, analyses show that the Training for Prevention is more effective for short programs. For the longer programs, it did not matter whether managers and implementers had participated in Training for Prevention or not; participants responded approximately equally well under either study condition. For the shorter programs, however, it mattered a great deal whether managers and implementers had participated in Training for Prevention. If managers and implementers from short programs were in the Training: shorter programs in intervention

group had more sessions than short programs in control conditions; short programs from intervention group had less negative ratings of quality of delivery, more positive ratings of participants responsiveness and more positive ratings of perceived program impact.

Besides length of program, **significant moderator models were found also for the moderator activity of the manager** and its impact on the quality of delivery and participants' responsiveness. Effects of manager on dosage and perceived program impact were not confirmed. Effects of those moderator models are presented in Figure 26 and Figure 27 below.

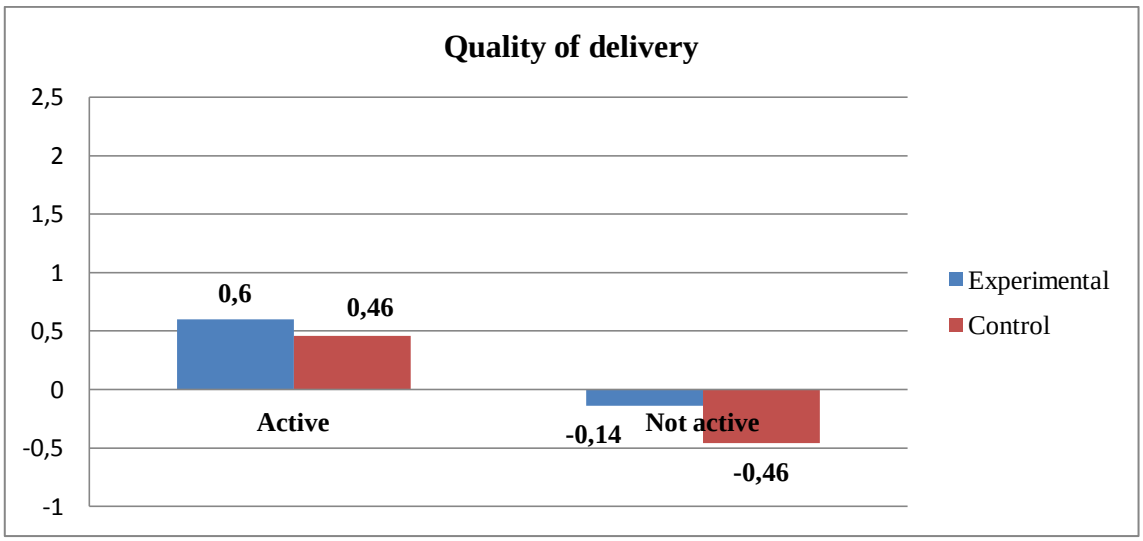


Figure 26. Effects of Training for Prevention and activity of manager on the participants report about quality of delivery.

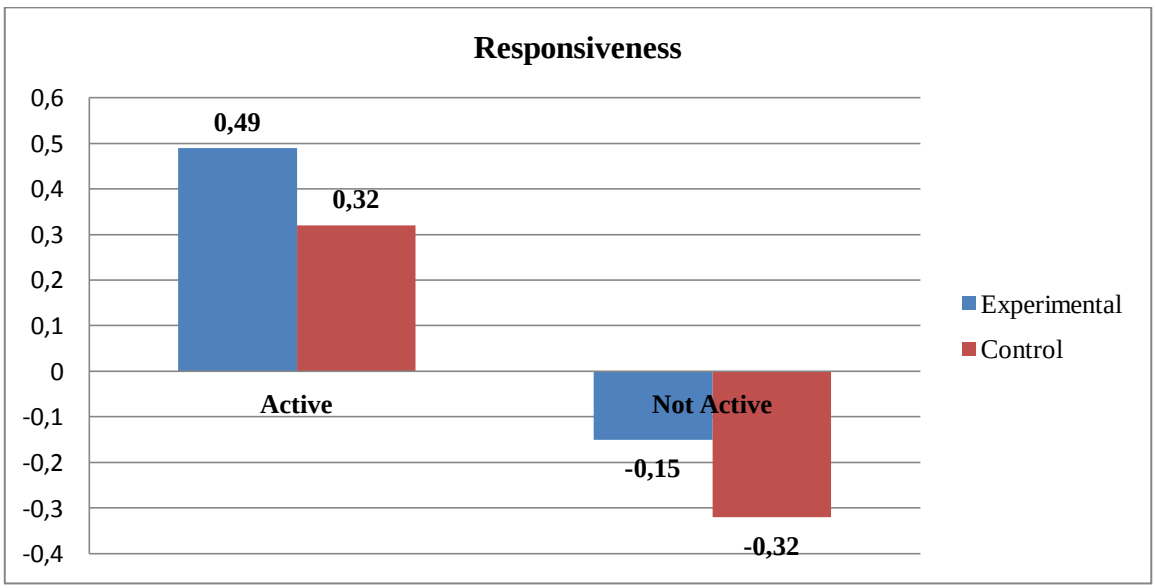


Figure 27. Effects of Training for Prevention and activity of manager on the participants report on responsiveness

As seen from Figures 26 and 27, programs where managers are active generally have higher levels of implementation quality, regardless of whether they were in experimental or control conditions. For programs whose managers are not actively participating in the life of the program and therefore did not even participate in the Training, participation of their implementers in the Training for Prevention affected the reports of participants – they were less negative about the quality of delivery and responsiveness than program participants from control conditions.

7. CHAPTER SEVEN: DISCUSSION AND CONCLUDING REMARKS

This final chapter is going to present an overview of findings which are the result of the studies conducted within this doctoral research. Results will be reflected with the recent literature, to place them within the current implementation research context. All research tasks will be reviewed, with conclusions and directions for future research. After the discussion of overall dissertation outcomes, limitations of the study will be presented followed by recommendations for mental health promotion and prevention practice.

7.1. General discussion

Measures of implementation quality

This doctoral research examined the implementation quality in mental health promotion and prevention programs being delivered in community settings in the County of Istria. It also examined factors related to the programs' characteristics and the support system that in previous research with similar programs are associated with higher levels of quality. It also tested whether an intervention designed to promote these factors and the quality of implementation had a positive effect on these outcome indicators. In order to achieve these goals, the first research task was **to construct valid and reliable measures of implementation quality based on implementation literature and existing measures.** Four measures were constructed according to theoretical definitions of implementation and the conceptual model of implementation created for this research (see Figure 8, Introduction section), distinguishing between implementation factors and indicators of implementation quality.

Studies of implementation have generally found that observational data is more reliable and objective than self-report (Lilehoj et al., 2004; Schoenwald et al., 2010, Domitrovich et al., 2010) which is often inflated due to social desirability on the part of those responsible for the program (Dusenbury et al., 2005; Cross and West, 2011). This is less of a concern for participants who are often unaware of the purpose of data collection regarding implementation. Since direct observation was not possible in this study it was decided to construct parallel measures of both sets of outcomes which would represent different perspectives as is suggested by Knoche, Sheridan, Edwards and Osborn (2010) or Cappella

and colleagues (2008). Managers and implementers were asked to report on implementation factors; while implementers and participants reported on program implementation quality. In that way, multiple ratters have assured greater reliability.

The ***Implementation Factors Questionnaire for Program Managers*** has 31 items in total and consists from items covering six implementation factors: standardization, implementers' skills, attitudes, training, support and monitoring. Results on this questionnaire are expressed in six subscales, each representing one implementation factor (the lowest of six α=.702 for attitudes subscale and the highest α=.870 for implementers' skills subscale). The ***Implementation Factors Questionnaire for Program Implementers*** has 33 items covering five implementation factors: standardization, attitudes, training, support and monitoring. Results on this questionnaire are expressed in five mentioned subscales, each representing one implementation factor (the lowest of six α=.714 for standardization subscale while the highest is α=.808 for support subscale). The ***Indicators of Implementation Quality Questionnaire for Program Implementers*** has 21 items describing constructs of fidelity, quality, responsiveness and perceived program impact. Results on this questionnaire are expressed in four mentioned subscales, each representing one indicator of implementation quality (the lowest of six α=.419 for fidelity subscale while the highest is α=.792 for perceived program impact). The measure for indicators of implementation quality reported by program participants had two versions. The ***Indicators of the Implementation Quality Questionnaire for Program Participants – adult version*** has 35 items covering dosage, quality of program delivery, participants' responsiveness and perceived program impact. Results on this questionnaire are expressed in four mentioned subscales, each representing one indicator of implementation quality (the lowest α=.803 for quality of delivery subscale and the highest α=.893 for responsiveness subscale). ***Indicators of Implementation Quality Questionnaire for Program Participants – child version*** has 20 items covering dosage, quality of program delivery, participants' responsiveness and perceived program impact. Results on this questionnaire are expressed in four mentioned subscales, each representing one indicator of implementation quality (the lowest is α=.689 for quality of delivery subscale and the highest is α=.857 for responsiveness subscale). Items of each measure load very highly on a single component in the factor analyses and show good test-retest reliability. It could be concluded that the first research task was fulfilled: reliability and construct validity analyses which have been undertaken show preliminary support for both reliability and construct validity of the constructed measures.

These four constructed questionnaires represent a contribution to implementation research since this study has provided beginning evidence in support of their validity and reliability. The content of the measures is theoretically driven and all procedures of measures construction were followed, but due to a limited sample size, these measures still have to be validated in future research, especially regarding predictive validity. Also, these measures were found feasible and ecologically valid in the context of our study. As Proctor and colleagues (2011) as well as Schoenwald and colleagues (2010) emphasized, measurement methods with a contextual fit could be easily incorporated within routine practice settings.

Study on implementation quality

The second research task was **to explore the level and variation of implementation quality in preventive programs in Istria**. Since the different informants describing the level of implementation factor or the level of implementation quality potentially have different perspectives on these outcomes, the second research task is closely related to the third research task, **to explore the differences in perception of implementation quality between program managers, program implementers and program participants**. The most important question for discussion here is connected with the definition of implementation success i.e. level of acceptable implementation quality which is often mentioned in literature: what should be considered as implementation quality (Proctor et al., 2010). Domitrovich and colleagues (2008) stress that criteria for high or low implementation quality are dependent on how they relate to program outcomes; although there are few programs which have defined them clearly and into depth. For example, Mihalic and colleagues (2004) in their review of Blueprints for delinquency prevention mention Life Skills Training Program, PATHS program (Greenberg et al., 1999) and Olweus Bullying Prevention Program. Blueprints findings show that programs had average implementation rate of 60 to 70% if they were effective, but it is harder to determine what levels of low, medium and high implementation quality are on the continuous scales. Criteria for low, moderate or high implementation quality need to be standardized to allow comparison between the multiple sites of the same program or between various programs and settings.

It is important to stress that in this study measures of both implementation factors and quality were on a scale ranging from 1 to 4, with exception of dosage which is represented both in the number of sessions and in the percentage of program delivered. In order to describe the level and variation in implementation quality in preventive programs in Istria, it was decided that the theoretical mean, i.e. a result of 2.5 (minimum of 1, maximum of 4)

would serve as a reference point. Results below that value for both implementation factors and indicators of implementation quality are considered low. This solution is somewhat arbitrary and primarily descriptive since this is the first implementation research conducted in Croatia, measures are used for the first time, and they have not yet been validated by being linked to program outcomes. Even though these results represent a contribution to the field of mental health promotion and prevention in Istria and Croatia in general, especially regarding the cohort of included programs, we are aware that the measured level and variability of implementation quality is mainly descriptive and still burdened with questions related to implementation measurement. It has to be stated that problems are connected with the question how many assessments of implementation have to be made in order to capture the process of change adequately and to make conclusions about the quality process of intervention delivery (Domitrovich et al., 2008; Domitrovich et al., 2010;), as well as with absence of a reference point for comparison of implementation quality level. Durlak (2010) states that implementation measurement at only one time point is inadvisable and that implementation should be collected at multiple occasions, weekly or monthly (Baker et al., 2010; Odom et al., 2010). When referring to findings of the study on implementation quality, i.e. second and third research tasks of this dissertation, it has to be stated that only one time assessment of implementation factors and indicators of implementation quality was taken into account.

In general, when looking at the pattern of scores presented in chapter five, results from our study show that implementation factors were consistently rated lower than indicators of implementation quality. This is surprising if we assume these factors strongly influence implementation quality (Fixsen et al., 2009). It was expected that lower ratings of implementation factors would follow with lower ratings of indicators of implementation quality. In our study, managers which have reported on the implementation factors tend to give more strict reviews on the level of factors than program implementers do. From the program managers' point of view, it could be stated that the level of program standardization, organization of training and investments in intervention knowledge are low and they do not represent satisfactory conditions for an implementation process of high quality. The average value of manager's reports for monitoring system are also really close to the cut-off point of 2.5, so one possibility might be that managers are monitoring program delivery poorly and do not communicate regularly with implementers. From the perspective of managers, all these findings could indicate the presence of a low level of implementation quality in the studied

programs. Average results for implementation factors collected from program implementers show a bit different and more positive perception than it was the case for program managers. Program implementers also report that program standardization is the lowest of all implementation factors, but for them the average results go beyond the cut-off point.

Average overall results for all indicators of implementation quality, including fidelity, quality of delivery, perceived participants' responsiveness and program impact are above the value of 3, seen from both the program implementers' and program participants' perspective. If we analyse results for dosage, which was either expressed raw or coded differently than other indicators, (i.e. as the percentage of delivered out of total number of sessions and meetings with participants), most of the programs have delivered around 40% of the whole content by the mid-intervention assessment time point. In general, these positive average results for all indicators of implementation quality support the conclusion that implementation quality is high. Schoenwald and colleagues (2010) address this issue stating that program implementers could be biased towards their own intervention. Since program participants in our study were not aware why implementation quality is being assessed, our expectation is that they are a more reliable source on implementation quality than implementers and this level of implementation quality may reflect the actual quality of the programs that were delivered.

Research task number four was **to explore the relationships between implementation factors and indicators of implementation quality**.When examining those associations, implementation factors and implementation quality rated by the same source were not examined as that would inflate the association. Results presented in chapter 5 on the association between implementation factors and indicators of implementation quality have revealed that these two concepts are not as inter-related as expected. Even though there was not a hypothesis about the level of relationship, there is a great deal of literature suggesting that the implementation factors measured in this study promote higher levels of implementation quality (Kam et al., 2003; Roberts-Gray et al., 2007; Mihalic et al., 2008). For the association of managers' ratings of implementation factors and implementers' ratings of implementation quality, significance is shown only for managers' report on implementers' skills ($B=0.43$, $p<.04$) and managers' report on support ($B=0.35$, $p<.05$) predicting implementers' report on fidelity. Regarding the association of managers' implementation factors and participants' indicators of implementation quality, results show that managers' ratings emerged as a statistically significant predictor for two participants' indicators of

implementation quality: in negative direction for dosage (predicted by managers' attitudes B=-0.42, p<.04) and being positive for participants' responsiveness (predicted by managers' support, B=0.17, p<.05). When inspecting the relationship of implementers' implementation factors and participants' indicators of implementation quality, a statistically marginally significant relation was found for implementers' report on attitudes for only one of the participants' indicators of implementation quality: dosage (B=-0.42, p<.07). There are several factors that make it difficult to see a relationship between implementation factors and indicators of implementation quality. The study sample was small and quality ratings were high, so there may not been enough variation to detect an association. It seems that manager ratings were more highly associated with implementation quality ratings made by participants' level compared with implementers as informants, but this needs additional research and clarifications in a different sample. These indications should be researched in the future, in the context of prospective study since they are not clear in our research, but can be found in previous studies, especially when it comes to contextual factors and leadership (Gingiss et al., 2006; Dariotis et al., 2008; Lochman et al., 2009; Wandersman et al., 2009; Ringwalt et al., 2010; Rohrbach et al., 2010).

Another important issue with our findings could be connected with the definition of implementation quality aspects and accessibility to data on the relevant contextual factors that influence this process. Although the indicators of implementation quality used in this study were based on literature, there are not so many studies, which have researched the relationship of the aspects of implementation quality. Berkel and colleagues (2011) address the issue of confound measures: fidelity and adaptation, fidelity and quality, and quality and responsiveness but the field of implementation research is still rather young and not all relevant implementation factors and aspects have yet been researched. In terms of measuring the relevant implementation factors, several studies from the work of Bosworth and colleagues (Bosworth et al., 1999; Gingiss et al., 2006; Roberts-Gray et al., 2007) suggest that there are almost 300 variables which can be identified in the process of implementation planning and monitoring so it is very possible that some important dimensions were not measured in this study. There is a strong need for further research on the association of implementation factors and aspects of implementation quality, in order to make this picture more complete and evidence-based.

Study on the impact of Training of Prevention

The fifth research task was **to assess the effects of the Training for Prevention on the implementation factors and implementation quality reported by program managers, implementers and program participants**. As it was described, the Training for Prevention was designed to inform program managers and implementers about the importance of effect predictors and encourage them to improve their practices in order to achieve higher levels of implementation. Apart from theoretical background, the Training has offered some exercise and practical work connected with the understanding of logic modelling, importance of objectives and precise defining of intended outcomes, implementation strategy and interactive techniques, evaluation plan, as well as steps for assuring better financial support and recognition in community. In order to make this discussion easier and more transparent, results of statistical tests comparing the intervention and control group are summarized in four summary tables below (Tables 7.1, 7.2, 7.3 and 7.4). Analyses have shown that according to program managers and implementers implementation factors did not improve when compared to the control group, so Hypothesis 5.1 had to be rejected.

It is important to note that analyses have shown that the intervention group of managers and implementers has rated implementation factors lower than the control group. Managers' analyses are showing significant differences at post-test for three out of six implementation factors: standardization, training and intervention knowledge as well as monitoring. The differences between the groups on two of these three dimensions (standardization and monitoring) were already evident at mid-intervention time-point, while implementers have reported on significantly lower standardization at mid-intervention assessment. Since Training took place before the mid-intervention implementation assessment, all of the mentioned factors were covered in the Training for Prevention themes, so these negative effects could be also viewed as an outcome of the intervention. It is possible that these results are the effect of insufficient knowledge and acknowledgment of science-based implementation principles in the control group. It is possible that managers and implementers from control conditions perceived implementation factors level more unrealistically and maybe they see themselves, their implementers and organizations in whole in overly positive light. Managers and implementers from intervention conditions might have raised their awareness through the Training and could have learned to be more critical in the evaluation of their own programs and organization.

Regarding indicators of implementation quality, program implementers from the experimental group did not report improved indicators of implementation quality in comparison with the control group. When participants' ratings of implementation quality were analysed, post-test differences have shown that participants from the experimental group report higher levels on all indicators of implementation quality than participants in control conditions, two of them being significant - quality of delivery and responsiveness. We can conclude that hypothesis 5.2 was partly confirmed.

Regarding the Training for Prevention intervention effectiveness and its implications on the level of implementation quality, it could be concluded that its impact still has to be further researched. It was expected that Training for Prevention effects would be visible on the ratings of implementers because they were directly included in the intervention. At the same time, results on the participant level were more distant, so not expected. It was surprising to find results on the ratings of implementation quality provided by participants, without sufficient explanation why those changes did happen. It was expected that improvements reported by participants would be caused by changes in implementation factors reported by managers and implementers but it is possible that changes took place that were the result of factors that were not measured in this study. Also, it may be that the influence on quality needs more time to emerge than the period of time that our study measured. Even though there were no significant group differences on implementation quality according to implementer ratings, the participant ratings are potentially valid and less biased since participants were not familiar with the fact that managers and implementers participated in the Training, and so they did not have pressure to report positively.

Table 7.1

Summary Group Comparisons of Manager Ratings of Implementation Factors

	Group Comparisons		
	Mid-Evaluation	**Mid to Post**	**Post-test**
Standardization	Control > Intervention	NS	Control > Intervention
Implementers' skills	NS	NS	NS
Attitudes	NS	NS	NS
Training	NS	NS	Control > Intervention
Support	NS	NS	NS
Monitoring	Control > Intervention	NS	Control > Intervention

Table 7.2

Summary Group Comparisons of Implementer Ratings of Implementation Factors

	Group Comparisons		
	Mid-Evaluation	**Mid to Post**	**Post-test**
Standardization	Control > Intervention	NS	NS
Attitudes	NS	NS	NS
Training	NS	NS	NS
Support	NS	NS	NS
Monitoring	NS	NS	NS

Table 7.3

Summary Group Comparisons of Implementer Ratings of Implementation Quality

	Group Comparisons		
	Mid-Evaluation	**Mid to Post**	**Post-test**
Fidelity	NS	NS	NS
Quality	NS	NS	NS
Responsiveness	NS	NS	NS
Perceived Impact	NS	NS	NS

Table 7.4

Summary Group Comparisons of Participant Ratings of Implementation Quality

	Group Comparisons		
	Mid-Evaluation	**Mid to Post**	**Post-test**
Fidelity	NS	NS	NS
Quality	NS	NS	Intervention > Control
Responsiveness	NS	NS	Intervention > Control
Perceived Impact	NS	NS	NS

These findings were supplemented by the moderator analyses. The purpose of these analyses was to determine whether variables like activity of a manager, length of a program, type of participants and type of program qualified the effects of Training for Prevention on both sets of outcomes. There were no significant moderation effects for measures of

implementation factors and the only significant effects for ratings of implementation quality were for participants' reports.

Results showed that the length of a program modified the effects of the Training on participants' report of dosage, quality of program delivery, responsiveness and program impact. The Training for Prevention was more effective for short programs. For the longer programs, it did not matter whether managers and implementers had participated in Training for Prevention or not; participants responded approximately equally well under either study condition. For the shorter programs, however, it mattered greatly whether managers and implementers had participated in Training for Prevention. If managers and implementers from short programs were in the Training: participants from shorter programs reported higher dosage levels than short programs in experimental conditions; short programs from intervention group had less negative ratings of quality of delivery, more positive ratings of participants responsiveness and more positive ratings of perceived program impact.

When going back to written project proposals collected from the short programs prior to the Training, to the qualitative data and experience with those organizations which were running the short programs, it can be stated that those organizations either had lower levels of expertise or had lower structural capacity within the organization to run the program. Regarding the mental health promotion and prevention practice, they had less knowledge on evidence and science based prevention, they had mostly delivered programs in ex-cathedra manner and did not do much to improve the quality of program delivery. Organizations from the experimental condition which were providing longer and more elaborate programs are also the organizations that already had more basic knowledge about mental health promotion and prevention as well as implementation quality. These moderator analyses suggest that in the future, the Training for Prevention might be most appropriate for managers and providers that have proposed short prevention programs. More in general, moderation findings may suggest that the Training may need to be less standardized for all participants, and tailored to the capacities and needs of the different organizations, managers and implementers.

Besides length of program, significant moderator models were found also for the moderator activity of the manager and its impact on the quality of delivery and participants' responsiveness. Programs where managers are actively involved generally have higher levels of implementation quality reported by participants, regardless of whether they were in

experimental or control conditions. For programs whose managers are not actively participating in the life of the program and therefore did not even participate in the Training, participation of their implementers in the Training for Prevention affected the reports of participants – they were less negative about the quality of delivery and responsiveness than program participants from control conditions.

These findings seem interesting not only from the perspective of the Training, but also from the perspective of the influence of managers on the quality of delivery and responsiveness from participants' perspective. Managers are regarded as the ones with the most important role for the implementation quality in other research as well. For example, Saunders and colleagues (2006) report on the measurement of implementation of LEAP, a multi-component public health intervention which was developed to promote physical activity in high school girls. Measurement of implementation included process measures of school organization and environment as unit of analysis. When comparing high-implementers and low-implementers, Saunders and colleagues (2006) saw that the differences were connected with organizational level components i.e. school environment: higher implementing schools had a more active school health promotion staff. Midthassel and Ertesvåg (2007) highlight the role of managers with the level of readiness for innovation i.e. new program: schools with higher readiness had better implementation quality. Readiness for prevention is a stage of preparation of a system, organization or personnel to meet a situation and lead to the planned change. The same study underlines the role of ecological factors, especially the role of school principle and leadership: having a plan of anti-bullying work in school which is integrated in everyday school life was particularly significant for implementation quality. Also, crucial was the headmasters' ability to inspire staff to commit to a program and follow the plan.

Another study which has found evidence for an impact of managers is the study of PATHS, empirically tested school-based curriculum for preventing aggression and promoting positive development, from Kam, Greenberg and Walls (2003). Implementation in their research was measured by observational ratings of classrooms which observed aspects of teacher mastery and principal support. The study has found significant effects for effectiveness only for those settings where overall implementation was high and had clear principal support.

7.2. Limitations of the study and suggestions for future research

There were several limitations and challenges in this study that have to be discussed, especially those connected with methodology and its repercussions on the analyses as well as those connected with the sample size. All presented limitations will be followed by ideas and recommendations for future research.

Limitations of the measures

The first is connected with the fact that the four measures of implementation factors and implementation quality were not constructed and pretested in preliminary research and then used in this research with more experience and confidence. Preliminary research would probably assure more data and orientation for more precisely defining different levels of low or high implementation quality. Also, our results which show adequate reliability and construct validity have to be taken as preliminary and confirmed in future research. Additionally, it has to be stated that all findings for the subscale of fidelity have to be considered with caution. Cronbach's alpha for fidelity was below our expectations (α=.419 for the first measurement and α=.446 for the post-test assessment). Internal consistency of the items under the construct of fidelity could be this low not because of unreliability of these items but because of the characteristics of Croatian programs which are not evidence-based, which is especially affecting fidelity.

This raises the question of understanding of core components by program managers and program implementers. Also, this research is one of the few studies that examine implementation quality within programs which are not evidence-based. Although dosage, quality of delivery and responsiveness are more connected with implementation process and could be studied also within programs whose outcomes are not confirmed yet, fidelity is more specific to the already proven interventions. This study aimed at capturing that aspect of implementation quality, but that still needs to be researched in future studies. As already stated, although all four measures have been proven valid and reliable at a preliminary level, the question what is low, moderate or high implementation quality still remains unanswered. This calls for further, more thorough research in future studies, not within a correlational design but with longitudinal assessments of implementation quality and program outcomes.

Issues of sampling and power

The most important limitation for all three presented studies in this doctoral research concerns sampling and power. Although it was challenging to organize the data collection for 24 programs, regarding their specific features, types of participants, settings of implementation and number of sessions, the limited sample size did not provide enough statistical power for some findings and conclusions. We have learned that it is also necessary in this type of studies to run a power analysis on the needed number of programs and subjects in this study to prove significant effects if these would be present in reality. The question of power is especially connected with the implementers and managers. We had only 22 managers in the post-test analyses and 55 program implementers nested.

It is important to mention that chapter six also presents findings in terms of effect sizes (many were small to moderate) as a way to help interpretation of group differences because of the lack of power to find statistically significant differences. Although we have presented all results, for those two sources of data on implementation quality, findings should be viewed as preliminary for general conclusions. However, our sample of programs and experts covers probably a very large part of the population of programs and practitioners in this field in the County of Istria. It could be stated that when interpreting data to make statements just about Istria, the strict $p>.05$ is less relevant since we are not generalizing the data for whole Croatia. Of course, additional studies on implementation quality and the impact of the Training of Prevention have to be conducted on many more programs outside the cohort of programs involved in this study. This could even be a national study on implementation quality in mental health promotion and prevention. In our study, we have combined programs that were very different, so in future larger studies efforts have to be made in order to make the program constant and examine one program with larger sample of implementers and participants.

Additionally to the problem of the relative small number of programs in the sample, there were inconsistencies with data collection and number of participants regarding the point of measurement. More specifically, the number of managers, implementers and participants differed between the study on implementation quality and study of the Training for Prevention impact, i.e. between the mid-implementation assessment and the post-implementation assessment. The greatest differences in the number of collected questionnaires are present at participants' level. At the mid-implementation assessment 434 participants rated indicators of implementation quality, while at post-test assessment data was collected from 744 participants. That is confusing for the presentation of data, complicated for the reader, it

prevents solid conclusions about all programs in the studied cohort, and it aggravates the tracking of implementation quality processes. Those differences in sample size were due to the fact that some programs were short (for example, only one session), so they had only the possibility of post-test measurement of implementation quality or could not organize two data collections for other reasons. As a consequence, five programs did not have mid-implementation assessment, so the first set of analyses on general implementation quality had only seventeen programs in the studied cohort. The data at post-test was somewhat biased: the programs with a post-test only were mostly short and to our experience had worked with participants differently; three out of the five programs saw their users only once and had only a very limited communication with them.

Problems with matching during the data collection complicate the situation with participant data. Many participants were not matched between two time points, while a range of them forgot their identification codes or did not code their ratings. Because of that, an analysis connecting the two time points to test the impact of the Training for Prevention was done only for participants from the eighteen programs that have had data collection in both time points. Because participants' forms were not matched through the two assessments, per each of the participants' ratings of implementation quality average program results for the whole sample from mid-intervention was used.

Timing of the implementation measurement and timing of the Training for Prevention

The measurement issue that has to be raised here is connected with the decision to assess implementation quality twice: after 1/3 of program delivery and at the end of program delivery. A program already had to be in the process of delivery to make an assessment of indicators of implementation quality possible. One third was chosen as an approximate point: programs in the studied cohort are so diverse that the actual implementation process that was performed at 1/3 of the planned process differed hugely between the programs.

The problem with the timing of measurement was additionally complicated by the timing of Training for Prevention intervention. It would have been ideal if the first assessment of the implementation quality had happened before the delivery of Training. That was not possible since there were great diversities between programs in the timing of when during the year program delivery starts as well as in the length of the programs. It was not possible to control all of these factors since we had real life settings and community based programs which had their tradition and habits before the project started. A significant limitation is also

connected with the fact that five out of twelve programs from the experimental conditions started with program delivery before the Training for Prevention was finished and delivered to them in its entirety. This heavily reduces the possibilities of implementing implications from the training messages in their already ongoing programs. An additional problem was the difference in participation level in the training between managers and implementers. All implementers from the 12 programs attended the Training regularly, while that was not the case for managers, only half of them attended the Training and were not present at every session.

Short study period

The conclusions of this study are also burdened since this research includes a cohort of projects that have not yet been proven effective, and since there was no possibility of including a control group for each of the programs. If there had been enough time to have at least a two year project, the first implementation measures would not have been confounded with the Training, and there would have been more possibility to organize control groups within each of the programs in the studied cohort. This would have led to a greater research rigour of the study design and better statistical power would be achieved. A longer study period would also have facilitated the incorporation of knowledge that was delivered throughout the Training for Prevention into practice. Our study design did not allow enough time for program changes and for incorporation of science-based principles in program settings. The time limit had also policy reasons, since the Department of Health and Social Services is making a new health plan for the County of Istria and it aims to use the findings from this study for making improvements in their policy and procedures for supporting prevention programs. The current findings can be perceived as a contribution to this purpose, but have to be researched thoroughly through more sophisticated designs in the future. The presented conclusions about the Training for Prevention and its outcomes cannot be stated with maximum certainty. Given the used training design, we are not able to argue that the Training and the outcome study have enough power to create and to prove behaviour change.

Statistical significance of effect sizes

With these limitations in mind, we expect that more of the small and moderate effect sizes for the managers and implementers reports might have been significant if the study would have had more statistical power. For example, when mid-assessment is controlled, participation in the Training for Prevention resulted in a d=.30 for the managers' report on implementers skills, while implementers' report on training and intervention knowledge is d=.28 although not statistically significant. The same is evident for control group of managers at post-test which had negative effect sizes d=-.55 for attitudes and d=-.42 for support, as well as implementers which had moderate negative effects at post-test, all of them not being significant. Implementers' ratings on indicators of implementation quality also show positive direction of Training's impact when mid-assessment is controlled: effect size for fidelity is d=.23 while d=.13 for perceived program impact. Durlak (2010) discusses the interpretation of effect sizes, mentioning that interpreting their magnitude in the field of implementation research is somewhat premature while we do not have yet a sufficient database for judging them. He clearly states that researchers should not use Cohen's (1988) conventions since they are only guidelines. Durlak and colleagues (2011) present meta-analysis of 213 school-based universal socio-emotional learning programs and present effect size for the SEL programs. For example, d=.57 was found for the effect of SEL on skills, d=.23 for attitudes, d=.24 for emotional distress and d=.27 for academic performance. Durlak (2010) warns that we need a lot more information on what effects can be achieved in implementation research and what their practical pay-off is.

Implementation outcomes vs. program outcomes

It has to be stated that when the thesis proposal for this dissertation was defended, one of the aims of the study was also to test if higher implementation quality is connected with higher program outcomes. Connection of implementation outcomes and program outcomes such as behaviour change, attitude change or knowledge enhancement would lead to the higher certainty on the predictive validity of our measures, a clearer picture what actually implementation quality is and of what importance the implementation is for program effects. As stated in the review of Durlak (2010), it would be very informative to clarify which aspects of implementation are the most important for different outcomes. Research on program outcomes was a part of our wider project "Preffi – Quality assurance in the County of Istria" and scope of the dissertation *Study of effectiveness of prevention programs* of a

colleague, Josipa Mihic. Due to the fact that participants were not matched adequately and that program effect sizes were calculated on group level while the analyses in this dissertation are on the individual level, it was not possible to address this interesting issue at this point. That remains to be resolved in future analyses.

Additional recommendations for future research

It is recommended that future research of implementation factors and indicators of implementation quality pay more attention to design rigour. If the Training for Prevention would be additionally researched, careful matching of programs in the control and experimental group has to be conducted, assuring that programs in the control group are not different as was the case in the presented study. Implementation quality and Training's impact have to be researched within the same programs, comparing the implementation outcomes in the group of participants whose implementers got the Training with the implementation outcomes of the group of participants which have been included in the same program but implementers did not receive the Training.

With regard to the study on implementation quality and taking into account that this research represents a novelty not just for Croatian but also for international conditions, we recommend to complement quantitative studies with qualitative study designs to get a more deeper insight into implementation processes and their facilitating conditions. A recommended and important research direction is the study of relationship between implementation factors and indicators of implementation quality as well as the study of relationship between different indicators of implementation quality. To date, the knowledge about these relationships is still very limited. For the field of mental health promotion and prevention to continue to grow, greater attention and better understanding of the implementation process and the factors that support it, is essential. Regardless of the mentioned agenda, the first step to help achieve more clarity in the implementation field is through the development of a comprehensive implementation theory model that integrates different perspectives presented in the literature (Domitrovich et al., 2010). Additional interesting questions concern the review on implementation by Durlak (2010): it is needed to determine who should provide necessary data, when implementation assessments should be done, and what ecological factors should be evaluated.

7.3. Practical implications

7.3.1. Recommendations for the County of Istria

This study on the implementation quality of 24 prevention projects in the County of Istria has direct practical repercussions, which follow from the level of implementation factors as well as from the level of indicators of implementation quality. According to the results of implementation factors, some recommendations could be drawn. It seems that several programs need serious investments in improving the capacity of structure and conditions for implementation. First, it could be stated that all organizations in the studied cohort would benefit from investments in the level of program standardization. That result was somewhat expected because from the researcher's experience with the studied practices, programs included in the sample do not have clear guidelines for program delivery. They sometimes even do not have an exact set of themes the program has to cover and sometimes changes in program are made without sufficient scientific base. The described findings of this study could be of great interest to the Department of Health and Social Services in Istria which is financing those programs and has the right of raising questions why invested resources are not being spent well.

Also, according to the results on implementation factors, the level of training and intervention knowledge is another important area for future investments. Both the organizations conducting interventions and the administration of the Department of Health and Social Services in the County of Istria should invest more in the training of implementers, organize more in-service trainings and assure the translation of up-to-date skills and knowledge necessary for program delivery. That is important because program implementers often have different professional backgrounds, levels of experience and skills, and at the same time, they are the most important stakeholders of mental health promotion and prevention policy. For some of the implementers, delivering interventions in the field of mental health promotion and prevention is not a permanent position but a part-time job, which is not controlled and standardized. This might affect the quality of implementation in general and indirectly the quality of life in end-users. Since those organizations and implementers are stakeholders which carry out the county's policy, The County of Istria should organize trainings which would contribute to the level of knowledge and professional readiness of those people because they are working with children, youth and families, and are required to do this with high quality and effectively.

Domitrovich and Greenberg (2000) emphasize that all implementation research should begin with specifying program components and active ingredients to reveal a theory behind an intervention. This still remains to be done for the researched set of programs in the cohort as well as with other mental health promotion and prevention programs in the County of Istria and Croatia in general. According to these results, the Administration of the Department of Health and Social Services could analyse the financial investments in the studied group of programs, and could set guidelines or even obligatory rules to safeguard sufficient levels of program and implementation quality. Such criteria could include specific guidelines on the level of needed capacity, i.e. investments in organizational support for adequate implementation for the programs financed by the Department. Domitrovich and colleagues (2008) emphasize that both the intervention and its support system have to be standardized, have to specify its core elements and a delivery model. First, considerably more research needs to focus on core intervention components to open up the "black boxes" of locally based practices and programs. There are a lot of possibilities for such studies which do not need to be expensive and mostly depend on the motivation and readiness of mental health promotion and prevention experts. The Administration of the Department of Health and Social Services has to lead those practitioners towards the conviction that working in this field is serious work which requires sufficient level of knowledge and professionalism. Eventually, these core intervention components (e.g., relationship development, skill teaching, and collaboration) could be taught to practitioners more generally (Meyers, Durlak and Wandersman, 2012), perhaps as part of secondary education curricula and other workforce development initiatives, in order to enhance the quality of implementation.

The findings of this study suggest that not only the organizations themselves, but also the Department of Health and Social Services have to invest in the studied implementation factors. Since all 24 programs included in this studied cohort have in the end received the Training for Prevention, knowledge and skills of science-based mental health promotion and prevention should also be transferred to the evaluators which are conducting the quality appraisal of project proposals of interventions for children, youth and families in the County of Istria. That would also imply that the Department has to invest in sufficient standardization, monitoring, training and intervention knowledge, support and skills of all personnel and external associates.

Also, if possible, the Department could employ a mental health promotion and prevention expert which could serve as a consultant (purveyor in the work of Fixsen et al., 2005), a

person taking care of implementation quality. An advantage of having a well-organized and persistent approach to implementation of evidence-based practices and programs may be that the consultant can accumulate knowledge over time. Based on their experience and expertise, consultants could learn to improve implementation approaches early in the process of program development and implementation to avoid some of the later problems. In addition, an experienced consultant can describe to the managers of an implementation site possible problems and possible solutions that can be applied. Developing sufficient capacity for implementation is essential for helping local providers to conduct new programs effectively. The extent of their success will depend on the interaction of multiple ecological factors that contribute to capacity (Durlak and Dupre, 2008).

It would be interesting if the Department's administration would also recommend to organizations to use the newly constructed measures for implementation assessment regularly. Such data could be used to test their value as a predictor of program outcomes, but at this stage it should serve as information for structural capacity and investments. There are possibilities to make those questionnaires available for online assessments. The collected data could immediately be included into a unique regional or national database, which can be used for research on implementation. That would be very valuable for the mental health promotion and prevention practice in the County of Istria and serve as a good example for Croatia as a whole. Similar suggestions are offered by Meyers, Durlak and Wandersman (2012) who explain that user-friendliness can be increased by manuals, guides, worksheets or other tools to aid the dissemination of that kind of innovation. That would lead to specific capacity for innovation and for capacity regarding organizational functioning. To conclude this part of practical recommendations, we refer to Fixsen and colleagues (2005) who state that implementation appears most successful when: 1) Carefully selected practitioners receive coordinated training, coaching, and frequent performance assessments; 2) Organizations provide the infrastructure necessary for timely training, skilful supervision and coaching, and regular process and outcome evaluations; 3) Communities and consumers are fully involved in the selection and evaluation of programs and practices; 4) State and federal funding avenues, policies, and regulations create a hospitable environment for implementation and program operations.

7.3.2. Perspectives of Training for Prevention capacity

Regarding the Training for Prevention, several practical recommendations could be drawn from the findings in the conducted study of Training's impact on implementation quality. First, the Training for Prevention should be more targeted at managers and providers that have proposed short prevention programs, since the study shows that the training especially benefited them. Secondly, findings may also suggest that the Training should be less standardized for all participants, but more tailored to the capacities and needs of the different organizations, managers and implementers. Thirdly, our findings suggest that the Training is more needed for those organizations where the role of the manager is just formal, without any engagement and activity connected with program implementation and quality assurance.

Capacity of the Training still needs further exploration. When the Training was delivered to the experimental group of managers and implementers, all participants were asked to report on their level of satisfaction with the intervention content, with pacing of delivery and knowledge of the trainers. The feedback gathered was positive but more related to process evaluation. Accordingly, to evaluate the Training delivery more into depth, while delivering the Training for Prevention to the control group of participants in the waiting-list condition, researchers have decided to apply pre and post measurements of Training's effects: a questionnaire on prevention readiness and a test of knowledge connected with the Training's content. Since only eighteen people from the control condition have participated in the second Training delivery, findings are just providing an indication of direction about the possible Training's impact. These participants have improved their knowledge, but there was no indication of an effect of Training on their level of readiness. Development and future investments in the Training for Prevention should also take into account conditions for Training's effectiveness.

Regarding the findings presented in this dissertation which are inconclusive, the development of the intervention should follow some changes in the pacing of Training's delivery: six main topics should in the future be delivered across a longer period of time, with more practical assignments which would assure that attendees understand the concepts and are able to apply them. That kind of approach would develop skills and behaviours besides knowledge and allow the incorporation of Training's principles into practice. According to our findings, it might be stated that more time has to be invested in the topics which deal with

the role of manager, importance of relationship between managers and implementers, systems for monitoring, standardization of delivery, core program components as well as the importance of fidelity and adequate dosage. The experience with delivering the Training for Prevention was absolutely positive for the authors and researchers of the intervention, but one significant practical implication, important for the Training as well as the Department of Health and Social Services, lies in the observation that the Training allows and spontaneously supports networking between organizations and programs. The capacity of the Training could be further enhanced in the future by the formalization of those networks and partnerships with the help of technology and virtual communities. In that way, Training materials, knowledge transfer, advocacy and impact could be more accessible for practice, more permanent and thorough.

7.4. Conclusions

This doctoral research examines the issue of implementation quality in mental health promotion and prevention programs that are being delivered in community settings in the County of Istria, taking into account the programs' characteristics and the support system surrounding them. **The general aim of this doctoral research was to study implementation processes and their outcomes in prevention programs in Croatia.** In order to achieve the above stated aim, this doctoral research was conducted through pre-research concerning the construction of implementation scales and two studies, 1) study on implementation quality and 2) study of the impact of Training for Prevention.

The first research task was **to construct valid and reliable measures of implementation quality based on implementation literature and existing measures.** Four measures were constructed according to theoretical definitions of implementation and the conceptual model of implementation created for this research; distinguishing among implementation factors and indicators of implementation quality. Managers were asked to report on implementation factors; implementers rated both implementation factors and indicators of implementation quality, and participants only reported on program implementation quality. The ***Implementation Factors Questionnaire for Program Managers*** has 31 items in total and consists of items covering six implementation factors: standardization, implementers' skills, attitudes, training, support and monitoring. The ***Implementation Factors Questionnaire for Program Implementers*** has 33 items covering five implementation factors: standardization, attitudes, training, support and monitoring. The ***Indicators of Implementation Quality Questionnaire for Program Implementers*** has 21 items describing constructs of fidelity, quality, responsiveness and perceived program impact. Measure for indicators of implementation quality reported by program participants had two versions: ***Indicators of Implementation Quality Questionnaire for Program Participants – adult version*** which has 35 items and ***Indicators of Implementation Quality Questionnaire for Program Participants – child version*** which has 20 items covering dosage, quality of program delivery, participants' responsiveness and perceived program impact. Four constructed questionnaires represent a contribution to the implementation research since they have been confirmed as valid and reliable, which should be considered still as preliminary since this is the first study on their validity and reliability, and the data are still based on a small number of managers and implementers.

The second research task was **to explore the level and variation of implementation quality in preventive programs in Istria**. Since that level and variation is closely connected with the source i.e. informant describing the level of implementation factor or the level of implementation quality, the second research task is closely related to third research task, **to explore the differences in perception of implementation quality between program managers, program implementers and program participants**. In general, results have shown that implementation factors are consistently rated lower than indicators of implementation quality. In our study, managers which have reported on the implementation factors tend to give more strict reviews on the level of factors than program implementers do. From the program managers' point of view, it could be stated that program standardization, organization of training and investments in intervention knowledge are low and they do not represent satisfactory conditions for an implementation process of high quality. The average value of manager's reports for monitoring system is also close to the chosen normative cut off point of 2.5, so it can be concluded that managers are not monitoring program delivery with quality and it seems that they do not communicate with implementers regularly. All of these findings would indicate lower implementation quality. Average overall results per indicators of implementation quality, including fidelity, quality, participants' responsiveness and perceived program impact, are all above the value of 3, both seen from program implementers' and program participants' perspective. On the general level, average results per indicators of implementation quality support the conclusion that implementation quality in the studied mental health and prevention programs in the County of Istria is satisfactory to high as perceived by implementers and participants. This still needs further research since the experience of the researcher was that implementation quality in the studied program cohort needed investments.

The research task number four **was to explore the relationships of implementation factors and indicators of implementation quality**. Results have revealed that these two concepts are not as inter-related as expected. It could be stated that our design does not allow us to make conclusions regarding the nature of relationship between implementation factors and indicators of implementation quality. It seems that role and standpoint of the manager is a better predictor of implementation outcomes for the participants' level in comparison to implementers as informants, but this needs additional research and clarifications in a different sample with use of an improved research design.

The fifth research task was **to assess the effects of Training for Prevention on the implementation factors and implementation quality reported by program managers, implementers and program participants**. The study on the impact of the Training for Prevention examined whether the newly designed Training for Prevention affected implementation factors and implementation quality in the experimental group. Analyses have shown that program managers and implementers from the experimental group do not report about improved implementation factors when compared to the control group, so **Hypothesis 5.1 had to be rejected**. Regarding indicators of implementation quality, program implementers from the experimental group do not report improved indicators of implementation quality in comparison with the control group. Analysing participants' ratings of implementation quality,, post-test differences have shown that participants from the experimental group report on the higher level of two indicators of implementation quality, quality of delivery and responsiveness than participants in the control conditions. **Hypothesis 5.2 was partly confirmed.** As discussed earlier in this final chapter, there are reasons to assume that these results are influenced by changed and more restricted quality norms among managers and implementers as a consequence of their participation in the Training.

These findings were supplemented by moderator analyses which intended to test if the Training for Prevention had a different effect in specific groups of managers, implementers and participants. The purpose of these analyses was to determine whether variables such as activity of a manager, length of a program, type of participants and type of program qualified the effects of Training for Prevention. Moderator analyses have shown significant results of some moderators only for the participants' report on indicators of implementation quality, but not for managers' report on implementation factors or implementers' report on implementation factors or indicators of implementation quality.

Regarding the significance of the Training for Prevention's impact, moderator effects and interaction analyses show that the Training for Prevention is more effective for short programs. For longer programs, it did not matter whether managers and implementers had participated in the current version of the Training for Prevention or not; participants responded approximately equally well under either study condition. For the shorter programs, however, it mattered a great deal whether managers and implementers had participated in the Training for Prevention. Besides program length, significant moderator models were found also for moderator activity of the manager and its impact on the quality of delivery and participants' responsiveness. Programs where managers are active generally have higher

levels of implementation quality reported by participants, regardless of whether they were in experimental or control conditions. For programs whose managers are not actively participating in the life of the program and therefore did not even participate in the Training, participation of their implementers in the Training for Prevention affected the reports of participants – they were less negative about the quality of delivery and responsiveness than program participants from control conditions.

This doctoral research contributes to the field of mental health promotion and prevention science in Croatia as a first example of implementation quality research. Regarding the fact that this field is emerging in worldwide context, scientific contribution could be seen in four newly constructed measures as well as in the invention of the new Training for Prevention intervention which could enhance the quality of mental health promotion and prevention programs.

REFERENCES

Ajzen, (1991).The Theory of planned behavior. *Organizational Behavior and Human decision processes, 50*, 179-211.

Baron, M. R., Kenny, D.A. (1986). The moderator-mediator variable distinction in social Psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51 (6)*, 1173-1182.

Bartholomew, L.K., Parcel, G.S., Kok, G., Gottlieb, N. S. (2006). *Intervention mapping: designing theory and evidence based health promotion programs*. San Francisco: Jossey-Bass.

Bašić, J. (2009). *Teorije prevencije: prevencija poremećaja u ponašanju i rizičnih ponašanja djece i mladih*. Zagreb: Školska knjiga.

Bašić, J., Ferić Šlehan, M., Kranželić Tavra, V. (2007a).*Zajednice koje brinu-model prevencije poremećaja u ponašanju: Strategijska promišljanja, resursi i programi prevencije u Istarskoj županiji.* Pula: Istarska županija.

Bašić, J., Ferić Šlehan, M., Kranželić Tavra, V. (2007b).*Zajednice koje brinu-model prevencije poremećaja u ponašanju: Epidemiološka studija – mjerenje rizičnih i zaštitnih čimbenika u Istarskoj županiji.* Pula: Istarska županija.

Bašić, J., Grozić-Živolić, S. (Eds) (2010). *Zajednice koje brinu - Model prevencije poremećaja u ponašanju djece i mladih: Razvoj, implementacija i evaluacija prevencije u zajednici*. Pula – Zagreb: Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu i Istarska Županija.

Bašić, J., Janković, J. (Eds) (2000). *Rizični i zaštitni čimbenici u razvoju poremećaja u ponašanju djece i mladeži*. Zagreb: Povjerenstvo Vlade Republike Hrvatske za prevenciju poremećaja u ponašanju djece i mladih,.

Bašić, J., Janković, J. (Eds) (2003). *Lokalna zajednica – izvorište Nacionalne strategije prevencije poremećaja u ponašanju djece i mladih.* Zagreb: Povjerenstvo Vlade Republike Hrvatske za prevenciju poremećaja u ponašanju djece i mladih.

Bašić, J., Koller-Trbović, N., Žižak, A. (1993).*Integralna metoda u radu s predškolskom djecom i njihovim roditeljima: teorijski pristup*. Zagreb: Alinea.

Bašić, J., Mihić, J., Novak, M. (2010). Risk analysis in the period of growing-up of children and youth: starting point for effective prevention. *Journal of Public Health, 19,* 3-11.

Berkel, C., Mauricio, A. M., Schoenfelder, E., Sandler, I.N. (2011). Putting the Pieces Together: An Integrated Model of Program Implementation. *Prevention Science, 12,* 23-33.

Bosworth, K., Gingiss, P., Pothoff, S., Roberts-Gray, C. (1999).A Bayesian Model to Predicts the Success of the Implementation of Health and Education Innovations in School-centered Programs. *Evaluation and Program Planning, 22 (1),* 1-11.

Botvin, G.J. (2004). Advancing Prevention Science and Practice: Challenges, Critical Issues and Future Directions. *Prevention Science, 5 (1),* 69-72.

Botvin, G.J., Griffin , (2007). School-based Programmes to Prevent Alcohol, Tobacco and Other Drug Use. *International Review of Psychiatry, 19 (6),* 607–615.

Brace, N., Kemp, R., Snelgar, R. (2006). *SPSS for Psychologists: A Guide to Data Analysis using SPSS for Windows.* Third Edition. New Jersey: Lawrence Erlbaum Associates.

Brown, C.H., Berndt, D., Brinales, J.M., Zong, X. and Bhagwat, D. (2000).Evaluating the evidence of effectiveness for preventive interventions: using a registry system to influence policy through science. *Addicitive Behaviors, 25(6),* 955-964.

Cappella, E., Frazier, S.L., Atkins, M.S., Schoenwald, S.K., Glison, C. (2008). Enhancing schools' capacity to support children in poverty: An ecological model of school-based mental health services. *Administration and Policy in Mental Health, 35*, 395-409.

Chinman, M., Gordon, H., Wandersman, A., Ebener, P., Hunter, S., Imm, P., Sheldon, J. (2005). Developing a Community Science Research Agenda for Building Community Capacity for Effective Preventive Interventions. *American Journal of Community Psychology, 35 (3-4)*, 143-157.

Coatsworth, J. D., Duncan, L. G., Pantin, H., & Szapocznik, J.(2006).Retaining ethnic minority parents in a preventive intervention: The quality of group process. *Journal of Primary Prevention, 27*, 367–389.

Cohen, J. (1988). *Statistical power analysis for behavioural sciences.2nd edition*. New Jersey: Lawrence Erlbaum Associates, Hillsdale.

Cohen, J. (1992). A power primer.*Psychological bulletin, 112 (1)*, 155-159.

Crea, T.M., Usher, C.L., Wildfire, J.B. (2009). Implementation fidelity of Team Decision making. *Children and Youth Services Review, 31*, 119-224.

Cross, W., West, J.C. (2011).Examining implementer fidelity: Conceptualizing and measuring adherence and competence. *Journal of Child Services, 6 (1)*, 18-33.

Dane, A.V., Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychological Review, 18 (1)*, 23-45.

Dariotis, J.K., Bumbarger, B. K., Duncan, L.G., Greenberg, T.M. (2008). How do implementation efforts relate to program adherence? Examining the role of organizational, implementer, and program factors. *Journal of Community Psychology, 36 (6)*, 744-760.

Derzon, J., Sale, E., Springer, J., Brounstein, P. (2005). Estimating Intervention Effectiveness: Synthetic Projection of Field Evaluation Results. *Journal of Primary Prevention, 26 (4),* 321-343.

DeVellis, R. F. (2003). *Scale Development: Theory and Applications. Second Edition*. Applied Social Methods research Series, 26. London: Sage Publications.

Domitrovich, C.E., Bradshaw, C.P., Poduska, J.M., Hoagwood, K., Buckley, J.A., Olin, S., Hunter Romanelli, L., Leaf, J. P., Greenberg, M.T., Ialongo, S.N. (2008). Maximizing the Implementation Quality of Evidence-Based Preventive Interventions Schools: A Conceptual Framework. *Advances in School Mental Health Promotion, 1(3),* 6-28.

Domitrovich, C.E., Gest, S.D., Jones, D., Gill, S., Sanford DeRousie, R.M. (2010). Implementation Quality: Lessons Learned in the Context of Head Start REDI Trial. *Early Childhood Research Quarterly, 25,* 284-298.

Domitrovich, C.E., Greenberg, M.T. (2000). The Study of Implementation: Current findings from Effective Programs that Prevent Mental Disorders in School Aged Children. *Journal of Educational and Psychological Consultation, 11(2),* 193-221.

Durlak, A.J., DuPre, E.P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation in Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology,41,* 327-350.

Durlak, J. A. (1998). Why program implementation is important. *Journal of Prevention & Intervention in the Community,17 (2),* 5-18.

Durlak, J.A. (2010). The importance of doing well in whatever you do: A commentary on the special section, Implementation research in early childhood education. *Early Childhood Research Quarterly, 25,* 348-357.

Durlak, J.A., Weissberg, R.P., Dymnicki, B.A., Taylor, R.D., Schellinger, K.B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school based universal interventions. *Child Development, 82 (1),* 474-501.

Dusenbury, L., Brannigan, R., Hansen, W.B., Walsh, J., Falco, M. (2005). Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research Theory and Practice*, *20 (3),* 308-313.

Ferron, J., Hess, M.R., Hogarty, K.Y., Dedrick, R.F., Kromrey, J.D., Lang, T.R., Niles, J. (2004). *Hierarchical Linear Modeling: A Review of Methodological Issues and Applications*. Paper presented at the 2004 annual meeting of the American Educational Research Association.

Fixsen, D. L., Blase, K.A., Naoom, S.F., Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, *19,* 531-540.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). *Implementation Research: A Synthesis of the Literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Flaspohler, P., Duffy, J., Wandersman, A., Stillman, L., Maras, M.A. (2008). Unpacking Prevention Capacity: An Intersection of Research-to-practice Models and Community-centered Models. *American Journal of Community Psychology*, *41,* 182-196.

Flaspohler, P.D., Meehan, C., Maras, M.A., Keller, K.E. (2012). Ready, Willing and Able: Developing a Support System to Promote Implementation of School-Based Prevention Programs. *American Journal of Community Psychology*, 428-444.

Flay, B.R., Biglan, A., Boruch, R.F., Gonzalez Castro, F., Gottfredson, D., Kellam, S., Moscicki, E.K., Schinke, S., Valentine, J.C., Galavotti, J. (2005). Standards of

Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*, *6 (3)*, 151-175.

Galavotti, C., Sebert Kulman, A., Kraft, J.M., Harford, N., Petraglia, J. (2008). From Inovation to Implementation: The Long and Winding Road. *American Journal of Community Psychology*, *41*, 314-326.

Gingiss, P.M., Roberts-Gray, C., Boerm, M. (2006). Bridge-it: A System for Predicting Implementation Fidelity for School-Based Tobacco Prevention Programs. *Prevention Science*, *7*, 197-207.

Gottfredson, D.C., Gottfredson, G.D. (2002). Quality of school-based prevention programs: results from a national survey. *Journal of Research of Crime and Delinquency*, *39*, 3-25.

Graczyk, P.A., Domitrovich, C. E., Small, M., Zins, J. E. (2006).Serving all Children: An Implementation Model Framework. *School Psychology Review*, *35 (2)*, 266-274.

Greenberg, M.T. (2004). Current and Future Challenges in School Based Prevention: A Researchers Perspective. *Prevention Science*, *5 (1)*, 5-13.

Greenberg, M.T., Domitrovich, C. E., Grazyk, P.A., Zins, J.E. (2005). *The Study of Implementation in School Based Preventive Interventions: Theory Research and Practice*. Rockville: Substance Abuse and Mental Health Services Administration.

Greenberg, M.T., Domitrovich, C.E., Bumbarger, B. (1999). *Preventing mental disorders in school aged children: a review of effectiveness of prevention programs*. State College: Prevention Research Center for the Promotion of Human Development, Pennsylvania State University.

Greenhalgh, T., Robert, G., MacFarlane, F., Bate, P., Kyriakidou, O. (2004). Diffusion of Innovation in Service Organizations: Systematic Review and Recommendations. *Milbank Quarterly*, *82 (4)*, 581-629.

Gregory, A., Henry, D., Schoeny, M.E. (2007). School Climate and Implementation of Preventive Intervention. *American Journal of Psychology*, *40*, 250-260.

Hallfors, D., Godette, D. (2002). Will the 'Principles of Effectiveness' improve prevention practice? Early findings from diffusion study. *Health Education Research Theory and Practice*, *17 (4)*, 461-470.

Han, S.S., Weiss, B. (2005). Sustainability of Teacher Implementation of School-Based Mental Health Programs. *Journal of Abnormal Child Psychology*, *33 (6)*, 665-679.

Hawkins, J.D., Catalano, R.F. and Arthur, M.W. (2002).Promoting science-based prevention in communities. *Addictive Behaviors, 27 (6)*, 951-976.

Hosman, C., Jané -Llopis, E., Saxena, S. (Eds) (2005).Prevention *of mental disorders: Effective interventions and policy options*. Oxford Oxford University Press.

Hosman, C.M. (1994). *Effectiveness and effect management in mental health promotion and prevention*.In D. Trent and C. Reed. (Eds) Mental Health Promotion.Aldershot Avebury.

Hosman, C.M. (2008). *Evidence of Effectiveness and Improving Effectiveness in Prevention and Mental Health Promotion*.Textbook ekstboek Inleiding Preventieve Geestelijke Gezondheidszorg.

Hosman, C.M., Engels, C. (1999). The value of model programmes in mental health promotion and mental disorder prevention. *International Journal of Mental Health Promotion*, *1 (2)*, 1-14.

Jané -Llopis, E., Barry, M. M. (2005). What makes mental health promotion effective? *Promotion and Education*, *2*, 47-54.

Jané-Llopis, E., Anderson, P. (2005).*Mental Health Promotion and Mental Disorder Prevention: A Policy for Europe*. Radboud University Nijmegen, Nijmegen.

Janković, J., Bašić, J. (Eds) (2001): *Prevencija poremećaja u ponašanju djece i mladih u lokalnoj zajednici*. Zagreb: Povjerenstvo Vlade Republike Hrvatske za prevenciju poremećaja u ponašanju djece i mladih.

Kam, C-M., Greenberg, T.M., Walls, C.T. (2003).Examining the Role of Implementation Quality in School-Based Prevention Using the PATHS Curriculum. *Prevention Science*, *4 (1)*, 55-63.

Kellam, S.G., Langevin, D. J. (2003). A Framework for Understanding „Evidence" in Prevention Research and Programs. *Prevention Science*, *4 (3)*, 137-153.

Knoche, L., Sheridan, S.M., Edwards, C.P., Osborn, Q.A. (2010). Implementation of a relationship-based school readiness intervention: A multidimensional approach to fidelity measurement for early childhood. *Early Childhood Research Quarterly*, *25*, 299-313.

Kok, G.J., Van den Borne, B., Dolan Mullen, P. (1997). Effectiveness of health education and health promotion: meta-analysis of effect studies and determinants of effectiveness. *Patient Education and Counselling*, *30*, 19-27.

Leykum, L.K., Pugh, J.A., Lanham, H.J., Harmon, J., McDaniel, R.R. (2009). Implementation research design: integrating participatory action research into randomized controlled trials. *Implementation Science*, *4 (69)*; 1-8.

Lillehoj, C.J., Griffin, K.W., Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and Relation of Youth outcomes. *Health Education and Behavior*, *31*, 242-257.

Lochman, J.E., Powell, N.P., Lewczyk Boxmeyer, C., Qu, L., Wells, K.C., Windle, M. (2008). Implementation of a School-Based Prevention Program: Effects of Counselor and School Characteristics. *Professional Psychology: Research and Practice*, *40 (5)*, 476-482.

Meyers, D.C., Durlak, J.A., Wandersman, A. (2012). The quality implementation framework: a synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, *50 (3-4),* 462-480.

Meyers, D.C., Katz, J., Chien, V., Wandersman, A., Scaccia, J.P., Wright, A. (2012). Practical Implementation Science: Developing and Piloting the Quality Implementation Tool. *American Journal of Community Psychology, 50 (3-4),* 481-496.

Midthassel, U.V., Ertesvåg, S. K. (2007). Schools implementing Zero: The process of implementing an anti-bullying program in six Norwegian compulsory schools. *Journal of Education Change, 9,* 153-172.

Mihalic, S.F., Fagan, A., Irwin, K., Ballard, D., Elliot, D. (2004). *Blueprints for violence prevention*. Office of Juvenile Justice and Delinquency Prevention, University of Colorado, Boulder.

Mihalic, S.F., Fagan, A.A., Argamaso, S. (2008). Implementing the LifeSkills Training drug prevention program: factors related to implementation fidelity. *Implementation Science*, *18,* 3-5.

Molleman, G. (2005). *Preffi 2.0:Health Promotion Effect Management Instrument. Development, validity, reliability and usability*. Woerden: Netherlands Institute for Health Promotion and Disease Prevention.

Molleman, G.R.M., Peters, L.W.H., Hosman, C.M.H., Kok, G. (2005). Implementation of a Quality Assurance Instrument to Improve the Effectiveness of Health Promotion in Netherlands.*Health Education Research Theory and Practice*, *20 (4),* 410-422.

Molleman, G.R.M., Peters, L.W.H., Hosman, C.M.H., Kok, G., Oosterveld, P. (2006). Project quality rating by experts and practitioners: experience with Preffi 2.0 as a quality assessment instrument. *Health Education Research*, *21 (2),* 219-229.

Mrazek, P., Haggerty, R.J., Eds. Committee on Prevention of Mental Disorders, Institute of Medicine. (1994). *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Washington, D.C National Academy Press.

Nacionalna strategija prevencije poremećaja u ponašanju djece i mladih od 2009. do 2012. godine. (2009). Zagreb: Povjerenstvo Vlade Republike Hrvatske za prevenciju poremećaja u ponašanju djece i mladih, Vlada Republike Hrvatske.

O' Connel, M.E., Boat, T., Warrner, K.E. (Eds) (2009).*Preventing Mental, Emotional and Behavioral Disorders among Young People: Progress and Possibilities. Part 3: New Frontiers*. Institute of Medicine, National Academy of Sciences.

Odom, S.L., Fleming, K., Diamond, K., Lieber, J., Hanson, M., Butera, G., Horn, E., Palmer, S., Marquis, J. (2010). Examining different forms of implementation in early childhood curriculum research. *Early Childhood Research Quarterly, 25*, 314-328.

O'Donnell, C. L. (2008). Defining, conceptualizing and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78 (1)*, 33-84.

Peters, L., Molleman, G., Hommels, L., Ploeg, M., Hosman, C., Llopis, E. (2003). *Explanatory Guide Preffi 2.0*. Woerden: Netherlands Institute for Health Promotion and Disease Prevention.

Prochaska, J.O., Redding, C.A., Evers, K.E. (2002).*The transtheoretical model and stagesof change.*In Glanz, K., Rimer, B.K., Lewis, F.M. (Eds.) Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass, Inc.

Proctor, E.K., Rosen, A. (2008). From knowledge production to implementation: research challenges and imperatives. *Research of social work practice, 18*, 285-291.

Proctor, E.K., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., Hensley, M. (2010). Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 65-76.

Raphael, B. (1999). The Value of model programmes in mental health promotion and mental disorder prevention. *Journal of Public Mental Health*, *1(2)*, 20-21.

Ringwalt, C. L., Pankratz, M., Jackson-Newsom, J., Gottfredson, N., Hansen, W. B., Giles, S. and colleagues (2010). Three-year trajectory of teachers' fidelity to a drug prevention curriculum. *Prevention Science*, *11*, 67–76.

Roberts-Gray, C., Gingiss, P.M., Boerm, M. (2007).Evaluating school capacity to implement new programs. *Evaluation and Program Planning*, *30*, 247-257.

Rogers, E.M. (1974). Communication in Development. *The Annals of American Academy of Political and Social Science*, *412*, 44-54.

Rogers, E. M. (1976). New product adoption and diffusion. *Journal of Consumer Research*, *2*, 290-301.

Rogers, E.M. (1995). *Diffusion of Innovation (5th edition)*. New York: The Free Press.

Rohrbach, L. A., Dent, C.W., Skara, S., Sun, P., Sussman, S. (2007). Fidelity of Implementation in Project Towards No Drug Abuse (TND): A Comparison of Classroom Teachers and Program Specialists. *Prevention Science*, *8*, 125-132.

Rohrbach, L.A., Gunning, M., Sun, P., Sussman, S. (2010). The Project Towards No Drug Abuse (TND) Dissemination Trial: Implementation Fidelity and Immediate Outcomes. *Prevention Science*, *11 (1)*, 77-88.

Saunders, R.P., Ward, D., Felton, M.G., Dowda, M., Pate, R. R. (2006).Examining the link between program implementation and behavior outcomes in the lifestyle education for activity program (LEAP).*Evaluation and Program Planning, 29,* 352-364.

Schoenwald, S.K., Garland, A.F., Chapman, J.E., Frazier, S.L., Sheidow, A.J., Southam-Gerow, M.A. (2010). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research, 38,* 32-43.

Shaw, I.E., Green, J.C. and Melvin, M.M. (Eds) (2007). *Handbook of Evaluation – Policies, Programs and Practices*. London: Sage Publications.

Singer, D.J. (1998). Using SAS PROC MIXED to fit Multilevel Models, Hierarchical Models, and Individual Growth Models. *Journal of Education and Behavioral Statistics, 23 (4),* 323-355.

Speller, V., Evans, D. and Head, M.J. (1997).Developing quality assurance standards for health promotion practice in UK. *Health Promotion International, 12 (3),* 215-224.

Spoth, R., Guyll, M., Trudeau, L., Goldberg-Lillehoj, C. (2002). Two studies of proximal outcomes and implementation quality of universal preventive interventions in a community-university collaboration context. *Journal of Community Psychology, 30,* 499-518.

Stice, E., Shaw, H., Bohon, C., Nathan Marti, C., Rohde, P. (2009). A meta-analytic review of depression prevention programs for children and adolescents: factors that predict magnitude of intervention effects. *Journal of Consulting and Clinical Psychology, 77 (3),* 486-503.

Stith, S., Pruitt, I., Dees, J., Fronce, M., Green, N., Som, A., Linkh, D. (2006). Implementing community based programming: A review of literature. *The Journal of Primary Prevention, 27 (6),* 599-617.

Tobler, N.S., Stratton, H.H. (1997). Effectiveness of school-based drug prevention programs: A meta-analysis of the research. *Journal of Primary Prevention*, *18*, 71-12.

Tones, K., Tilford, S. (2001). *Health Promotion: Effectiveness, efficiency and equity (3rd edition)*. Nelson Thornes Ltd, Cheltenham.

Wandersman, A. (2003). Community Science: Bridging the Gap between Science and Practice with Community- Centered Models. *American Journal of Community Psychology*, *31(3-4)*, 227-242.

Wandersman, A. (2009). Four keys to success (theory, implementation, evaluation and resources/system support): high hopes and challenges in implementation. American *Journal of Community Psychology*, *43*, 3-21.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., Clachman, M., Dunville, R., Saul, J. (2008). Bridging the Gap Between Prevention Research and Practice: The Interactive Systems Framework for Dissemination and Implementation. *American Journal of Community Psychology*, *41*, 171-181.

Woltman, H., Feldstein, A., McKay, C., Rocchi, M. (2012).An introduction to hierarchical linear modelling. *Tutorials in Quantitative Methods in Psychology, 8 (1)*, 52-69.

APPENDICES

LIST:

APPENDIX 1. ***Table A1. Set of parallel items representing implementation factors for different types of informants***

APPENDIX 2. ***Table A2. Set of parallel items representing indicators of implementation quality for different types of informants***

APPENDIX 3. ***Table A3. Research plan for implementation quality assessment per each of 24 included programs in experimental and control conditions***

APPENDIX 4. ***Table A4. Number of participants from each program for mid-intervention and post-intervention implementation assessment***

APPENDIX 5. ***Tables with descriptive statistics of implementation factors and indicators of implementation quality for managers, implementers and participants***

APPENDIX 6.***Preffi 2.0 instrument***

APPENDIX 1

Table A1. Set of parallel items representing implementation factors for different types of informants

Table A1. Set of parallel items representing implementation factors for different types of informants

CHARACTERISTICS OF A PROGRAM	
PROGRAM STANDARDIZATION	
ITEMS FOR PROGRAM MANAGERS	**ITEMS FOR PROGRAM IMPLEMENTERS**
Program deliverers are provided with specific guidelines for program delivery.	I am provided with specific guidelines for program delivery.
Program deliverers are provided with a course of themes for this program which have to be touched on.	I am provided with a course of themes for this program which have to be touched on.
There are activities specific for each program theme in our program.	I know which activities are specific for each program theme in our program.
Program deliverers are told to follow a schedule of themes.	I keep up with themes according to the schedule.
Program deliverers are expected to keep up with set of themes without making changes.	I am expected to keep up with set of themes without making changes.
Program deliverers are told to conduct program in the same way for all participants.	I conduct program in the same way for all participants.
Program deliverers are told that only small changes should be made to the program.	If I make changes to the program, that should be only small changes.
	There is a detailed written description of the preventive program conducted by our organization. YES NO I use structured written materials in the program implementation. YES NO Our program has a handbook for implementation of activities. YES NO
PROGRAM IMPLEMENTER'S SKILL	
ITEMS FOR PROGRAM MANAGERS	
The program deliverer is skilled at delivering this program.	
Program deliverer keeps most of participants active and engaged.	
Program deliverer is prepared for the program sessions/ meetings/activities. Program deliverer is conducting core components of our preventive intervention.	

ATTITUDES TOWARDS THE INTERVENTION	
ITEMS FOR PROGRAM MANAGERS	**ITEMS FOR PROGRAM IMPLEMENTERS**
This intervention makes a real difference in the lives of participants.	This intervention makes a real difference in the lives of participants.
I am afraid that program effects are short-lived and fade out quickly after the program ends.	I am afraid that program effects are short-lived and fade out quickly after the program ends.
Our intervention meets the needs of participants in sufficient manner.	Our intervention meets the needs of participants in sufficient manner.
This program needs to include more meetings or workshops with participants to be effective.	This program needs to include more meetings or workshops with participants to be effective.
This program needs to cover more themes to have more impact.	This program needs to cover more themes to have more impact. This program could be more effective if it cover more themes.
Our program affects behaviour or attitudes of program participants.	
	I like this program very much.
	This program is a good model for dealing with this problem.
	I think there are changes that could be made to improve this program.
	Most of the activities in whole program were in place and adequate.
	This program only scratches the surface of the problem it is designed to prevent.
	The activities in this program are comprehensive.
ORGANIZATION CAPACITY	
TRAINING AND INTERVENTION KNOWLEDGE	
ITEMS FOR PROGRAM MANAGERS	**ITEMS FOR PROGRAM IMPLEMENTERS**
I as organization manager invest enough in the development of skills and knowledge the program deliverer needs for program implementation.	My organization invests enough in the development of skills and knowledge I need for program implementation.
I organize in-service trainings where program deliverers practice skills needed for program implementation.	My organization provides in-service trainings which give me the possibility to practice skills needed for program implementation.
I send our program deliverers to different seminars and trainings which can benefit our program implementation.	My organization sends me to different seminars and trainings which can benefit my program implementation.
My program deliverer is well prepared with respect to the demands of this program.	I feel prepared to deliver the intervention.
Program deliverer training is covering skills and knowledge needed for program implementation.	The training I was provided gave me the skills and knowledge needed for program implementation.
SUPPORT TO PROGRAM IMPLEMENTER	
ITEMS FOR PROGRAM MANAGERS	**ITEMS FOR PROGRAM IMPLEMENTERS**
When deliverer encounters difficulties in program implementation, I am	When I am insecure about program implementation, I can consult the

available to provide advice.	manager of the organization.
I provide sufficient administrative and technical support for program deliverers throughout program implementation.	Management of my organization provides me with sufficient administrative and technical support throughout program implementation.
I provide enough emotional support to program deliverers through different phases of program implementation.	When needed in different phases of program implementation, I can get enough emotional support from my superiors.
Program deliverer is included in supervision arranged by our organization or similar human services.	I am included in supervision of my work where I can talk about experiences and problems connected with program implementation.
When a problem in implementation arises, I as an organization manager work collaboratively with program deliverer.	When a problem in implementation arises, organization manager works with me collaboratively.
Program deliverers perceive me as supporting and someone he/she can rely on.	I perceive organization manager as a person of trust I can rely on. Organization manager possesses skills needed for quality management.
Program has enough financial resources and support.	In my view the program has enough financial resources and support.
MONITORING	
ITEMS FOR ORGANIZATION MANAGERS	**ITEMS FOR PROGRAM IMPLEMENTERS**
I regularly communicate with program deliverer regarding the program implementation.	I regularly communicate with organization manager to share the information about the program implementation.
I am along with phases of program delivery and I know what is happening on the field.	Organization manager is along with phases of program delivery and knows what is happening on the field.
Program deliverer sends me in written feedback about the program implementation regularly.	I regularly send written feedback to organization manager about the program implementation.
I regularly hold meetings with program deliverer to talk about important steps in the process of program implementation.	I regularly hold meetings with my organization manager to talk about important steps in the process of program implementation.
Our organization has a structured employee appraisal form to assess deliverers working quality.	Quality of my work is assessed by structured employee appraisal system which is used in our organization.
I come to the field and watch my staff delivering the program.	Program manager comes to the field and watches me delivering the program.
Someone in our organization observes program deliverers conducting the program.	Someone in our organization observes me while conducting the program.

APPENDIX 2

Table A2. Set of parallel items representing indicators of implementation quality for different types of informants

Table A2. Set of parallel items representing indicators of implementation quality for different types of informants

FIDELITY	
	ITEMS FOR PROGRAM IMPLEMENTERS
	I know what the core components of the program are.
	I deliver program activities as planned.
	I think that it is o.k. to leave out some activities as long as they are not core elements of the program.
	I need to make changes to this program to meet the needs of participants.
QUALITY OF PROGRAM DELIVERY	
ITEMS FOR PROGRAM PARTICIPANTS	**ITEMS FOR PROGRAM IMPLEMENTERS**
Program deliverer is skillful in program implementation. *Child version: Program deliverer is doing a good job.*	I think that I am skilled at delivering this program.
We sometimes ran out of time during activities/workshops/meetings.	In activities/workshops/meetings we often run out of time.
Program deliverer gives us a feedback about the way we have conducted certain activity or exercise. *Child version: Program deliverer talks with us kids about the way we have done certain activity.*	I give feedback to the participants about the way they have conducted certain activity or exercise.
Program deliverer seemed underprepared.	I am prepared for the program sessions/meetings/activities.
Program deliverer represents activities in highly engaging manner. *Child version: Program deliverer represents activities cheerfully.*	During sessions I am able to keep most participants active and engaged in the program. I assure active participation of all the participants during the program (discussion, opinion expression).
I perceive the rhythm of program implementation as adequate.	
If needed, program deliverer repeats some program activities for participants. *Child version: Program deliverer repeats some program activities if I ask him to.*	
I like the working style of program deliverer. *Child version: I like program deliverer.*	
It is evident that program deliverer is positive towards the program and that he/she believes in its impact.	
If I have some questions, I can talk to program deliverer. *Child version: If I have some questions, I can talk to program deliverer.*	
Program deliverer is doing a good job and I trust him.	
PARTICIPANT'S RESPONSIVENESS	

ITEMS FOR PROGRAM PARTICIPANTS	ITEMS FOR PROGRAM IMPLEMENTERS
I am interested in themes presented in this program. *Child version: This program is interesting to me.*	Participants are interested in themes presented in this program.
In general, I stay engaged during the whole meeting/workshop.	In general, participants stay engaged during the whole meeting/workshop.
I am highly collaborative during meetings/workshops/activities.	Participants are highly collaborative during meetings/workshops/activities.
If I get homework or assignment on the meeting, I fulfil it. *Child version: If I get homework or assignment on the meeting, I usually fulfil it.*	If I give some homework or assignment to participants, they fulfill it.
Atmosphere on the meetings/workshops/activities is positive. *Child version: I have a lot of fun during this program.*	Atmosphere on the meetings/workshops/activities is positive.
I perceive others from the group as supportive.	Participants from the group are supportive to each other.
I seek additional materials and sources of information about specific program themes.	Participants seek additional materials and sources of information about specific program themes.
There are activities in this program that I refuse to participate in.	There are activities in this program that participants refuse.
Program deliverer likes some of the participants more than others. *Child version: Program deliverer likes some of the participants more than others.*	I like some of the participants more than others.
During the activities conduction, program deliverer assures active participation of all the participants (discussion, opinion expression). *Child version: During the program, deliverer asks me what I think.*	
I am bored in this program. *Child version: I am bored in this program.*	Participants are bored in this program.
I feel excited when going to workshop/activities of this program. *Child version: I feel excited when going to workshop/activities of this program.*	Participants are excited when going to workshop/activities of this program.
I like this program very much. *Child version: I like this program.*	
This program is a good model for dealing with this problem. *Child version: Activities in this program are good.*	
This program has fulfilled my expectations.	
I meet interesting people because of this program.	
I think there are changes that could be made to improve this program.	
This program needs to include more meetings with us participants to be effective. *Child version: It would be great if this program lasted longer.*	

This program could be more effective if it cover more themes.	
Most of the activities in whole program were in place and adequate.	
This program only scratches the surface of the problem it is designed to prevent.	
Program activities are comprehensive.	
When I think about all activities/workshops/meetings which were implemented until now, I was present at: <20% 40% 60% 80% 100% *Child version: How many workshops/meetings have you attended?*	When you think about all activities/workshops/meetings which were implemented until now, how many participants have attended regularly: < 25% 25%-50% 50% 50-75%>75%
	How many participants attend each workshop/meeting on average? Very few Less than half about half more than half almost all
DOSAGE	
ITEMS FOR PROGRAM PARTICIPANTS	
How many workshops/meetings were there until now? *Child version: How many workshops/meetings were there until now?*	
PERCEIVED PROGRAM IMPACT	
ITEMS FOR PROGRAM PARTICIPANTS	**ITEMS FOR PROGRAM IMPLEMENTERS**
I have gained from experience of other participants.	Participants gain from learning about the experiences of other participants.
I was changing behaviour in different phases of this program. *Child version: I have improved my behaviour because of this program.*	Participants are changing behaviour in different phases of this program.
This program has helped me in my functioning. *Child version: This program has helped me.*	This program has helped participants in their functioning.
This program helped me to learn something important and relevant to my life. *Child version: This program taught me something important.*	This program has helped participants to learn something important and relevant to their lives.
This program has improved my relationships with others.	This program has improved participant's relationships with others.
I have a feeling that I have gained after each workshop/meeting. *Child version: On each program meeting/activity, I learn something new.*	
I am afraid that program effects are short-lived and fade out quickly after the program ends.	
This program met my needs.	
I think about some themes of this program in my everyday life.	
I will change something in my behavior in the future because of this program.	

APPENDIX 3

Table A3. Research plan for implementation quality assessment per each of 24 included programs in experimental and control conditions

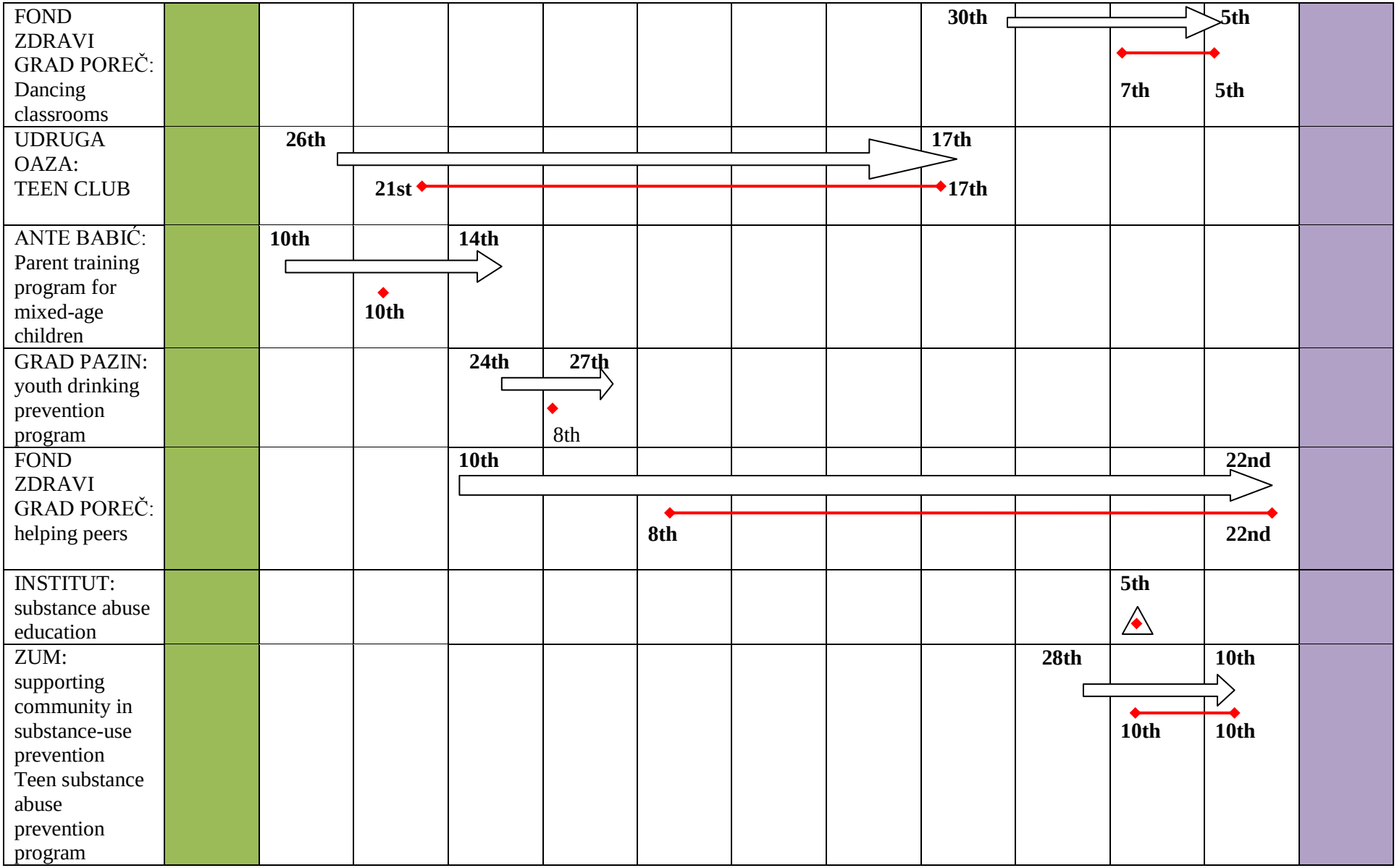
Table A3. Research plan for implementation quality assessment per each of 24 included programs in experimental and control conditions

Legend: Preffi 2.0 baseline assessment of all programs, Training for Prevention intervention, Preffi 2.0 post-test of all programs

Duration of interventions Only one lecture in the program Dates of 1st and 2nd implementation quality measurement

NAME OF THE ORGANIZATION AND PROGRAM	JAN 2011	FEB 2011	MAR 2011	APR 2011	MAY 2011	JUN 2011	JUL 2011	AUG 2011	SEP 2011	OCT 2011	NOV 2011	DEC 2011	JAN 2012
ZIID TEATAR: Teen theatre workshop		19th		16th					28th				
OBITELJSKI CENTAR: Mentor program*				11th	9th			16th					
LABIN ZDRAVI GRAD: Parent training				20th 27th	6th 6th								
DRUŠTVO PSIHOLOGA: Media literacy program										17-25th 25th	24th 28th	6th 9th 2-19th 7th-21st	
SUNCOKRET: Program for academic support			9-30th 15-30th										
ZAVOD ZJZ: substance abuse education, parent-based											1st		

ZAVOD ZJZ: substance abuse education, teacher-based					7th							
DND PULA: Parent training	9th		14th									
		2nd	14th									
UDRUGA LABIN ZDRAVI GRAD: Health promotion program for self-confidence training				3rd	6th							
				23rd	6th							
					6th	13th						
					23rd	13th						
OBITELJSKI CENTAR: Structured free time health promotion program using art techniques						12th	11th					
						19th	11th					
GRAD BUZET: Parent training program for mixed age of children		10th	7th									
		3rd	7th									
OBITELJSKI CENTAR: Parent training program for pre-school children		24th	5th									
		10th	5th									


FOND ZDRAVI GRAD POREČ: Dancing classrooms
30th
5th
7th
5th
UDRUGA OAZA: TEEN CLUB
26th
17th
21st
17th
ANTE BABIĆ: Parent training program for mixed-age children
10th
14th
10th
GRAD PAZIN: youth drinking prevention program
24th
27th
8th
FOND ZDRAVI GRAD POREČ: helping peers
10th
22nd
8th
22nd
INSTITUT: substance abuse education
5th
ZUM: supporting community in substance-use prevention Teen substance abuse prevention program
28th
10th
10th
10th

GRAD PAZIN: Parent training programs, risk and universal					20th risk	17th		18th 18th		10th uni 17th	21st 21st		
DND PAZIN: Let's grow up together			15th	5th	17th 17th								
ART STUDIO: Parent-child art classes								22nd	7th	13th 13th			
ODISEJA: Successful parenting Parent training for parents of elementary school children		8th 22nd	1st 14th 14th 15th 15th	4th 4th 18th 1 - 18th						13th 20th	17th 17th 10th 24th	15th 15th	
ASANDO CHER: Parent training for Roma parents			8th 15th	12th 12th							8th 15th	20th 20th	

APPENDIX 4

Table A4. Number of participants from each program for mid-intervention and post-intervention implementation assessment

Table A4. Number of participants from each program for mid-intervention and post- intervention implementation assessment

NAME OF THE ORGANIZATION AND PROGRAM	NUMBER OF ORGANIZATION MANAGERS	NUMBER OF PROGRAM DELIVERERS	NUMBER OF PROGRAM PARTICIPANTS
ZIID TEATAR: Teen theatre workshop	1	1	12 in first measurement of IQ 11 in post measurement of IQ
OBITELJSKI CENTAR: Mentor program	1	1	3 in first measurement of IQ 3 in post measurement of IQ
LABIN ZDRAVI GRAD: Parent training	1	2	5 in first measurement of IQ 4 in post measurement of IQ
DRUŠTVO PSIHOLOGA: Media literacy program	1	3	139 in post measurement of IQ
SUNCOKRET: Program for academic support	1	3	10 in first measurement of IQ 9 in post measurement of IQ
ZAVOD ZJZ: substance abuse education, parent-based	1	1	29 in post measurement of IQ
ZAVOD ZJZ: substance abuse education, teacher-based	1	1	63 in post measurement of IQ
DND PULA: Parent training	2	1	23 in first measurement of IQ 21 in post measurement of IQ
UDRUGA LABIN ZDRAVI GRAD: Health promotion program for self-confidence training	1	4	30 in first measurement of IQ 32 in post measurement of IQ
OBITELJSKI CENTAR: Structured free time health promotion program using art techniques	1	3	19 in first measurement of IQ 19 in post measurement of IQ
GRAD BUZET: Parent training program for mixed age of children	1	1	12 in first measurement of IQ 12 in post measurement of IQ
OBITELJSKI CENTAR: Parent training program for pre-school children	1	1	8 in first measurement of IQ 6 in post measurement of IQ

FOND ZDRAVI GRAD POREČ: Dancing classrooms	1	2	113 in first measurement of IQ 115 in post measurement of IQ
UDRUGA OAZA: TEEN CLUB	1	4	8 in first measurement of IQ 8 in post measurement of IQ
ANTE BABIĆ: Parent training program for mixed-age children	1	1	10 in first measurement of IQ 1 in post measurement of IQ
GRAD PAZIN: youth drinking prevention program	1	1	45 in post measurement of IQ
FOND ZDRAVI GRAD POREČ: helping peers	1	1	27 in first measurement of IQ 11 in post measurement of IQ
INSTITUT: substance abuse education	1	2	50 in post measurement of IQ
ZUM: supporting community in substance-use prevention Teen substance abuse prevention program	1	2	8 in first measurement of IQ 7 in post measurement of IQ
GRAD PAZIN: Parent training programs, risk and universal	1	2	Universal: 8 in first measurement of IQ 8 in post measurement of IQ Risk: 7 in first measurement of IQ 11 in post measurement of IQ
DND PAZIN: Let's grow up together	1	2	52 in first measurement of IQ 58 in post measurement of IQ
ART STUDIO: Parent-child art classes	1	1	10 in first measurement of IQ 10 in post measurement of IQ
ODISEJA: Successful parenting Parent training for parents of elementary school children	1	10	50 in first measurement of IQ 47 in post measurement of IQ
ASANDO CHER: Parent training for Roma parents	1	2	13 in first measurement of IQ 10 in post measurement of IQ

APPENDIX 5

Tables with descriptive statistics of implementation factors and indicators of implementation quality for managers, implementers and participants

Table A5.1 Descriptive statistics of implementation factors from mid-intervention assessment and post-assessment for managers and implementers

Variable	M	SD	Min	Max	Skewness	Kurtosis
MANAGERS						
standardization						
mid	2.55	0.65	1	3.33	-0.83	0.29
post	2.46	0.77	1	3.83	-0.55	0.23
implementer skills						
mid	3.42	0.48	2.33	4	0.08	-1.46
post	3.68	0.43	3	4	-0.82	-1.18
attitudes						
mid	2.75	0.40	2.17	3.67	0.21	-0.64
post	2.86	0.42	2.17	3.67	0.14	-0.11
training						
mid	2.25	0.58	1	4	0.78	2.18
post	2.65	0.67	1.75	4	0.57	-0.33
support						
mid	3.24	0.51	2.17	4	-0.42	0.87
post	3.36	0.53	2.50	4	-0.22	-1.37
monitoring						
mid	2.64	0.43	1.83	4	0.52	1.02
post	2.75	0.66	1.83	4	0.59	-0.74

Variable	M	SD	Min	Max	Skewness	Kurtosis
IMPLEMENTERS						
standardization						
mid	2.87	0.57	1.50	3.83	-0.49	-0.40
post	3.01	0.57	1.50	4	-0.38	-0.24
attitudes						
mid	2.93	0.38	2.18	3.34	-0.13	-0.73
post	2.96	0.38	2.09	3.73	-0.03	-0.43
training						
mid	2.88	0.64	1.50	4	0.09	-0.75
post	2.99	0.64	1.50	4	-0.22	-0.41
support						
mid	3.42	0.51	2.14	4	-0.74	-0.37
post	3.46	0.57	1.43	4	-1.39	2.16
monitoring						
mid	2.98	0.60	2	4	0.01	-1.15
post	2.90	0.59	1.20	4	-0.45	0.51

Table A5.2 Descriptive statistics of indicators of implementation quality from mid-intervention assessment and post-assessment

Variable	M	SD	Min	Max	Skewness	Kurtosis
IMPLEMENTERS						
fidelity						
mid	3.11	0.97	2.25	4	-0.02	0.09
post	3.26	0.38	2.75	4	0.18	-0.89
quality of delivery						
mid	3.68	0.35	2.83	4	-0.29	-1.19
post	3.53	0.38	2.83	4	-0.38	-1.16
responsiveness						
mid	3.55	0.25	2.63	4	-0.64	-0.47
post	3.42	0.39	2.63	4	-0.42	-0.77
program impact						
mid	3.46	0.37	2.75	4	0.73	-0.18
post	3.31	0.47	2.50	4	0.39	-1.28
PARTICIPANTS						
dosage						
mid	4.17	2.87	1	14	-0.39	-0.48
post	5.99	7.81	1	48	3.84	16.72

quality of delivery						
mid	3.65	0.26	1.17	4	-1.77	5.89
post	3.52	0.45	1	4	-1.68	4.32
responsiveness						
mid	3.46	0.52	1	4	-1.73	3.99
post	3.28	0.59	1	4	-1.24	1.86
program impact						
mid	3.20	0.59	1	4	-0.98	1.01
post	3.08	0.70	1	4	-0.76	0.21

APPENDIX 6

Preffi 2.0 instrument

PREFFI 2.0

OPERATIONALISATION AND NORMS

The purpose of this document is to help you fill out the Preffi 2.0 evaluation sheet. We recommend that you first read the user manual in the Preffi 2.0 assessment package.

The document includes one or more "yes" or "no" questions for each Preffi criterion (unit). Based on the answers to those questions, the criterion (unit) can be categorized as „weak", „moderate" or „strong".

Some questions may be difficult to answer, especially if the project plan does not provide enough information or if you yourself lack knowledge related to certain fields of expertise. In any case, you should answer as many questions as possible. The answer „not available" is offered in a limited number of criteria, usually in the cases when the criteria are difficult to put in words in project descriptions or when they are not obvious to those who are not themselves included in project implementation (for example, „competence and characteristics of the project manager" and „adjusting to the culture"). If some criteria allow "not available" as an answer, this will be explicitly noted.

The document provides space for comments on every criterion; for example, your comments on why you answered a certain question with "yes" or "no". You may also specify and describe aspects you believe need improvement, and you can also transfer this to the Answer Sheet.

Criteria in the document are listed in the same order as in the Evaluation sheet. So the document starts with Problem analysis, cluster 2 and ends with Contextual conditions and feasibility, cluster 1. The User manual explains the rationale behind such an order.

Cluster 2, Problem analysis

2.1. Nature, severity and scope of the problem

Operationalisation:

1. Is the problem or the topic clear?
2. Is it clear whether the problem or the topic is frequent within the group or community?
   Additional questions:
   Is the prevalence of the problem known (=number of existing cases)?
   Is the incidence of the problem known (=number of new cases in a certain period of time)?
3. Is the interrelatedness of health and social problems clear? This includes indicators like rate of unemployment, income, fear of crime, racial discrimination, drug addiction, number of welfare cases and housing conditions.
4. Is what is known about immaterial costs of the problem clearly stated – such as mortality (mortality rate, life expectancy), diseases and disorders, limitations, disabilities, harmful impact, medicine use and absence from work?
5. Is what is known about material costs clearly stated – such as cost of services, health care costs, measurement costs, loss of revenue due to attempts to solve or contain the problem.

Norms:

- Weak: question 1 = no and /or question 2 = no and/or question 3 = no
- Moderate: questions 1 - 3 = yes and question 4 and/or 5 = no or not available
- Strong: questions 1 - 3 = yes and questions 4 and/or 5 = yes

2.2. Distribution of the problem

Operationalisation:

1. Is it clear how the problem is distributed regarding:
   - age?
   - sex?
   - socio-economic status?
   - ethnical background?
   - religious background?
   - cultural or subcultural origin?
   - time (seasons, days of the week, hours of the day)?
2. Is anything known about the geographical distribution of the problem, in terms of a certain region, city or area? (For example, the unusually high mortality from cancer in a certain region; traffic accidents on certain intersections; fear of crime in certain streets or buildings, etc.)
3. Are data available for a specific target area at which the project is aimed (designed for the whole country or a province, region, city, town district)? If not, has data been correctly extrapolated from general data?

Note: each question enumerates many points of interest, but not all of these need to be of importance for every project situation.

Norms:

- Weak: question 1 = no
- Moderate: question 1 = yes and question 2 = no and question 3 = no
- Strong: question 1 = yes and question 2 = yes and/or question 3 = yes

2.3. Perception of the problem by key people

Operationalisation:

1. Is it known to what extent the problem is actually perceived by the target group as a problem?
2. Has it been established which individuals, groups, agencies and parts of the social sector are involved in the process of tackling or solving problems?
3. Has it been established to what extent these individuals, groups, agencies and parts of the social sector agree about the source and cause of the problem?
4. Has it been established how major social subgroups, such as ethnic or cultural groups, men and women or different types of schools, perceive the problem?
5. Has it been checked whether politicians and the public opinion are interested in or pressure for certain steps to be taken for solving the problem?

Norms:

- Weak: question 1=no and/or question 2=no
- Moderate: at least question 1 = yes and question 2 = yes (with the possibility that the answer to some of the remaining questions is also yes)
- Strong: at least question 1 = yes and question 2 = yes and two more questions = yes

Note: In this cluster the 'target group' always implies the final target group to which the project refers to.

Cluster 3. Determinants of (psychological) problems, behaviour and environment

3.1. Theoretical model

Operationalisation:

1. Have the theoretical assumptions or the model used for explaining the (psychological) problem, risk and desired behaviour or environmental factors been clearly stated?
2. Has it been clearly shown that the selected model is most suitable for approaching these (psychological) problems, behaviour or environmental factor (for example, because the model has been specifically developed for a specific problem, behaviour or environmental factors, because the model has already been successfully applied or it has been discussed in a scientific journal or because its applicability can be supported by theoretical arguments)?
3. Has it been clearly described how factors affect each other, how they affect behaviour, environmental factors and/or the problem – favourably or unfavourably?

Norms:

- Weak: questions 1 and/or 2 = no
- Moderate: 1 = yes, 2 = yes, 3 = no
- Strong: all questions = yes

3.2. Contribution of determinants to psychological problems, behaviour or environmental factors

Operationalisation:

1. Is it known which determinants influence desired and undesired behaviour, environmental factors or the (psychological) problem (on a personal level, on the level of social environment and psychological environment)?
2. Is it clear which determinants are the most important?

3. Is it clear how reliable is the evidence of determinants?
4. Is it clear to what extent determinants can be applied to relevant subgroups (e.g. according to age, sex, ethnicity, religion, etc.)

Norms:

- ❖ Weak: question 1 = no (making the other questions irrelevant)
- ❖ Moderate: 1 = yes and at the most one more question = yes
- ❖ Strong: question 1 = yes and at least two more questions = yes

3.3. Susceptibility of determinants to change

Operationalisation:

1. Has it been estimated to what extent determinants are susceptible to change in the described situation (on the level of an individual and on the level of social and physical environment)?
2. Has this estimate been based on theoretical and/or scientific knowledge about the variability of determinants? (Suggestion: consult relevant literature, co-workers or experts, conduct preliminary testing)

Norms:

- ❖ Weak: question 1 = no (making the second question irrelevant)
- ❖ Moderate: question 1 = yes, question 2 = no
- ❖ Strong: question 1 = yes and question 2 = yes

3.4. Priorities and selection

Operationalisation:

1. Have the target behavioural or environmental factors or (psychological) problems been specified?
2. Has it been explained to which health problem(s) or life quality problem(s) these factors are related?
3. How have the target determinants for behavioural or environmental factors or (psychological) problems been explained?
4. Have the groups in risk and/or target groups been mentioned and specified?

Norms:

- Weak: maximum of two questions = yes
- Moderate: three questions = yes
- Strong: all questions = yes

Cluster 4. Target group

Note: In this cluster, the expression 'target group' always refers to the final target group.

4.1. General and demographic characteristics of the target group

Operationalisation:

Suggestion: Much of the data collected during problem analysis is also probably relevant in this chapter.

1. Is it clear which general and demographic characteristics are relevant for this specific project? An affirmative answer requires that at least the first five characteristics from the following list apply:
   - the size of the target group
   - age
   - sex
   - socio-economic status (level of education, income, profession, work status)
   - ethnical background
   - cultural background
   - religious background
   - marital status, housing conditions
   - number of family members
   - geographic position
   - language (spoken and written), illiteracy
2. Are concrete figures available about relevant characteristics of the target group in this project?

Norms:

- Weak: both questions = no
- Moderate: question 1 = yes, question 2=no
- Strong: both questions = yes

4.2. Motivation and options of the target group

Operationalisation:

1. Is it known to what extent the target group is ready for change?
2. Is it known which factors influence the motivation of members of the target group to change? (These can include awareness of the problem, attitude, self-efficiency, obstacles, etc. Suggestion: see also cluster 3)
3. Is it known for the purpose of this specific project, which desires, needs, limitations and obstacles for change the group is aware of?

Norms:

- Weak: question 1 = no, regardless of the answers to questions 2 and 3
- Moderate: question 1 = yes and question 2 or 3 = no
- Strong: all questions = yes

4.3. Accessibility of the target group

Operationalisation:

1. Is it clear by what means the target group can be covered? (Suggestion: think about locations, media, intermediary persons)
2. Is the selection of the means (locations, media, intermediaries) corroborated by the project?

Norms:

- ❖ Weak: both questions = no
- ❖ Moderate: question 1 = yes and question 2 = no
- ❖ Strong: both questions = yes

Cluster 5. Objectives

5.1. Objectives are adjusted to the analysis

Operationalisation:

1. Does the objective description make a clear distinction between different objective levels? The levels may refer to health/life quality, behaviour/environment/problems and determinant's level as well as the level of objectives for creating preconditions.
2. Do the objectives adjust and are they in accordance with the analysis conducted in previous clusters? (see clusters 2 and 3)

Norms:

- Weak: both questions = no
- Moderate: one question = yes and one question = no
- Strong: both questions = yes

5.2. Objectives are specific, time-limited and measurable

Operationalisation:

1. Do objectives specify factors that need to be changed? (Suggestion: This question has been analysed in 5.1)
2. Has for the objectives a target group been specified in which these objectives need to be achieved?
3. Do objectives specify the desired magnitude of effects that wants to be achieved (e.g.: 10% decrease)?
4. Do objectives specify the time period in which they need to be realised?

Norms:

- Weak: questions 1. and/or 2.=no
- Moderate: question 1=yes, question 2=yes, 3. question=no, 4. question=no
- Strong: question 1=yes, question 2=yes and questions 3 and/or 4=yes

5.3. Objectives are acceptable

Operationalisation:

1. Are the project theme and the set objectives in accordance with the objectives of your organisation?
2. Are the intervention objectives acceptable (or can they become acceptable) for financing/to the evaluation board or maybe to the medical ethical board/institutional board for evaluation?
3. Are the objectives of the intervention acceptable (or can they become acceptable) to possible partners and implementers?
4. Are the intervention objectives acceptable (or can they become acceptable) to the target group?

We are aware that questions related to this criterion are not easy to answer. If you can answer them use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- Weak: at least one negative answer in questions 1-3
- Moderate: questions 1 - 3 = yes
- Strong: questions 1 – 4 = yes

5.4. Objectives are achievable

Operationalisation:

1. Has the necessary personnel, money and time for achieving the set objectives been estimated? (Suggestion: data from criterion 3.3 can be useful here)
2. Is there a sufficient number of available experts, competent persons and partners for achieving the set objectives?

We are aware that questions related to this criterion are not easy to answer. If you can answer them use the given norm. If you cannot give an answer, mark "not available" on the Answers Sheet.

Norms:

- Weak: question 1 = no, regardless of the answer to question 2
- Moderate: question 1 = yes and question 2 = no
- Strong: both questions = yes

Cluster 6. Intervention development

Note: If the project includes more interventions, you can answer the questions in general. However, if you are interested in assessing each specific intervention, it is possible to answer each question separately. (see User Manual, section 3.3)

6.1. Rationale for the intervention strategy

6.1.a. Adjusting the strategies and methods to objectives and target groups

Operationalisation:

1. Have the intervention methods been specified?
2. Has it been established how intervention methods are appropriate and adequate for achieving the desired objectives (e.g. through research or theoretical considerations)?

Norms:

- Weak: both questions = no
- Moderate: question 1 = yes and question 2 = no
- Strong: both questions = yes

6.1.b. Previous experience with intervention(s)

Operationalisation:

1. Are any reports available about a successful or unsuccessful use of this intervention by someone else (in literature or through other experts)?
2. Do you as an individual have some experience of a successful or unsuccessful application of intervention?
3. Does the suggested method seem potentially efficient for this specific situation? (Suggestion: you have to consider the extent to which your situation can be compared to other situations where some experience has already been gathered, especially concerning objective terms/determinants, themes/problems, target groups and contextual conditions)

Norms:

- Weak: question 1 = no and question 2 = no
- Moderate: question 1 and/or 2 = yes and question 3 = no
- Strong: question 1 and/or 2 = yes and question 3 = yes

6.2. Duration, intensity and chronology

6.2.a Duration and intensity of the intervention

Operationalisation

1. Are some research data or practical experiences available about the duration and intensity in which the intervention should be implemented in order to achieve the set objectives?
2. Has this data been used in deciding on the optimum duration and intensity of the proposed intervention?

Norms:

- ❖ Weak: question 1 = no (making the answer to the second question irrelevant)
- ❖ Moderate: question 1 = yes and question 2 = no
- ❖ Strong: both questions = yes

6.2.b. Intervention chronology

Operationalisation:

1. Has it been established whether target groups react better to the intervention in a certain time of year? (This can for example refer to education about sunbathing and skin cancer during summer. Religious and state holidays should be taken into account or periods like Ramadan).
2. Has it been established whether the chronology of the intervention is compatible with specific important experiences of target group members? (This can for example refer to care interventions for mental health in crisis situations or to the level of sexual experience of young people included in the AIDS prevention program).
3. Has it been established to what extent the time period of the intervention agrees with the age or development stage of the target group? (This can for example refer to the information that interventions for preventing aggression with children are most effective if they are conducted when the children are 3 or 4 years of age).
4. In the case when the intervention is to be implemented with the help of intermediary persons, has the chronology of the intervention been adjusted to these persons?

Norms:

- ❖ Weak: maximum one question = yes
- ❖ Moderate: maximum two questions = yes
- ❖ Strong: at least three questions = yes

6.3. Adjusting to the target group

6.3.a. Participation of the target group

Operationalisation:

1. In the case when the intervention has been developed somewhere else (for example, on the national level): has the general target group been at least consulted during intervention development?
2. For any project: has the specific target group (e.g. residents of a target district) for the ongoing project at least been consulted during intervention development or before selecting the intervention model?
3. For any project: regarding the project's characteristics, has the target group been sufficiently involved in development and intervention selection?

Norms:

- Weak: question 1 = no or not available and question 2 = no (making the third question irrelevant)
- Moderate: questions 1 and/or 2 = yes, question 3 = no
- Strong: question 1 and/or 2 = yes, question 3 = yes

6.3.b. Adjusting to 'culture'

Note: The term 'culture' is used in the broadest sense; it can include adapting to age, sex, socio-economic status, etc. For example, it might be necessary to address young people differently than adults and older people.

Operationalisation:

1. Is the content (message) in accordance with knowledge, views, customs, roles and capacities of members of the cultural or subcultural groups?
2. Is the way of reaching members of the cultural or subcultural group adequate and does it adequately convey the messages? Is the medium for communication frequently used and attractive?
3. Is the target group accessible to the source or message transmitter (e.g. intermediary)?

4. Has the source or message transmitter shown proof of sufficient understanding and knowledge about culturally determined customs and social norms of the target group?
5. Does the target group perceive the intervention as being in accordance with their culture?

We are aware that questions related to this criterion are not easy to answer. If you can answer them use the given norm. If you cannot give an answer, mark "not available" on the Evaluation sheet.

Norms:

- ❖ Weak: question 1, 2, 3 and/or 4 = no
- ❖ Moderate: questions 1 - 4 = yes and 5 question = no
- ❖ Strong: all questions = yes

6.4. Effective techniques

Operationalisation:

1. Have the following techniques been used in the project, considering the importance they have for the project to be assessed?

Effective techniques

- a room for personalised approach
- feedback (about the situation in the target group, behaviour or intervention effects)
- use of rewarding strategies
- removal of obstacles towards the desired behaviour
- mobilising social support/commitments, involving the social environment
- training skills

- ensuring follow-up
- defining objectives and implementation intentions
- interactive approach

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- ❖ Weak: none or few of the effective strategies have been used
- ❖ Moderate: some effective strategies have been used
- ❖ Strong: many effective strategies have been used

6.5. Feasibility in existing practice

6.5.a. Adjusting to the intermediary target group

Operationalisation:

1. Have the members of the intermediary target group been consulted during the development process of the intervention (for the final target group)?
2. Is the intervention in accordance with ways of operating, procedures, standards and values of intermediaries and their organisation?

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- ❖ Weak: both questions = no
- ❖ Moderate: one of two questions = no
- ❖ Strong: both questions = yes

6.5.b. Intervention/s feasibility characteristics

Operationalisation:
The following questions can be answered for every intermediary group separately:

1. Has it been established/recorded to what extent intermediaries feel that the use/implementation of the intervention will improve their current practice?
2. Has it been established/recorded to what extent intermediaries feel that the new intervention is in accordance with the current procedure?
3. Has it been established/recorded to what extent intermediaries possess the necessary skills for implementing the intervention?
4. Has it been established/recorded whether the intervention procedure is clear to the intermediaries, i.e. whether they know what is expected of them?
5. Has it been established/recorded whether the intermediaries think that the new intervention gives them enough space for experimenting? Can intermediaries test the intervention without being strictly bound by the content of the intervention?
6. Has it been established/recorded whether the intermediaries feel they can immediately notice the effects of the intervention?
7. Has it been established/recorded to what extent intermediaries feel the intervention to be affordable?

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- Weak: 0 - 2 questions = yes
- Moderate: 3 - 5 questions = yes
- Strong: 6-7 questions = yes

6.6. Coordination between intervention/activity

Operationalisation:

1. Is the program comprehensive enough to reach the set objectives? In other words, does it make sufficient use of available segments of intervention methods, ways and determinants of the target group?
2. If the program/project includes multiple interventions (segments of intervention methods, ways and determinants of the target group), are these different interventions coordinated in a satisfying manner?

Norms:

- ❖ Weak: both questions = no
- ❖ Moderate: one of two questions = yes
- ❖ Strong: both questions = yes

6.7. Preliminary testing

Operationalisation:

1. Has preliminary testing been used?
2. Have conclusions been made and steps taken in accordance, in terms of communication and/or effects, based on preliminary testing? In other words, has the intervention been adjusted where necessary?

Norms:

- ❖ Weak: question 1 = no (making the second question irrelevant)
- ❖ Moderate: question 1 = yes and question 2 = no
- ❖ Strong: both questions = yes

Cluster 7. Implementation

7.1. The selection of the implementation strategy adjusted to intermediaries

7.1.a. Implementation model: top down and/or bottom-up

Operationalisation:

1. Have certain implementation models been selected deliberately?
2. Do intermediaries have the chance of adjusting the intervention to their own situation?
3. If intermediaries have the chance of adjusting the intervention, is it clear which parts of the intervention need to be preserved?

Norms:

- Weak: question 1 = no or question 1 = yes and question 2 = no (making the third question irrelevant)
- Moderate: question 1 = yes, question 2 = yes and question 3 = no
- Strong: all questions = yes

7.1.b. Adjusting intervention implementation to intermediaries

Operationalisation:

1. Is it clear how members of the intermediary group are distributed during different expansion and innovation application phases (awareness of innovation; decision to apply the innovation; reporting the innovation; continued innovation application)
2. Have specific objectives for each expansion and innovation application phase been set, for every segment of the intermediary or target group?
3. Do the implementation interventions fit in with the objectives that have been set for each stage of diffusion and use and for each intermediary target group or target groups segment?
4. Are the set objectives realistic considering the fact that the intermediary group can be divided into 'innovators', 'early adopters', 'early majority', 'late majority' and 'laggards'?

Norms:

- ❖ Weak: question 1 = no and question 2 = no
- ❖ Moderate: question 1and/or 2 = yes
- ❖ Strong: at least questions 1-3 = yes

7.1.c. Appropriateness of the supplier for intermediaries

Operationalisation:

1. Is it known whether the planned supplier is appropriate in the eyes of the intermediary target group? Aspects of appropriateness include:

- support/commitment
- authority
- competence
- image
- the size of the supplier agency
- position within the network
- financial capacity and other available resources

2. Are different contact persons used, when appropriate, for different segments of the intermediary target group?

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- ❖ Weak: question 1 = no (characteristics have not been taken into consideration)
- ❖ Moderate: question 1 = yes (some characteristics have been take into consideration) and question 2 = no (different persons were looked for but were not included)
- ❖ Strong: question 1 = yes (some characteristics have been take into consideration) and question 2 = yes or it is irrelevant

7.2. Monitoring and generating feedback

Operationalisation:

1. Has it been specified in how many points of time the expansion progress and intervention implementation will be assessed, e.g. by collecting feedback from intermediaries and the final target group?
2. Does the assessment lead to an active adjustment of the expansion process and intervention implementation?

Norms:

- Weak: both questions = no
- Moderate: question 1 = yes and question 2 = no
- Strong: both questions = yes

7.3. Incorporation into existing structures

Operationalisation:

1. Has the intervention been incorporated into the existing structure?
2. Has an effort been made, or is it made right now, to fit the intervention into already existing structures?
3. Are these activities and attempts strong enough, i.e. are they aimed at the right hierarchical level? (e.g. it is easier to influence business people through other business people)

Norms:

- Weak: question 1 = no and question 2 = no (making the third question irrelevant)
- Moderate: question 1 = no, question 2 = yes and question 3 = no
- Strong: question 1 = yes or question 2 = yes and question 3 = yes

Cluster 8. Evaluation

Note: In the case of the project including more interventions and/or evaluations, the questions can be answered generally. However, if you are interested in evaluating each specific intervention or evaluation, it is possible to provide answers for each intervention or evaluation separately, for example with the help of the matrix. (see User Manual, section 3.3)

8.1. Explicitness and agreement on evaluation principles

Operationalisation:

1. Have important individuals, groups and/or organisations been included in designing the evaluation? This refers to commission organisations, the ones who need to implement the intervention, members of the target group and potential external experts.
2. Do all key people have a clear idea about the questions that the evaluation must answer and do they agree on these questions?
3. Is it clear which form/s of the evaluation is/are necessary in order to answer the questions?
4. Do key people agree about the strength of proof that needs to be obtained through the evaluation and is this level of proof achievable?

Norms:

- Weak: question 2 = no
- Moderate: question 2 = yes and question 1 and/or question 3 = no
- Strong: at least questions 1, 2, 3 = yes

8.2. Process evaluation

Operationalisation:

1. Does the process evaluation allow insight into the degree to which the activities have been implemented according to plan?
2. Does the process evaluation allow insight into user's opinions (final and/or intermediary target group) about activities and materials?
3. Does the process evaluation allow insight into intervention coverage (which people have been included, how representative are they, who was excluded from the intervention and why)?
4. Does the process evaluation allow insight into the degree to which the objectives of creating preconditions for the project have been reached?
5. Does the process evaluation allow insight into possible unpredictable circumstances and side-effects?
6. Does the process evaluation reveal conditions for success?

Norms:

- Weak: maximum three questions = yes
- Moderate: three or four questions = yes
- Strong: at least five questions = yes

8.3. Effect evaluation

Note: We are aware that answering questions in this cluster requires certain professional knowledge about effect evaluation, which can make it more difficult for individuals to answer. It is a problem we are not able to solve at this moment. It is in this sense our goal to offer support through the Internet version of Preffi 2.0. which is to be developed in the future.

8.3.a. Has any change been measured or is being measured at this moment?

Operationalisation:

1. Has it been measured (or is it being measured now) to which degree the objectives of the intervention have been reached (or are reached)? It is necessary to take into account different objectives (emphasized in section 5.1), especially momentary (or intermediate) intervention objectives. This will mostly not include end objectives of the intervention on a public health level since their realisation requires a longer period.
2. Are the used measuring methods valid and reliable? This concerns questions referring to outcome measures, measuring methods, measuring instruments and the size and representative quality of the sample/group that is being studied.

Norms:

- Weak: both questions = no
- Moderate: question 1 = yes and question 2 = no or not available
- Strong: both questions = yes

8.3.b. Is it likely that the change was caused by intervention?

Operationalisation:

1. Is it clear which of the alternative explanations for noticed changes can be excluded (out of the six possible explanations listed in the criteria in the manual)? Special attention should be directed towards information about study design and use of multiple measuring methods and multiple sources (e.g. results of the process evaluation and effect evaluation) and to the degree of agreement between their findings.
2. Is the level of credibility of the made conclusions justified by the level of security offered by the study design? Conclusions must be aligned with the measure in which alternative explanations cannot be excluded.

We are aware that questions related to this criterion are not easy to answer. If you can answer them use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- ❖ Weak: both questions = no
- ❖ Moderate: question 1 = yes and question 2 = no
- ❖ Strong: both questions = yes

8.4. Feedback to key people

Operationalisation:

1. Have key people been introduced to important feedback acquired in the evaluation process? This includes the following aspects:

- Do the findings agree with problems noticed and/or questions asked by key people?
- Does the provided information include aspects the key people have the power to change? (Can this information be used to derive some policy recommendations?)
- Have any side-effects been clearly shown?
- Are the proposed measures acceptable to key people?
- Will findings be available within a reasonable time?

2. Is the manner of presenting the findings adjusted to key people (in terms of readability and conciseness)?

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- Weak: question 1 = no
- Moderate: question 1 = yes and question 2 = no
- Strong: both questions = yes

Cluster 1. Contextual conditions and feasibility

1.1. Support/commitment

Operationalisation:

1. Has it been established which internal and external partners are necessary for ensuring adequate support and commitment during every phase of the project?
2. Is there sufficient support and commitment among the required partners?
3. Have agreements been made and confirmed about involving internal and external collaborators in the project?

We are aware that questions related to this criterion are not easy to answer. If you can answer them use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- Weak: question 1 = yes or no, question 2 = no and question 3 = no
- Moderate: question 1 = yes, question 2 = yes and question 3 = no
- Strong: all questions = yes

1.2. Capacity

Operationalisation:

1. Have available resources for the project been established?
2. Are the available resources in line with the objectives of the project?
3. Have the available resources in every phase of the project been used in the most efficient way?

Norms:

- Weak: all questions = no OR question 1 = yes or no, question 2 = no and question 3 = no or 'not available'
- Moderate: question 1 = yes, question 2 = yes and question 3 = no or 'not available'
- Strong: all questions = yes

1.3. Management by the project manager

1.3.1. Expertise and characteristics of the project manager

Operationalisation:

1. Is only one person responsible for the project?
2. Does the person with exclusive responsibility have the necessary competence for implementing the project?
3. Is the work style of the person with exclusive responsibility for the project compatible with the specific phase and peculiarities of the project?
4. Does the person with exclusive responsibility for the project have appropriate personal characteristics for implementing the project?

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- Weak: question 1 = no or question 2 = yes and at least one more question yes
- Moderate: question 1 = yes and two questions out of 2, 3 and 4 = yes
- Strong: all questions = yes

1.3.b. Key points for management

Operationalisation:

1. Is the project being implemented in accordance with the project plan which includes clear moments for making decisions?
2. Is the communication plan being actively implemented?
3. Has the project manager the opportunity to use available resources in a flexible manner?
4. Is the project manager ensuring that his/her competence and the competence of other staff involved is up-to-date by organising additional training, intervision, etc.?

We are aware that questions related to this criterion are not easy to answer. If you can answer them, use the given norm. If you cannot give an answer, mark "not available" on the Answer Sheet.

Norms:

- Weak: question 1 = no; or question 2 = yes and at least one more question = yes
- Moderate: question 1 = yes and two of the remaining questions = yes
- Strong: all questions = yes

BIOGRAPHY OF THE CANDIDATE

Miranda Novak was born on 8 May 1981 in Čakovec where she attended elementary school and high school. She studied at the Faculty of Philosophy from 1999-2005 obtaining a degree in Psychology. She enrolled in postgraduate Prevention Science and Disabilities Studies, specialization „Prevention Science: prevention of mental and behavioural disorders and mental health promotion" at the Faculty of Education and Rehabilitation Sciences, University of Zagreb, in 2007/2008. Professionally she is interested in the prevention of mental and behavioural disorders, socio-emotional and healthy development of children and youth, mental health promotion, community prevention programs, research of prevention program effectiveness, social marketing, internalized disorders and counselling.

Her research interests are connected with field of prevention and mental health. She is working as a research assistant at the Department of Behavioural Disorders at the Faculty of Education and Rehabilitation Sciences, University of Zagreb. She was an associate on several research projects from which the most distinguished ones are the project of the Ministry of Science, Education and Sport, "Communities That Care: development, implementation and evaluation of community prevention", and the project "Implementation of evidence-based prevention program of socio-emotional learning through science evaluation and its application into Croatian kindergartens and primary schools (PATHS-RASTEM)" which was financed by the Unity Through Knowledge Fund, Ministry of Science, Education and Sports and World Bank. She has published six scientific and one expert paper; she was co-author of five chapters in the book and contributed several dozens of abstracts and presentations on domestic and international conferences.

She got additional training within the field of counselling and psychotherapy and she is working as a counsellor with children, youth and families. She is a member of the Croatian Psychological Association, Croatian Psychological Chamber, Society for Prevention Research, European Network for Socio-Emotional Competencies and European Society for Prevention.

LIST OF PUBLISHED PAPERS:

1. Novak, M. i Bašić, J. (2008). Internalizirani problemi kod djece i adolescenata: obilježja i mogućnosti prevencije. *Ljetopis socijalnog rada, 15(3),* 473-498.

2. Novak, M. (2008). Neki pokazatelji psihičkog zdravlja nezaposlenih mladih. *Kriminologija i socijalna integracija. 16 (2),* 61-73.

3. Bašić, J., Novak, M., Grozić-Živolić, S. (2008). Percepcija rizika i potreba zajednice: percepcija građana i ključnih ljudi u istarskoj županiji. *Kriminologija i socijalna integracija, 16 (2),* 73-85.

4. Bašić, J., Novak, M., Grozić-Živolić, S. (2008). Spremnost i mobilizacija zajednice za prevenciju rizičnih ponašanja djece i mladih: perspektiva ključnih ljudi u Istarskoj županiji. *Kriminologija i socijalna integracija, 16 (2),* 85-97.

5. Mihić, J., Novak, M., Bašić, J., (2010). Zajednice koje brinu: CTC Upitnik za djecu i mlade u procjeni potreba za preventivnim intervencijama. *Ljetopis socijalnog rada, 17 (3),* 391-412.

6. Horvat, M., Kolačko, D., Novak, M., Bašić, J. (2011). Promišljanja o etičkim dilemama u pripremi studenata socijalne pedagogije za djelovanje u praksi. *Kriminologija i socijalna integracija, 19 (2),* 91-102.

7. Bašić, J., Mihić, J., Novak, M. (2011). Risk analysis in the period of growing-up of children and youth: starting point for effective prevention. *Journal of Public Health, 19 (1),* 3-11.