

Wie wirkt langfristig?

Entwicklung eines Forschungsdesigns für eine langfristige Wirkungsanalyse am Beispiel des Mentoring-Programmes ‚Nightingale‘ Wien.

How to Measure the Long-term Outcome?

Development of a Research Design for a Long-term Impact Assessment using the Mentoring Program ‘Nightingale’ Vienna.

Masterarbeit

Zur Erlangung des akademischen Grades

Master of Arts in Social Sciences

der Fachhochschule Campus Wien

Masterstudiengang: Sozialwirtschaft und Soziale Arbeit

Vorgelegt von:

Verena Kofler, BA

Personenkennzeichen:

1510600034

Erstbegutachterin:

FH-Prof.ⁱⁿ Mag.^a Florentina Astleithner

Zweitbegutachter:

Prof. Dr. Herbert Bassarak

Eingereicht am:

31.08.2017

Erklärung:

Ich erkläre, dass die vorliegende Masterarbeit von mir selbst verfasst wurde und ich keine anderen als die angeführten Behelfe verwendet bzw. mich auch sonst keiner unerlaubter Hilfe bedient habe.

Ich versichere, dass ich diese Masterarbeit bisher weder im In- noch im Ausland (einer Beurteilerin / einem Beurteiler zur Begutachtung) in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Weiters versichere ich, dass die von mir eingereichten Exemplare (ausgedruckt und elektronisch) identisch sind.

Datum: 31.08.2017

Unterschrift:

Danksagung

An dieser Stelle möchte ich zuerst dem Projektleiter von ‚Nightingale‘ Wien, Herrn Mag. Günther Leeb danken, dank dessen Interesse und Unterstützung diese Arbeit erst entstehen konnte. Auch bedanke ich mich bei allen InterviewpartnerInnen für die Bereitschaft und die Zeit, die sie sich genommen haben. Ebenfalls möchte ich mich bei Frau Mag.^a Florentina Astleithner für die Betreuung der Masterarbeit und ihre Anregungen bedanken. Abschließend möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Masterarbeit unterstützt haben, allen voran Mag. Raphael Asada, BA und Clemens Loidl BA BA BA.

Kurzfassung

Soziale Programme sehen sich mehr und mehr gezwungen mit ihrer Wirkung beschäftigen und diese nachweisen zu müssen. Dabei wird die Langfristigkeit häufig vernachlässigt. In der vorliegenden Arbeit wurde der Frage nachgegangen, wie kurz- und langfristige Outcomes im Mentoring-Programm ‚Nightingale‘ nachgewiesen werden können. Dazu wurden zwei Wirkungsziele betrachtet: die Steigerung der Bildungsmotivation und die Stärkung des Selbstwertes der TeilnehmerInnen. Im Zuge dieser Arbeit wurde ein Forschungsdesign für eine langfristige Wirkungsanalyse entwickelt, ohne dieses durchzuführen. Im Quasi-Experiment-Design anhand der *Double-Difference*-Methode und mit einer Vergleichsgruppe können kurz- und langfristige Veränderungen (Outcomes) aufgezeigt und auf das Programm zurückgeführt werden. Hierbei handelt es sich um ein nicht-erinnerungsbasiertes Verfahren, das sowohl intendierte als auch nicht-intendierte persönlichkeitsorientierte Outcomes berücksichtigt. Mit dem ‚Fünf-Faktoren-Modell‘ (Bleidorn und Ostendorf 2009) und der ‚Selbstwertkala‘ (Schütz und Selling 2006) wurden dabei etablierte Instrumente gewählt, die den wissenschaftlichen Gütekriterien entsprechen. Das Forschungsdesign wurde anschließend mittels qualitativer teilstrukturierter ExpertInnen-Interviews mit relevanten AkteurInnen auf Akzeptanz und Umsetzbarkeit diskutiert. Zusätzlich wurden die Interviews mit ausgewählten Daten der aktuellen Evaluierung von ‚Nightingale‘ ergänzt. Sowohl Akzeptanz als auch Umsetzbarkeit des Forschungsdesigns wurden hierbei (bedingt) bestätigt. Die Arbeit zeigt Möglichkeiten und Grenzen der Planung und hypothetischen Umsetzung einer langfristigen Wirkungsanalyse im Bereich der Sozialwirtschaft auf.

Abstract

Social programs have to deal increasingly with their effects and are required to prove them. The aspect of long-term effects is often neglected. The present study investigated the question of how short- and long-term outcomes can be measured by an impact assessment on the mentoring program 'Nightingale', focusing on two of its main objectives: the increase of the educational motivation and the strengthening of self-esteem. A research design has been developed, which is able to show short and long-term changes (outcomes) using a quasi-experiment design and the *double-difference method*. By involving a comparison group, it allows the outcomes to be causally linked to the program. This non-memory-based approach takes into account both intended and non-intended personality-oriented outcomes. With the 'five-factor model' (Bleidorn and Ostendorf 2009) and the 'self-value scale' (Schütz and Selling 2006), established instruments were chosen that correspond to the scientific main quality criteria. This research design has then been discussed with relevant stakeholders on acceptance and practicability, based on qualitative, partially structured expert interviews. Information resulting from the interviews were extended with selected data deriving from analysis of current evaluations of 'Nightingale'. Thereby both, the acceptance and the feasibility of the research design, could (conditionally) be confirmed. The work presents the possibilities and limitations regarding planning and hypothetical implementation of a long-term impact analysis in the field of social economy.

Abkürzungsverzeichnis

AHS	Allgemeinbildende Höhere Schule
EBP	Evidence Based Practice
ECEB-Mentoring	European Center of Evidence-Based Mentoring
FFM	Fünf-Faktoren-Modell
HiPIC	Hierarchical Personality Inventory for Children
KMS	Kooperative Mittelschule
MGT	Measurement Guidance Toolkit
MSWS	Multidimensionale Selbstwertkala
NMS	Neue Mittelschule
NPM	New Public Management
RE	Realistic Evaluation
SROI	Social Return of Investment
SRS	Social Report Standard
SWS	Selbstwertkala
WFA	Wirkungsorientierte Folgenabschätzung

Schlüsselbegriffe

Langfristigkeit

Wirkungsanalyse

Mentoring-Programm

Wirkungsmessung

Evidence-Based-Practice

Bildungsmotivation

Selbstwert

Inhaltsverzeichnis

Danksagung	I
Kurzfassung	II
Abstract	III
Abkürzungsverzeichnis	IV
Schlüsselbegriffe	V
1 Einleitung	1
1.1 The ‚Nightingale‘ Mentoring Network	3
1.1.1 ‚Nightingale‘ Wien.....	4
1.1.2 Ziele und Evaluierung – ‚Nightingale‘ Wien.....	6
2 Forschungsstand	9
2.1 Theorie und Implementierung der Wirkungsorientierung in der Sozialwirtschaft ...	9
2.1.1 Was ist Wirkung? Begriffserklärung, Verortung und Abgrenzung.....	9
2.1.2 Wirkungsorientierung und -steuerung im Kontext des NPM.....	16
2.1.3 Wirkungsorientierung in der Sozialwirtschaft und Sozialen Arbeit	17
2.2 Erkenntnisse der empirischen Wirkungsforschung im Mentoring	19
2.2.1 Wirkungsorientierung in Mentoring-Programmen.....	19
2.2.2 Forschungsstand ‚Nightingale‘	22
3 Erkenntnisinteresse und Forschungsfragen	24
4 Methode	27
4.1 Erstellung eines Forschungsdesigns für eine langfristige Wirkungsanalyse	28
4.1.1 Operationalisierung der Ziele nach dem SMART-Ansatz	28
4.2 Datenerhebung durch die qualitative Befragung von AkteurInnen	29
4.3 Zugang zum Feld und Beschreibung des Feldes.....	30
4.3.1 Durchführung Interviews.....	32
4.3.2 Leitfadengestützte ExpertInnen-Interviews	34
4.3.3 Transkription	35
4.3.4 Analyse der Interviews – Qualitative Inhaltsanalyse	35
4.3.5 Kodierung mittels MAXQDA	37
4.3.6 Schwerpunktmäßige Auswertung des Onlinefragebogens von ‚Nightingale‘	37
5 Ergebnisse	39
5.1 Entwicklung des hypothetischen Forschungsdesigns	39

5.1.1	<i>Wirkungsziele als SMART-Ziele</i>	40
5.1.2	<i>Auswahl der Versuchsanordnung und Methode</i>	40
5.1.3	<i>Auswahl der Instrumente</i>	42
5.1.3.1	<i>Bildungsmotivation</i>	42
5.1.3.1.1	<i>Fünf-Faktoren-Modell</i>	42
5.1.3.1.2	<i>Measurement Guidance Toolkit</i>	45
5.1.3.2	<i>Selbstwert</i>	47
5.1.3.2.1	<i>Selbstwertskalen</i>	47
5.1.3.3	<i>Literaturrecherche</i>	48
5.1.4	<i>Darstellung des hypothetischen Forschungsdesigns</i>	48
5.2	<i>Organisation</i>	52
5.2.1	<i>Dauer des Programmes</i>	52
5.2.2	<i>Auswahl der Schulen</i>	53
5.2.3	<i>„Nightingale“ an der Pädagogischen Hochschule</i>	53
5.2.4	<i>Auswahl der Mentees</i>	54
5.2.5	<i>Einverständnis der Eltern</i>	55
5.2.6	<i>MentorInnen und Kooperationen mit Universitäten und Hochschulen</i>	56
5.2.7	<i>Fördergeber und verpflichtende Evaluierung</i>	57
5.2.8	<i>Evaluierung des „Nightingale“-Netzwerkes</i>	58
5.2.9	<i>Schwerpunkt MentorInnen</i>	59
5.2.10	<i>Befragung von Kinder/Mentees</i>	61
5.2.11	<i>Das „Nightingale“-Netzwerk als Mitglied im ECEB-Mentoring</i>	64
5.3	<i>Einflussfaktoren auf die Mentoring-Beziehung</i>	65
5.3.1	<i>Incomes MentorInnen</i>	65
5.3.2	<i>Incomes Mentees</i>	65
5.3.3	<i>Dauer des Programmes</i>	66
5.3.4	<i>Familie und Umgebung</i>	66
5.4	<i>Umsetzbarkeit</i>	67
5.4.1	<i>Vergleichsgruppe</i>	67
5.4.2	<i>Psychologische Instrumente</i>	68
5.4.3	<i>Terminierbarkeit – dritter Messpunkt</i>	70
5.4.4	<i>Akzeptanz</i>	73
6	Diskussion der Ergebnisse	78
6.1	<i>Rekapitulation des Forschungsdesigns</i>	78

6.2	Aktuelle Evaluierung ‚Nightingale‘ Wien	79
6.3	Vergleichbarkeiten im Double-Difference-Ansatz	79
6.4	Bildung von Ziel- und Vergleichsgruppe	81
6.5	Selbst- und Fremdevaluierung	82
6.6	Langfristigkeit	83
6.7	Anwendbarkeit der Instrumente	83
6.7.1	<i>Normierung und Altersbegrenzung</i>	84
6.8	Akzeptanz	86
7	Conclusio und Ausblick	88
	Literaturverzeichnis	91
	Abbildungsverzeichnis	100
	Tabellenverzeichnis	101
	Anhang	102

1 Einleitung

Wirkungen von sozialen Interventionen nachzuweisen wird ein immer bedeutsameres Thema im Bereich der Sozialwirtschaft und der Sozialen Arbeit. Es ist ein Thema, das in seinen Vor- und Nachteilen respektive Chancen und Risiken stark umstritten ist und große Relevanz einnimmt. Dabei ist die Frage, *wie* etwas wirkt oder wirken soll nicht einfach zu beantworten. Im sozialen Gefüge existieren selten eindeutige Kausalmechanismen, die pfadgenau den Einfluss einer sozialen Intervention aufzeigen. Dennoch birgt die Wirkungsorientierung in ihrem Ansatz, Projekte oder Programme auf ihre Wirkung hin zu planen, zu evaluieren und zu verbessern, ein großes Entwicklungspotential für die Sozialwirtschaft und Soziale Arbeit. Die Wirkungsorientierung ist seit 2013 ein fixer Bestandteil im österreichischen Haushaltsgesetz (vgl. Sozialministerium o. J., S. 4)¹ und verpflichtet, konkrete Wirkungsziele zu benennen und deren Erfüllung zu überprüfen. Dies kann künftig auch Einfluss auf die Vergabe von öffentlichen Geldern und auf die Finanzierung von Programmen im sozialwirtschaftlichen Bereich haben. Somit wird Wirkungsorientierung vermehrt zu einem konstitutiven Teil von Projekten und Programmen und es erscheint sinnvoll, sich *vor* einem verpflichtenden Wirkungsnachweis zu überlegen, wie die Wirkung einer sozialen Intervention nachgewiesen werden kann. Ein solches Vorhaben benötigt schließlich beträchtliche Ressourcen und soll aus dem Programm heraus entwickelt werden.

Vor diesem Hintergrund behandelt die vorliegende Arbeit die Frage, wie Wirkungsorientierung im Mentoring-Programm ‚Nightingale‘ zur Stärkung bildungsbenachteiligter Kinder umgesetzt werden kann. ‚Nightingale‘ wurde in Schweden entwickelt und bildet heute ein internationales Netzwerk, das insgesamt sieben Länder umfasst. In Wien wird das Programm seit 2010 ausgeführt. Hierbei werden vor allem bildungsbenachteiligte Kinder (Mentees) im Eins-zu-eins-Tandemsetting mit Studierenden (MentorInnen) zusammengebracht, die als Bildungs-Vorbild fungieren sollen. Die Tandems treffen sich während den sieben Monaten einmal wöchentlich für ca. drei Stunden in der Freizeit. Das Interesse und der Zugang zu ‚Nightingale‘ Wien ist auf die eigene Teilnahme als Mentorin im Jahrgang 2014/15 rückzuführen, wodurch der Fokus auf den Standort Wien zu erklären ist. Ausgehend von den Zielen des Mentoring-Programmes wurden zwei von sechs (Wirkungs-)Zielen genauer betrachtet: die ‚Steigerung der Bildungsmotivation‘ und die ‚Stärkung des Selbstwertes‘. Diese Ziele

¹ Zitiert nach Citavi-Basisstil.

werden als Hauptziele des Programmes angesehen und werden auch in anderen Mentoring-Programmen häufig als elementare Wirkungsziele angegeben. Die Frage lautet dabei nicht, ob die Ziele erreicht wurden, sondern wie eine langfristige Wirkung in diesen Bereichen festgestellt werden kann.

Soziale Interventionen wie Mentoring-Programme zielen häufig auf die Persönlichkeitsveränderung der Zielgruppe, wobei Evaluierungen jedoch oft nur einmal (am Ende des Programmes) durchgeführt werden. Für die Feststellung von Veränderungen bedarf es jedoch mindestens zweier Beobachtungszeitpunkte. Zudem können Persönlichkeitsveränderungen oder vergleichbare langfristige Zielsetzungen nicht immer direkt am Ende der Intervention festgestellt werden. Neben dieser ‚Praxislücke‘ besteht eine Forschungslücke in der Analyse von langfristigen Wirkungen, auf die in dieser Arbeit eingegangen werden soll.

Die Forschungsfrage lautet daher, *wie* langfristige Veränderungen in der Persönlichkeit der teilnehmenden Kinder festgestellt und auf die Teilnahme am Programm zurückgeführt werden können. Dies erfordert die Entwicklung eines Forschungsdesigns für eine Wirkungsanalyse, das im Rahmen dieser Arbeit geleistet wird, ohne die Analyse tatsächlich auszuführen. Vielmehr soll anhand qualitativer ExpertInnen-Interviews mit den OrganisatorInnen und unterschiedlichen AkteurInnen (Mentees, MentorInnen, Eltern und Lehrpersonen) herausgefunden werden, wo die Möglichkeiten und Grenzen in der Planung und der Durchführung einer solchen langfristigen Wirkungsanalyse sind. Ziel dieser Arbeit ist somit nicht, eine Wirkungsanalyse durchzuführen, sondern durch einen explorativen Zugang die Möglichkeiten und Grenzen bei der Planung einer langfristigen Wirkungsanalyse aufzuzeigen.

Dazu wird in Kapitel 1.1 das Mentoring-Netzwerk vorgestellt und konkret auf ‚Nightingale‘ Wien und dessen derzeitiges Evaluierungsverfahren eingegangen. Daraufhin werden in Kapitel 2 der Forschungsstand zur Wirkungsorientierung in Mentoring-Programmen sowie der Forschungsstand mit konkretem Bezug zu ‚Nightingale‘ aufgezeigt. Kapitel 2.1 verweist auf relevante Begriffe der Wirkungsforschung. Zugleich stellt dieses Kapitel Ausschnitte der Diskussion um Wirkungsorientierung in der Sozialwirtschaft und Sozialen Arbeit dar, die relevant für die methodischen Entscheidungen sind. In Kapitel 3 werden die Erkenntnisinteressen dargestellt und die Forschungsfragen ausgeführt. Kapitel 4 behandelt methodische Überlegungen zur Literaturrecherche, zur qualitativen Befragung der AkteurInnen und zur ergänzenden Auswertung von bereits vorliegenden Evaluierungsergebnissen. Kapitel 5 prä-

sentiert zunächst das hypothetische Forschungsdesign für eine langfristige Wirkungsanalyse, und geht dann auf die Ergebnisse der qualitativen Befragung ein, die in Kapitel 6 schließlich diskutiert werden.

1.1 The ‚Nightingale‘ Mentoring Network

Das Mentoring-Programm ‚Nightingale‘ wurde 1997 in Malmö, Schweden, von der Pädagogin und Soziologin Carina Sild Lönroth entwickelt. Das Programm ist nach dem israelischen Vorbild ‚Perach‘ aufgebaut und bringt bildungsbenachteiligte Volksschulkinder (Mentees) und Studierende (MentorInnen) im Eins-zu-eins-Tandemsetting zusammen, um gemeinsam Freizeitaktivitäten zu unternehmen. Der Name ‚Nightingale‘ ist das englische Wort für Nachtigall und entstammt einem Gedicht, das beschreibt, dass die Nachtigall nur dann singt, wenn sie sich sicher und wohlbehütet fühlt (vgl. Lönroth 2007, S. 13).

Dank der Comenius-Förderung der Europäischen Union konnte ‚Nightingale‘ ab 2007 auf mehrere Länder in Europa übertragen werden – mit diesem Jahr wurde das Mentoring-Programm auch erstmals in Österreich, Linz, ausgeführt (vgl. Lönroth 2007, S. 13). Seit 2010 wird das Programm in Österreich durch die Kinderfreunde in der Abteilung ‚Gemeinsam – Initiative Interkulturelle Arbeit‘ organisiert und zusätzlich in den Städten Wien, Salzburg und Graz realisiert. Neben Schweden und Österreich findet ‚Nightingale‘ in Island, Norwegen, Spanien, der Schweiz und in Uganda statt (vgl. The Nightingale Mentoring Network 2017a) und hatte 2013 insgesamt fünfhundert Tandems (vgl. Lönroth 2013, S. 2). Für den Austausch zwischen den Ländern gibt es neben der Homepage (vgl. The Nightingale Mentoring Network 2017c) mit eigenem Log-In für Mitglieder jährliche Konferenzen – zuletzt in Barcelona. Jede/r NetzwerkpartnerIn muss einen Jahresbericht abgeben und quantitative wie qualitative Angaben einreichen (vgl. The Nightingale Mentoring Network 2017a, S. 2). Das ‚Nightingale‘-Konzept ist patentiert und als Marke eingetragen. Da für die Umsetzung eine Gebühr zu bezahlen ist (vgl. The Nightingale Mentoring Network 2017b), kann es als *Social Franchise* angesehen werden. Dennoch steht es ausführenden KoordinatorInnen frei, eigene Zielsetzungen bzw. Schwerpunkte zu setzen. Im Folgenden soll zunächst auf das in Malmö entwickelte Konzept eingegangen werden, um dann die Umsetzung des Konzepts in Wien zu betrachten.

Das Basiskonzept aus Malmö sieht vor, dass Tandems aus Studierenden aller Studienrichtungen und (vorzugsweise) Kindern zwischen 8 und 12 Jahren gebildet werden. Das Programm baut auf Freiwilligkeit auf. Nicht nur die MentorInnen nehmen freiwillig teil, auch

die Kinder (respektive Eltern) entscheiden, ob sie mitmachen möchten. Ein Großteil der teilnehmenden Kinder besucht „multicultural schools which are situated in a varied socioeconomic area“ (The Nightingale Mentoring Network 2017d, S. 1). In diesen Schulen steht es allen Kindern frei, sich für die Teilnahme am Programm anzumelden. Für die Auswahl der MentorInnen sieht das Konzept Bewerbungsgespräche vor, in denen ein Strafregisterauszug abgegeben werden muss. Alle MentorInnen nehmen an einem verpflichtenden Training teil. Zudem sind sie veranlasst, monatliche Berichte über ihre Tätigkeiten mit dem Mentee abzugeben und an einer Evaluierung teilzunehmen. Für die Mitarbeit erhalten die MentorInnen ein Budget zur Finanzierung der Aktivitäten sowie Studienpunkte (ECTS oder Vergleichbares) und ein Abschlusszertifikat. Um ein passendes *Matching* zu ermöglichen wird auch mit den Kindern ein persönliches Gespräch geführt. Weiter sieht das Konzept vor, dass Mädchen ausschließlich mit Mentorinnen Tandems bilden (vgl. The Nightingale Mentoring Network 2017d, S. 1). In den Vertragsunterlagen finden sich Ziele für Mentees und MentorInnen. Mentees sollen durch die Möglichkeit, ein/e StudentIn als Vorbild an die Seite gestellt zu bekommen, Selbstvertrauen entwickeln und in einer engen Beziehung zur/zum MentorIn neue Erfahrungen machen sowie neues Wissen erlangen. Der Austausch zwischen Kindern und Studierenden soll zu erhöhtem gegenseitigem Verständnis, Respekt und Toleranz führen. Zudem ist es ein langfristiges Ziel „that the child will develop more positively and they will be more likely to apply to university studies when the time comes“ (The Nightingale Mentoring Network 2017d, S. 3).

1.1.1 ‚Nightingale‘ Wien

Während die Projektentwicklerin Carina Sild Lönroth eines der langfristigen Hauptziele des Programmes darin sieht, dass die bildungsbenachteiligten Kinder schließlich einen Zugang zur Universität Malmö anstreben (Lönroth 2007, S. 17), ist dieses Ziel im Wiener Programm nicht explizit enthalten. Das erklärte Ziel von ‚Nightingale‘ Wien ist die „Stärkung von bildungsbenachteiligten Kindern durch Mentoring“ (Leeb 2014, S. 1). In Wien stellt das segregative österreichische Schulsystem einen Grund für die Bildungsbenachteiligung von Kindern mit Migrations- oder Fluchthintergrund dar. Bereits nach vier Jahren Grundschule werden Kinder, die an diesem Punkt meist zehn Jahre alt sind, nach ihrer Leistung in die nächsthöhere Schulstufe eingeteilt. Die nächsten Schulstufen sind die ‚Neue Mittelschule (NMS)‘, die 2015/16 ausgelaufene ‚Kooperative Mittelschule (KMS)‘ und die ‚Allgemeinbildende Höhere Schule (AHS)‘. Während die AHS mit der Matura abschließt, enden NMS und KMS ein Jahr vor Ende der Schulpflicht mit dem Pflichtschulabschluss – dieser ist Voraussetzung

für den Besuch einer weiterführenden Schule oder für die Absolvierung einer Berufsausbildung. Da die Unterrichtspflicht zwischen 6 und 15 Jahren gilt, kann es sein, dass Kinder, die aufgrund von Flucht oder Migration später eingeschult werden oder Schuljahre verlieren, nicht mehr schulpflichtig sind, bevor sie ihren Pflichtschulabschluss erhalten haben: „Sie sind dann bereits über 15 Jahre alt bzw. haben schon neun oder zehn Pflichtschuljahre absolviert, sind dann aber nicht mehr berechtigt die vierte Klasse Hauptschule/NMS (noch einmal) zu machen und abzuschließen“ (Wöhler et al. 2017, S. 50). Kinder, die erst in der Schulzeit die deutsche Sprache lernen, erleben so Bildungsbenachteiligung, da die Deutschkenntnisse oft nicht ausreichen um den Anforderungen zu entsprechen (vgl. Wöhler et al. 2017, S. 49 ff.). Hier knüpft ‚Nightingale‘ an: Mentoring „bietet [...] im Kinder- und Jugendbereich eine enormes Potential, soziales Lernen zu unterstützen und migrationsbedingte Nachteile im Bildungsprozess zumindest teilweise und temporär auszugleichen“ (Leeb 2014, S. 26). Dementsprechend ist die Auswahl der teilnehmenden Schulen für die Kooperation mit ‚Nightingale‘ an den Problemstellungen der Zielgruppe orientiert – die Schulen werden anhand folgender Merkmale ausgewählt:

„Verortung in einem Stadtgebiet mit hohem Anteil an Familien mit Migrationshintergrund und niederem sozioökonomischem Status und (zu) geringe Ressourcen der Schule zur Bewältigung der Herausforderung durch die vielsprachige und kulturell heterogene Schüler/innen-Population“ (Leeb 2014, S. 14).

Wie eingangs erklärt, können sich Kinder freiwillig für die Teilnahme melden. Abseits der freiwilligen Anmeldung werden Kinder von Lehrpersonen vorgeschlagen bzw. eingeladen, wenn sie von den LehrerInnen als schüchtern, ängstlich oder wenig selbstbewusst wahrgenommen werden:

„Die Auswahl der Schüler/innen, die eine/n Mentor/in erhalten sollen, obliegt grundsätzlich der Schule, allerdings nach festgelegten Kriterien bezüglich der Schüler/innen. Diese sind:

- geringer Selbstwert, wenig Selbstvertrauen,
- Förderbedarf im Deutsch-Spracherwerb – Integrationsbedarf im Klassenverband,
- schwierige familiäre Verhältnisse, wenig familiäre Förderressourcen,
- erst kurz in Österreich oder an der Schule,
- andere für die Schule wichtige Gründe“ (Leeb 2014, S. 14).

Um am Programm tatsächlich teilnehmen zu können, bedarf es der Zustimmung bzw. Unterschrift der Eltern. Hierbei gibt es häufig Schwierigkeiten, da den Eltern trotz mehrsprachiger Informationsfolder und -gespräche das Angebot und die Ziele des Programmes oft

nicht ausreichend nähergebracht werden können. Die Eltern sehen ihr Kind oft als zu jung oder zu unreif an oder möchten es nicht einer fremden Person überlassen. Jene Eltern allerdings, die sich auf das Programm einlassen, sind meist sehr zufrieden damit und möchten die Teilnahme am Programm verlängern. Auch die Evaluierungsergebnisse der Eltern sind durchwegs positiv (vgl. Leeb 2014, S. 19).

Da der Schwerpunkt dieser Arbeit in der möglichen Wirkungsorientierung des Programmes in Wien liegt, ist es relevant, auf dessen (Wirkungs-)Ziele und derzeitige Evaluierungsform besonders einzugehen.

1.1.2 Ziele und Evaluierung – ‚Nightingale‘ Wien

Allgemein formuliert der Projektleiter von ‚Nightingale‘ Wien, Günther Leeb, das Ziel des Mentoring-Programmes in der Stärkung und Förderung von Selbstvertrauen und sozialen Kompetenzen der Mentees (vgl. Leeb 2014, S. 3). Als konkrete Ziele werden im Informationsblatt (Leeb 2016) folgende angegeben:

- „Selbstwert stärken, Orientierung in Wien verbessern, Kennenlernen der Bildungs- und Freizeitmöglichkeiten der Stadt.
- Kompetenzen erwerben wie Mobilität, Selbsteinschätzung, kulturelle Einrichtungen kennenlernen.
- Bildungsmotivation stärken durch den Kontakt mit einer Studentin, die den in Österreich höchstmöglichen Bildungsabschluss anstrebt.
- Unterstützung der Familie bei Lern- und Bildungsfragen.“ (Leeb 2016, S. 1)

Bezüglich der Konkretheit von Zielsetzungen zeigt Herr Leeb zwei unterschiedliche Mentoring-Stile auf (vgl. Karcher und Nakkula 2010, S. 18, zit.n. Leeb 2014, S. 5): Das sogenannte *developmental mentoring* baut auf die Tandembeziehung, um in einem kooperativen Setting Fortschritte und Veränderungen in der Persönlichkeitsentwicklung zu erreichen. Das *instrumental mentoring* hingegen hat klare Zielsetzungen, die konsequent verfolgt und erreicht werden sollen. Leeb ordnet ‚Nightingale‘ dem *developmental mentoring* zu (vgl. Leeb 2014, S. 7). Daher sind die Ziele offen formuliert und nicht ohne weiteres überprüfbar. Die Schwierigkeiten und Grenzen der Evaluierung zeigt Herr Leeb in den veröffentlichten Evaluierungsergebnissen des Jahrganges 2013/14 auf. Diese Evaluierung wurde anhand a) von Monatsberichten der Studierenden, b) eines Evaluierungsfragebogens, beantwortet durch die Studierenden, c) von an Eltern und Kinder adressierten Fragebögen, d) eines Evaluierungs-

Workshops mit einer Auswahl von Mentees und e) eines Reflexionsgesprächs mit den teilnehmenden Lehrpersonen durchgeführt (vgl. Leeb 2014, S. 15 ff.). Im evaluierten Programmjahr nahmen 33 Tandems teil. Die Ergebnisse der Evaluierung wurden beispielhaft zusammengefasst und entlang der Kategorien ‚Erwerb von Wissen über Kultur- und Bildungseinrichtungen‘, ‚praktische Fähigkeiten‘ und ‚soziale Kompetenzen‘ dargestellt. Darüber hinaus wurde die Verbesserung der Sprachkenntnisse besprochen. Den über die Intervention hinaus bestehenden Kontakt zwischen den Tandems sieht der Leiter von ‚Nightingale‘ Wien als wichtigen Aspekt der Nachhaltigkeit des Programmes an (vgl. Leeb 2014, S. 26). Die vorgestellten Ergebnisse konzentrieren sich hauptsächlich auf die Aussagen zu den unternommenen Aktivitäten und deren Bedeutung für die Lernfortschritte der Mentees. Diese Angaben wurden aus den von Studierenden beantworteten Online-Evaluierungsbogen gewonnen und stellen damit die Perspektive der MentorInnen dar (vgl. Leeb 2014, S. 16 ff.). Günther Leeb geht in diesem Zusammenhang konkret auf die Schwierigkeit der Ermittlung von Kompetenzgewinnen ein:

„Generell ist es schwierig bis unmöglich, exakt zu ermitteln, welche Kompetenzgewinne im Sozialverhalten, im Bereich Selbstwert oder im Bereich Deutsch-Spracherwerb die teilnehmenden Schüler/innen durch die Aktivitäten mit ihren Mentor/innen erworben haben und welche Entwicklungsfortschritte sich durch den Schulunterricht und durch das soziale Umfeld ergeben haben“ (Leeb 2014, S. 16).

Der Projektleiter beschreibt damit die Herausforderung einer Wirkungsanalyse, die Wirkungen eines Programmes unter Ausschluss anderer Einflüsse direkt auf ebendieses zurückführen zu können. Dennoch fasst Leeb zusammen: „Das Gesamtbild aus den erwähnten Ermittlungsmethoden nährt jedoch die Gewissheit, dass das Programm praktisch bei allen Kindern eine ganze Reihe von positiven Auswirkungen hinterlässt.“ (Leeb 2014, S. 16). Die bisherige Evaluierung von ‚Nightingale‘ illustriert eine typische Problemlage – viele Mentoring-Programme berichten von positiven Auswirkungen, ohne aber entsprechende Nachweise erbringen zu können (vgl. DuBois et al. 2011, S. 59). Dies mag zwar auch durch einen Mangel an Ressourcen begründet sein, der eine genaue und wissenschaftliche Auseinandersetzung mit den Wirkungen nicht zulässt, allerdings weist Herr Leeb drauf hin, dass mitunter auch in großen Studien zu Mentoring nicht konkret geklärt werden kann, welche Wirkung warum eingetreten ist (vgl. Leeb 2014, S. 26). Damit spricht Günther Leeb die Forschung zu Jugend-Mentoring-Programmen in den Vereinigten Staaten an, die sich äußerst aktiv einer Reihe

von Fragen zu Mentoring und dessen Wirkungen widmet. Auf diese Forschung und die relevanten Aussagen aktueller Studien wird im empirischen Teil der Forschungsstandarddarstellung (Kapitel 2.2.) eingegangen.

2 Forschungsstand

Der in diesem Kapitel dargestellte Forschungsstand gliedert sich in einen eher theoretisch-historischen und einen empirischen Teil. Es sollen einerseits konstitutive Begriffe aus dem Bereich der Wirkungsforschung dargestellt werden, die eine Verortung und Abgrenzung des Zugangs dieser Arbeit zulassen. Darüber hinaus wird auf die breite Diskussion zur Wirkungsorientierung und Wirkungsanalyse in der Sozialwirtschaft und Sozialen Arbeit verwiesen, die einen Einfluss auf die ausgewählte Methodologie hat. Anschließend wird zu den Erkenntnissen aus der empirischen Erforschung der Wirksamkeit von Mentoring-Programmen im Allgemeinen übergeleitet. Abgeschlossen wird das Kapitel mit der Präsentation der bisher durchgeführten Wirkungsanalysen zum ‚Nightingale‘-Programm.

2.1 *Theorie und Implementierung der Wirkungsorientierung in der Sozialwirtschaft*

Dieses Kapitel wird zunächst einen Überblick über die Theorie und deren wichtigste Termini und Definitionen geben und zugleich ausgewählte Zugänge zur Wirkungsanalyse vorstellen. Daran anschließend wird auf die jüngere Vergangenheit und heutige Situation der Wirkungsorientierung und deren Besonderheiten in Sozialwirtschaft und Sozialer Arbeit eingegangen. Das Kapitel schließt mit einer Diskussion um die Möglichkeiten und Risiken der Wirkungsorientierung im sozialen Bereich.

2.1.1 Was ist Wirkung? Begriffserklärung, Verortung und Abgrenzung

Die Wirkung sozialer Maßnahmen wird immer häufiger in den Mittelpunkt gestellt. Gerade bei *Social Entrepreneurships* und *Start-Ups* kommt kaum ein Konzept ohne die Begrifflichkeit des (*social*) *impacts* aus. Impact wird zumeist als Synonym für alle Wirkungen, die eine Maßnahme haben soll, verwendet. Da Wirkungen in der wissenschaftlichen Auseinandersetzung differenzierter betrachtet werden, ist eine Behandlung der zentralen Begriffe Voraussetzung für diese Arbeit.

Der Begriff der *Wirkungsorientierung* umfasst die Ausrichtung eines Projektes nach konkreten Wirkungszielen – die Wirkung des Projektes steht also von Anfang an im Vordergrund. Die gesamte Projektplanung wird anhand dieser Ziele ausgerichtet und anhand von Monitoring und Evaluierung regelmäßig überprüft. Ein wirkungsorientiertes Projekt besteht aus der Planung der Wirkung, aus deren Analyse sowie aus deren Verbesserung bzw. der Verbesserung der Planung (vgl. Kurz und Kubek 2013, S. 6 ff.). Solche Verbesserungen sind laut Kurz und Kubek (vgl. ebd.) ein wichtiger Aspekt von Wirkungsorientierung, die leider

allzu häufig eher der Legitimierung des eigenen Tuns diene und für die Eigendarstellung genutzt würde: Die „wesentliche Bedeutung von Wirkungsorientierung liegt aber im Lernen, in der kontinuierlichen Verbesserung der eigenen Arbeit“ (Kurz und Kubek 2013, S. 7). Kurz und Kubek, die Autorinnen des ‚Kursbuch Wirkung‘ stellen dabei Wirkungsorientierung als Projektmanagement-Tool vor, welches sich an „Projekte, Programme und Organisationen [richtet], [...] [die] mit geringen Ressourcen für Monitoring und Evaluation mit einfachen Schritten die Wirksamkeit ihrer Arbeit überprüfen und die Ergebnisse zum Lernen nutzen [können]“ (Kurz und Kubek 2013, S. 9). Die Wirkungsorientierung dient hierbei, eine Projektidee holistisch zu betrachten, konkrete Ziele festzulegen und durch Monitoring zu beobachten. Dabei sollen auch Veränderungen von Faktoren, nicht-intendierte Wirkungen und Fehler berücksichtigt werden. Wirkungsorientierung stellt einen Rahmen dar, der unterstützen kann, ein (sozialwirtschaftliches) Projekt zu realisieren.

Wirkungsanalyse hingegen ist gemäß der ‚Deutschen Gesellschaft für Evaluation‘ eine

„systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes [...], die erzielten Ergebnisse, Schlussfolgerungen oder Empfehlungen [...] müssen nachvollziehbar auf empirisch gewonnenen qualitativen oder quantitativen Daten beruhen“ (Uebelhart und Zängl 2013, S. 269).

In diesem engeren Sinne erfordert die Wirkungsanalyse das Erheben von Daten zur nachvollziehbaren Darstellung von Wirkungen. In einem weiteren Sinne zielt die Wirkungsanalyse darauf ab, festzustellen, welche Einflussfaktoren diese Wirkung ermöglichen. Die Instrumente der Wirkungsanalyse sind Monitoring und Evaluierung, welche in ähnlicher Weise unterscheidbar sind. Monitoring erlaubt durch eine regelmäßige Erhebung von Daten eine Überprüfung der geplanten Schritte sowie leicht erfassbarer Wirkungen und beschreibt *was* passiert. Komplexere Wirkungen, deren Einflussfaktoren und die Qualität des Projektes insgesamt können durch Evaluierung festgestellt werden. Für Abschlussberichte oder zur Erfüllung von Berichtspflichten gegenüber Fördergebern werden die Daten aus Monitoring und Evaluierungen herangezogen. Letztere werden häufig am Ende eines Projektes durchgeführt. Mitunter wird die Wirkungsanalyse auch *Wirkungsmessung* genannt, das jedoch exakte und quantifizierbare Messung der eingetretenen Wirkungen impliziert. Da dies nicht immer möglich ist präferieren Kurz und Kubek den Begriff der Wirkungsanalyse (vgl. Kurz und Kubek 2013, S. 48 ff.).

Wirkungen werden von Plé (2015) wie folgt definiert: „Nach wissenschaftlichem Verständnis sind Wirkungen alle Veränderungen, die nach einer Intervention (Maßnahme) auftreten“

(Plé 2015, S. 26). Diese allgemeine Sicht auf Wirkung wird von Kurz und Kubek mit der sogenannten ‚Wirkungstreppe‘ (vgl. Kurz und Kubek 2013, S. 35) ausdifferenziert, wie in Tabelle 1 gezeigt wird. Die Autorinnen unterscheiden in der Entwicklung und Durchführung von sozialen Projekten sieben Stufen, wobei erst ab der vierten Stufe von Wirkungen gesprochen wird:

Tabelle 1: Wirkungstreppe nach Kurz und Kubek 2013, S. 35 (eigene Darstellung).

Inputs	Ressourcen, die in das Projekt investiert werden (Arbeit, Geld, Zeit, Ausstattung, Kooperationen etc.)		
Outputs	1. Leistungen, die vom Projekt erbracht werden (Angebote).	2. Die Nutzung dieser Angebote.	3. Die Zufriedenheit der TeilnehmerInnen mit dem Angebot.
Outcomes (Wirkungen auf die Zielgruppe)	4. Neues Wissen und gefestigte Fähigkeiten.	5. Veränderungen im Handeln.	6. Veränderungen in der Lebenslage.
Impact (Wirkungen auf gesellschaftlicher Ebene)	7. Soziale, ökonomische oder ökologische Veränderungen im angestrebten Ausmaß (Stadtteil, Bevölkerungsgruppe, Gesamtgesellschaft).		

Die Unterscheidung zwischen den Stufen *Input*, *Output*, *Outcome* und *Impact* ist hilfreich bei der Strukturierung der verschiedenen Ebenen auf denen Wirkungen auftreten können. So kann der Fokus dieser Arbeit explizit auf die Ebene der Outcomes, d.h. auf die Wirkungen bezüglich der Zielgruppe, gelegt werden. Diese Ebene zeigt auf „auf welche positiven Veränderungen bei den am Projekt Teilnehmenden das Projekt abzielt“ (Kurz und Kubek 2013, S. 37). Wirkungen werden aber noch nach einem weiteren Gesichtspunkt differenziert, indem die intendierten von den nicht-intendierten Wirkungen unterschieden werden. Im Gegensatz zu Ersten sind nicht-intendierte Wirkungen solche, die nicht geplant oder erwartet werden und können sowohl unerwünscht als auch unverhofft erwünscht sein (vgl. Plé 2015, S. 26). Neben verschieden differenzierten Wirkungen muss das „Kontrafaktische“ (Plé 2015, S. 28) in der Wirkungsanalyse berücksichtigt werden. Dabei handelt es sich um eine ‚Nichtwirkung‘, d.h. um jenen Zustand, der ohne die getätigte Maßnahme eingetreten wäre. Erst in der Differenz zur Nichtwirkung wird die eigentliche Wirkung sichtbar. Hierfür werden Datenerhebungsdesigns benötigt, die eine Kontroll- oder Vergleichsgruppe beinhalten.

Während Kurz und Kubek von Wirkungstreppe sprechen, ist im Allgemeinen von einem Wirkungsmodell oder einer Wirkungslogik („logic model“) die Rede. Es gibt – im Vergleich zu Kurz und Kubek – differenziertere Wirkungslogiken (vgl. Uebelhart und Zängl 2013), wie beispielsweise der ‚Programmbaum‘ von Beywl (2009). Dieser beinhaltet zusätzlich zu den hier bereits genannten Elementen noch weitere, etwa die *Incomes*, die im weiteren Verlauf dieser Arbeit von Bedeutung sind. *Incomes* stellen die Voraussetzungen, Erfahrungen, Einstellungen und Werte der Zielgruppe dar. Das sogenannte ‚Rhodes-Model‘ oder ‚Model of youth mentoring‘ (vgl. Rhodes 2005, S. 30 ff.) kann als verkürzte Wirkungslogik in Mentoring-Programmen betrachtet werden. Es beginnt direkt mit den Outputs (Mentoring-Beziehung) und endet mit den Outcomes (den Wirkungen auf die Mentees). Die Hauptkomponenten im Mentoring-Modell von Jean Rhodes stellen die Beziehung zwischen einem jungen Menschen (Mentee) und einem/einer Erwachsenen (MentorIn) dar. Diese Beziehung bildet die Grundlage für die Entwicklung des Kindes in den Bereichen soziale-emotionale sowie kognitive Kompetenz und Identität. Das Mentoring kann es dem Kind erleichtern sich auf fremde Personen einzulassen, die kognitive Entwicklung fördern und neue Sichtweisen eröffnen. Positive Veränderungen können auch hinsichtlich des Zugangs zur Schule und zum Lernen erreicht werden. Im Weiteren kann die Mentoring-Beziehung die Eigenwahrnehmung des Kindes verstärken und die zukünftige Identität des Kindes beeinflussen, beispielsweise durch den Beruf (im Falle von ‚Nightingale‘ das Studium) des/der MentorIn oder durch die ausgeführten Aktivitäten. Im Rhodes-Modell werden die Einflüsse auf die Mentoring-Beziehung aufgegriffen, die für eine Wirkungsanalyse relevant sind (vgl. Rhodes 2005, S. 30 ff.):

- a) persönliche zwischenmenschliche Geschichte bzw. Erfahrungen (positive wie negative) des Mentees,
- b) soziale Kompetenzen des Mentees,
- c) die Entwicklungsstufe des Mentees,
- d) die Dauer der Beziehung,
- e) die Programmgestaltung (hier sind das Matching, die Vorbereitung der Kinder und MentorInnen und die Supervision der Beziehungen zu nennen, aber auch die Ziele des Programms, die Auswahl der TeilnehmerInnen etc.) sowie
- f) Familie und Umfeld.

Alle Faktoren bis auf d) und e) können als *Incomes* betrachtet werden, da sie berücksichtigen, mit welchen Voraussetzungen das teilnehmende Kind in das Programm startet. Die *Incomes* der MentorInnen werden im Rhodes-Model vernachlässigt, spielen aber ebenso eine wichtige Rolle in der Beziehungsarbeit. Auf *Incomes* und weitere Einflussfaktoren wird in Kapitel 5.3 noch genauer eingegangen.

Einen standardisierten Vorschlag, wie ein Wirkungsmodell erstellt, aufbereitet und kommuniziert werden kann, bietet seit 2011 der ‚Social Report Standard‘ (SRS) an. Dabei bietet der eingetragene Verein der ‚Social Report Initiative‘ sämtliche Materialien unentgeltlich auf der Homepage an und unterstützt zudem mit Beratung. Durch die Zurverfügungstellung einer Berichtsvorlage wird dem Ziel nachgegangen, die „Resultate von Engagement im sozialen Sektor sichtbar und vergleichbar“ (SRI - Social Reporting Initiative e.V. 2017b) zu machen. Diese Vorlage wird bereits von unterschiedlichsten sozialwirtschaftlichen Unternehmen genutzt. Darunter befindet sich auch ein Mentoring-Programm aus Deutschland, ‚Balu und Du‘ (vgl. SRI - Social Reporting Initiative e.V. 2017a). Der SRS bietet dabei einen guten Einstieg in das Thema der Wirkungsorientierung, vernachlässigt aber die Methoden bzw. die Messbarkeit der Wirkungen und die Langfristigkeit.

Eine weitere Form der Wirkungsanalyse ist der ‚Social Return of Investment‘ (SROI), der die ökonomische Perspektive darstellt. Beim SROI wird zwischen Brutto- und Nettowirkungen unterscheiden, wobei von Erster ein sogenannter ‚Deadweight‘ subtrahiert wird, um zur ‚tatsächlichen‘ Wirkung zu gelangen (vgl. Schober und Rauscher 2014, S. 35 ff.). Dieser Begriff ist dem des Kontrafaktischen ähnlich und steht für Wirkungen, „die ohnehin eingetreten wären, auch ohne die konkreten Aktivitäten“ (Schober und Rauscher 2014, S. 8). Die Analyse des SROI versucht, der Investition in Geldeinheiten gemessene Wirkungen gegenüberzustellen. Diese Methode der ökonomischen Evaluierung erfreut sich aktuell großer Beliebtheit, da sie einfache Antworten auf komplexe Sachverhalte gibt. Dabei wird die Methode von Schober und Rauscher kritisch betrachtet und darauf verwiesen, dass die Werte als ‚Benchmark‘ (Vergleichswert) „nur in sehr seltenen Fällen und unter Berücksichtigung der jeweiligen Vorgehen zulässig“ (Schober und Rauscher 2014, S. 37) sind. Hierfür sei ein Beispiel aus dem Mentoring-Bereich gegeben: Das Mentoring-Programm ‚Big Brothers Big Sisters Deutschland‘ wurde seit der Gründung 2005 u.a. von der Benckiser Stiftung gefördert. 2013 wurde die Förderung eingestellt, um das Konkurrenz-Projekt ‚Balu und Du‘ zu finanzieren, das bei gleicher Leistung weniger Kosten verursachte. 2014 wurde eine SROI-Berechnung für ‚Balu und Du‘ durchgeführt, die den monetären Mehrwert des Programms

für die TeilnehmerInnen, die Gesellschaft und den Staat bis in das Jahr 2071 darstellen sollte – bis zum Jahr des Pensionseintritts der teilnehmenden Kinder. Die Berechnungen erforderte eine Fülle an Annahmen. So wurde zum Beispiel davon ausgegangen, dass die Kinder überdurchschnittlich ehrenamtlich engagiert sein werden, weil die Programm-Teilnahme sie für ein solches Verhalten inspiriert hätte. Dieses Beispiel verdeutlicht, dass der SROI mögliche Sozialrenditen aufzeigen kann – die als solche nicht oder nur schlecht überprüft werden können – aber keine eigentlichen Wirkungen repräsentiert (vgl. Täubner 2016, 65 ff.). Da der Fokus in der vorliegenden Arbeit nicht auf der monetären Bewertung von Wirkungen liegt, soll auf den SROI und andere monetären Wirkungsmethoden nicht weiter eingegangen werden. Stattdessen soll ein weiterer Begriff der Wirkungsorientierung diskutiert werden, der ungleich größere Relevanz für diese Arbeit hat: *Evidence-based Practice*.

Das *Evidence-based Practice*-Konzept (EBP) basiert auf der Argumentation, dass „Entscheidungen über Maßnahmen, Angebote, Hilfen und Interventionen auf eine völlig neue – durch empirische Forschung fundierte – Basis gestellt werden müsse“ (Oelerich und Otto 2011, S. 244). Quantitative Studien, konkret randomisierte Kontrollstudien und Metaanalysen, gelten dabei als *Gold-Standard*, während qualitative Zugänge am unteren Ende der Forschungshierarchie zu finden sind: „Im Wesentlichen (wenngleich stark vereinfacht) geht es um die Nutzung des besten verfügbaren wissenschaftlichen Wissens („best available evidence“) in der sozialarbeiterischen Praxis“ (Reidinger 2015, S. 173). Dieser forschungsbaasierte Ansatz ist aber aufgrund der komplexen Interventionsbereiche der Sozialen Arbeit umstritten. Einerseits sei es in den Handlungsfeldern der Sozialen Arbeit teilweise nahezu unmöglich, spezifische Ursache-Wirkungs-Logiken zu bestimmen (vgl. ebd. 2015, S. 174). Wo dies dennoch gelänge, stelle sich andererseits die Frage, welchen Erklärungswert die EBP aufweise:

„Die Testung klar definierter Handlungsprogramme im Rahmen einer Evidence-based Practice lässt zwar statistische Aussagen zu, unter welchen kontrollierten Bedingungen ein bestimmtes Programm bei einem spezifischen Klientel (wahrscheinlichkeitstheoretisch!) zu einer festgelegten Zielerreichung geführt hat, aber der Erklärungsgehalt dieser Ergebnisse und die handlungsrelevanten Erkenntnisse für die Praxis lassen zu wünschen übrig.“ (Oelerich und Otto 2011, S. 246).

Als Alternative zur EBP führen Oelerich und Otto die *Realistic Evaluation* (RE) an (vgl. Oelerich und Otto 2011, S. 247 ff.). Die RE – zunächst für die Kriminalprävention entwickelt – ist ein theoriebasierter Ansatz zur Wirkungsmessung, der sich besser zur Bearbeitung der

zumeist komplexen und mehrdimensionalen Probleme der Sozialen Arbeit eigne. Er sei in der Lage, sowohl die Outcome-generierenden Mechanismen als auch zugleich deren Kontextabhängigkeit zu berücksichtigen (vgl. Haunberger und Baumgartner 2017, S. 123):

„Ein Programm bzw. eine Intervention erzeugt nur dann Wirkungen, wenn diese in einem konkreten Kontext Wirkungszusammenhänge aktivieren, die sich beobachten lassen und Veränderungen erzeugen“ (Haunberger und Baumgartner 2017, S. 126).

Der Schwerpunkt der RE liegt im Verständnis der Wirkungsprozesse, während im sogenannten Black-Box-Verfahren, das häufig in der EBP angewandt wird, ausschließlich die (intendierten) erreichten Effekte betrachtet werden, ohne die zugrundeliegenden Kausalstrukturen zu erörtern. Neben den spezifischen Eigenheiten ihres Forschungszugangs wird die EBP auch als Steuerungsinstrument thematisiert. Im Diskurs um ‚Managerialisierung‘, ‚Ökonomisierung und Professionalisierung der Sozialen Arbeit, in dem das *New Public Management* (NPM) immer mitschwingt, entwirft Polutta einen Gegensatz zwischen ökonomischer und wissenschaftlicher Steuerung:

„Das wegen seiner forschungsmethodisch und professionstheoretisch höchst umstrittene Steuerungsmodell der EBP geht in seiner Zielsetzung über die Ansätze des New Public Management hinaus, indem es ein neues Verhältnis von Praxis und Wissenschaft sowie Sozialpolitik entwirft. Gegenüber den wirkungsorientierten Steuerungsversuchen über Marktmechanismen und wettbewerbliche Anreize setzt EBP auf eine rationale, an wissenschaftlicher Wirkungsforschung ausgerichtete Steuerung von Leistungen, Hilfen und Interventionen“ (Polutta 2010, S. 51 f.).

Welche Steuerungsmodi wirksam werden, ist stark davon abhängig, wie Wirkungsorientierung im privaten und im öffentlichen Sektor institutionalisiert wird. Oelerich und Otto (vgl. 2011, S. 244) zeigen anhand von Beispielen aus Deutschland und Großbritannien die voranschreitende Institutionalisierung des EBP-Konzepts auf, das in immer mehr gemeinnützigen und staatlichen Organisationen Anwendung findet. Dies betrifft den Mentoring-Bereich gleichermaßen. In den USA gibt es bereits einige EBP-Zentren (beispielsweise das ‚National Mentoring Ressource Center‘ oder das ‚Center for Evidence-Based Mentoring‘), in Europa wurde ein solches 2016 erstmals gegründet (‚European Center for Evidence-Based Mentoring‘ [ECEB Mentoring]² in Leeuwarden, Niederlande). Auch in Österreich ist dieser Steuerungsmodus angekommen. Welchen Stellenwert die Wirkungsorientierung und -steuerung

² Auf das ECEB-Mentoring im Zusammenhang mit Nightingale wird im Kapitel 5.2.11 eingegangen.

in der österreichischen Verwaltung einnimmt und welche Bedeutung dies für die vorliegende Arbeit hat, wird im Folgekapitel dargestellt.

2.1.2 Wirkungsorientierung und -steuerung im Kontext des NPM

Wirkungssteuerung bzw. wirkungsorientierte Steuerung stehen mit jener Neugestaltung öffentlicher Verwaltungen in engem Zusammenhang, die in den 1990er Jahren den Beginn des *New Public Managements* (NPM) markierte. NPM ist ein weltweites Phänomen, das sich privatwirtschaftlicher Methoden bedient und die Effizienz und Effektivität öffentlicher Sektoren steigern möchte. Erreicht werden soll dies durch eine stärkere Markt- und Wettbewerbsausrichtung, dezentrale Grundstrukturen, eine verstärkte KundInnen- und BürgerInnenorientierung und schließlich anhand von ziel- und ergebnisorientierter Steuerung bzw. Outcome- und Outputorientierung (vgl. Krause und Proeller 2017). Die Finanzierung sozialer Dienste wird an neue Bedingungen geknüpft, wie etwa Qualitätsnachweise und Erfolgskontrollen, welchen zumeist nicht fachliche sondern ökonomische Indikatoren zugrunde liegen (vgl. Zierer 2017, S. 74).

Die Wirkungsorientierung ist seit 2013 im Rahmen des Bundeshaushaltsgesetzes mit der ‚wirkungsorientierten Folgenabschätzung‘ (WFA) in Österreich implementiert. Die WFA umfasst neun Politikbereiche – darunter auch die ‚Wirkungsdimension‘ Soziales (vgl. Sozialministerium o. J., S. 4). Die bisherige ressourcenorientierte Steuerung in der österreichischen Bundesverwaltung basierte auf ‚negativen Leistungsanreizen‘, wie beispielsweise der Budgetkürzung bei dessen unvollständiger Ausschöpfung. Mit der Verankerung der Wirkungsorientierung in der Bundesverfassung und im Bundeshaushaltsgesetz 2013 soll der Schwerpunkt der Tätigkeiten in der Verwaltung nun auf die Wirkung bzw. auf das Erreichen von Wirkungszielen gelegt werden (vgl. Geppl et al. 2011, S. 421ff.). Dabei soll die wirkungsorientierte Steuerung folgende Vorteile haben: a) gesteigerte Transparenz, b) stärkere Außenorientierung, c) klare Verantwortlichkeiten und daher Förderung der Ergebnisverantwortlichkeit der Führungskräfte, d) Berücksichtigung der Gleichstellung von Mann und Frau und e) Schaffung der Voraussetzungen für eine ganzheitliche Steuerung und Verwaltung (vgl. Geppl et al. 2011, S. 422 ff.). Die Außenorientierung der Verwaltung bedeutet eine Schwerpunktverlegung von der Input- zur Ergebnisorientierung:

„Die knappen Ressourcen sollen für jene Tätigkeiten eingesetzt werden, welche die höchste Wirkung erwarten lassen. Das Hauptaugenmerk muss daher auf die Identifikation der effektivsten Stellhebel zur Erreichung der Wirkung gelegt werden“ (Geppl et al. 2011, S. 423).

2.1.3 Wirkungsorientierung in der Sozialwirtschaft und Sozialen Arbeit

Vor dem Hintergrund der Implementierung der Wirkungssteuerung in der öffentlichen Verwaltung ist die Kritik von Oelerich und Otto (vgl. 2011, S. 245) zu verstehen, dass der Wirkungsdiskurs nicht nur auf wissenschaftlicher Forschung und Aufklärung fokussiert, sondern einen erheblichen Einfluss auf wohlfahrtsstaatliche Angebote und die Soziale Arbeit an sich hat, „für die Soziale Arbeit [...] wird Wirkungsorientierung zum Kernelement von Ökonomisierungsprozessen“ (Polutta 2010, S. 50). Aufgrund der Entwicklungen in der öffentlichen Verwaltung ist die These naheliegend, dass sich das Prinzip der Wirkungsorientierung auch im Bereich der Fördermittel niederschlägt (vgl. Gössler 2014, S. 67). Soziale Programme werden in Zukunft stärker unter Druck geraten, Wirkungsnachweise zu erbringen und diese unter Umständen regelmäßig zu wiederholen. Dabei kann es zu einer Wettbewerbssituation kommen, in der die Organisation gewinnt, die ihre Wirkung am besten aufzeigen kann. Es ist daher sinnvoll, bereits frühzeitig und vor der Einführung einer allfälligen verpflichtenden Wirkungsanalyse zu eruieren, wie Wirkungen bestmöglich nachgewiesen werden können. Dies erlaubt es, ohne Zeitdruck Überlegungen aus dem Projekt/Programm und aus der Sozialen Arbeit mit entsprechender fachlicher Expertise entwickeln zu können.

Trotz aller Kritik und Sorge um die Wirkungsorientierung in der Sozialwirtschaft und Sozialen Arbeit, lassen sich auch einige positive Entwicklungen ausmachen. So stellt Gössler (2014, S. 68 ff.) fünf mögliche positive Begleiterscheinungen dar:

1. Die KundInnen stehen (wieder) im Mittelpunkt,
2. es können neue Steuerungskriterien zwischen Geldgebern und sozialwirtschaftlichen Organisationen ausgehandelt werden,
3. der Selbstwert der sozialen Arbeit wird gestärkt, wenn der Erfolg nachvollziehbar wird,
4. durch konkrete Wirkungsziele wird die Aufmerksamkeit der Arbeit fokussierter und
5. die Arbeit wird professionalisiert (auch wenn Gössler dies als ‚professionellen Fassadebau‘ sieht und damit meint, dass sich Organisationen und Projekte v.a. nach außen gut präsentieren und legitimieren [vgl. Gössler 2014, S. 69]).

Auch Polutta (2010) beschäftigt sich mit der Frage, ob die Wirkungsorientierung zu einer Professionalisierung oder De-Professionalisierung der Sozialen Arbeit führt und schließt mit der Erkenntnis, dass „Wirkungsforschung, die das Wissen um Wirkfaktoren der eigenen professionellen Praxis in einen reflexiven Verwendungszusammenhang und Wissensbestand

überführt, dringend benötigt wird“ (Polutta 2010, S. 59). Er sieht große Chancen einer Professionalisierung der Sozialen Arbeit (und Sozialpädagogik) durch „analytische Aufklärung, verfügbarem Erklärungswissen und [...] wissenschaftlicher Theoriebildung“ (Polutta 2010, S. 59). Ebenfalls betrachten Oelerich und Otto (2011, S. 246) eines der zentralen Potentiale des Wirkungsdiskurses in der besonderen Platzierung der Wissensgenerierung. Sicherlich ist darauf zu achten, wo der Schwerpunkt der Wissensgenerierung liegt, da gerade in der sozialen Interaktion bzw. Ko-Produktion die Einflüsse außerhalb der Maßnahme schwer zu erfassen sind und daher in Forschungsdesigns oft ausgeklammert werden. Die Wirkungsmodelle können dabei nicht (immer) in direkte Handlungsanweisungen für die Praxis umgesetzt werden,

„aber es kann einen Hintergrund für professionelle (Ermessens-)Entscheidungen darstellen, um evidenzbasiert zu begründen, welches Vorgehen mit Blick auf die je fallspezifischen Konstellationen und soziale Einbettung der AdressatInnen, angemessen ist“ (Ziegler 2010, S. 1066).

Die Diskussion um Wirkungsorientierung in der Sozialen Arbeit zeigt eine Reihe von (Heraus-)Forderungen auf, die für die konkrete Planung und Umsetzung von Wirkungsanalysen direkte Bedeutung haben. Von großer Relevanz ist, dass wirkungsorientierte Schritte aus der Praxis heraus gestaltet werden müssen: „Die Forderung nach ‚wirkungsorientierter Professionalisierung‘ [soll] nicht nur an die Soziale Arbeit heran getragen, sondern auch aus der Sozialen Arbeit heraus formuliert [werden]“ (Polutta 2010, S. 57). Die Wirkungsorientierung bietet große Chancen für die Sozialwirtschaft und die Soziale Arbeit, wenn eine reflexive und kritische Perspektive beibehalten wird. Eine Kooperation zwischen Praxis und Forschung ist dafür die Voraussetzung – diese bedeutet auch, dass die AkteurInnen in die Wirkungsorientierung miteinbezogen werden. Es bedarf konkret formulierter Wirkungsziele und Erhebungsdesigns, die mit dem Wesen der Sozialwirtschaft und Sozialen Arbeit vereinbar sind. Beim Implementieren einer wirkungsorientierten Perspektive muss Platz sein für die situativen und individuellen Anforderungen dieser Branche, ein „Titanic-Syndrom“ (Gössler 2014, S. 69), bei dem Unternehmen aufgrund von zu starrer Verfolgung falscher Ziele „versenkt“ werden, ist zu vermeiden.

All diese Aspekte sind in der Erstellung einer Wirkungsanalyse zu beachten und erfordern ein hohes Maß an transdisziplinärer Zusammenarbeit Seitens Forschung und Praxis. Wie

angedeutet, ist die Zusammenarbeit mit den AkteurInnen ein konstitutiver Teil der Forschung. Sie zielt darauf ab, die Praxis des Programmes zu durchblicken und neben intendierten auch nichtintendierte Wirkungen erkennen zu können:

„Die Perspektive der Wirkungsorientierung mit all ihren Facetten hinsichtlich professioneller Interventionen zur sozialen Problemlösung erhöht die Wahrnehmung der Komplexität, Situationsgebundenheit und vielfältigen Variierbarkeit sozialer Prozesse“ (Astleithner 2012, S. 10).

Mit dem Bewusstsein um die Möglichkeiten, Herausforderungen und Sorgen der Sozialwirtschaft und Sozialen Arbeit in Bezug auf Wirkungsorientierung wird der Frage nachgegangen, wie ein langfristiges Forschungsdesign für eine Wirkungsanalyse im Mentoring-Programm ‚Nightingale‘ geplant und umgesetzt werden kann. Hierfür werden im Anschluss Erkenntnisse aus der empirischen Wirkungsforschung im Mentoring vorgestellt.

2.2 *Erkenntnisse der empirischen Wirkungsforschung im Mentoring*

In diesem Kapitel soll der empirische Forschungsstand zu Wirkungsorientierung bzw. Wirkungsanalyse in Mentoring-Programmen aufgezeigt werden. Zunächst werden die Forschungsergebnisse allgemeiner ausgewählter Studien kurz beschrieben, anschließend werden aktuelle Wirkungsstudien zu ‚Nightingale‘ dargestellt.

2.2.1 Wirkungsorientierung in Mentoring-Programmen

Mentoring-Programme im Tandemsetting sind seit der Gründung von ‚Big Brother Big Sister‘ im Jahre 1904 immer weiter verbreitet und erforscht. In den USA gibt es 5.000 Mentoring-Programme mit über drei Millionen TeilnehmerInnen [vgl. DuBois et al. 2011, S. 58 (Stand 2006)]. Dabei dienen diese Programme als Intervention in unterschiedlichen sozialarbeiterischen Bereichen: „Accordingly, mentoring has enjoyed relatively wide use as an intervention strategy in diverse spheres of policy and practice, including education, juvenile justice, and public health“ (DuBois et al. 2011, S. 58). Sie gelten als anschauliche Praxisbeispiele für die positive Jugendentwicklung, indem sie das Problem ‚fehlender Vorbildfunktionen‘ mit der Zurverfügungstellung eines Erwachsenen lösen (vgl. DuBois et al. 2011, S. 58). Dabei sind Wirkungen häufig Gegenstand der Forschung. In aktuellen Evaluierungen diverser Mentoring-Programme wird häufig auf die Ergebnisse einer Metastudie von DuBois et al. (2011) eingegangen und versucht, die von den AutorInnen aufgezeigten Lücken für die eigene Evaluierung zu schließen (vgl. Gettings und Wilson 2014; Herrera et al. 2013;

Melling und Gurjee 2013; Lönroth 2013). Dank der Größe dieser Metastudie kann im Folgenden ein recht breiter Forschungsüberblick skizziert werden.

DuBois et al. (2011) fragen nach den typischen Wirkungen von Mentoring-Programmen (Tandem Kind und Erwachsener), sowie nach den Kriterien, die ein Programm erfolgreich machen, also einen positiven Outcome ermöglichen. Dafür untersuchten die AutorInnen 73 Studien zu Mentoring-Programmen, die zwischen 1999 und 2010 publiziert wurden. Die Studie baut auf der bereits abgeschlossenen Forschung von DuBois et.al. (2002) auf, welche Studienergebnisse von Mentoring-Programmen bis 1998 untersuchte. Inkludiert sind ausschließlich quantitative Evaluierungen, die eine Kontroll- oder Vergleichsgruppe beinhalten. Auch wurden nur Studien einbezogen, die ausreichend Daten zur Verfügung stellen konnten, um eine Effektstärke darzustellen. Die Analyse wurde auf Evaluierungen limitiert, die am Ende des jeweiligen Programmes stattfanden, da nur sieben Studien ein *Follow-Up*-Treffen nach dem Programm vorgesehen hatten (vgl. DuBois et al. 2011, S. 66 ff.).

Die AutorInnen schließen, nachdem sie unterschiedliche Einflüsse, Wirkungen und Kriterien für Mentoring-Programme ausmachen, dass eine Weiterführung solcher sozialen Interventionen sinnführend ist: „we feel safe in concluding that mentoring is, by and large, an effective mode of intervention for young people“ (DuBois et al. 2011, S. 80). Dies bedeutet, dass quantitative Methoden in Verbindung mit der Anwendung von psychologischen Instrumenten durchaus in der Lage sind, positive Outcomes von Mentoring-Programmen aufzuzeigen.

Dennoch weisen DuBois et al. darauf hin, dass Praxis und Forschung enger zusammenarbeiten müssen. Obwohl sich die Forschung im Verständnis um die Qualitätsfaktoren und Wirkungen von Mentoring ständig weiterentwickelt, kommen die Erkenntnisse dieser Arbeiten in der Praxis nur langsam an. Zudem wird mehr in die Neugründung oder die Ausweitung bestehender Mentoring-Programmen investiert, als in die Qualitätserhaltung bzw. -steigerung. Die Forschung wird nicht oder nur wenig in die Entwicklung und Evaluierung von Programmen involviert oder genutzt (vgl. DuBois et al. 2011, S. 75). Ferner verweisen DuBois et al. darauf, dass Mentoring-Programme nicht ausschließlich positiv verlaufen und teilweise sogar schaden können. Dieser Aspekt würde in den Evaluierungen zu wenig berücksichtigt bzw. als Möglichkeit überhaupt nicht in Betracht gezogen. So kann beispielsweise das Beziehungsende negative Wirkungen haben, wenn die Tandems nicht gut darauf vorbereitet sind und dem Abschied nicht genügend Aufmerksamkeit, z.B. in Form einer ri-

tuellen Begehung, entgegengebracht wird (vgl. Zilberstein und Spencer (2014)). Auch Gettings und Wilson (2014) nehmen diesen Punkt auf und beleuchten die Faktoren, welche die zu erzielende Wirkung positiv oder negativ beeinflussen. Der Einsatz der MentorInnen wird stark von Variablen des *Investment Model* geleitet, das aus der Zufriedenheit der MentorInnen, deren Alternativen und ihrer eingebrachten Investition (materielle Güter, Zeit) besteht. Diese drei Variablen sind elementar für die Qualität der Beziehung und tragen somit zur Qualität des Mentorings insgesamt bei.

Zugleich bemängeln DuBois et al. einerseits das häufige Fehlen von Wirkungsforschung an sich und andererseits die fehlende Beobachtung der Nachhaltigkeit von Wirkungen. Diese empirische Forschungslücke zeigt deutlich, dass die aufgrund der Beziehungsarbeit in Mentoring-Programmen erhofften langanhaltenden Wirkungen derzeit eher den Status von Annahmen haben (vgl. DuBois et al. 2011, S. 74). Um diese nachweisen zu können, müssten Langzeitstudien bis in das Erwachsenenalter der TeilnehmerInnen durchgeführt werden, was von DuBois et al. als wichtig erachtet wird (vgl. DuBois et al. 2011, S. 79 ff.). Die Autoren fordern darüber hinaus eine Weiterentwicklung von theoretischen Zugängen im Wirkungsbzw. Evaluierungsbereich von Mentoring-Programmen, einerseits bezüglich der Methodologie von Wirkungsanalysen, andererseits mit Blick auf das Wirkungsmodell. Hinsichtlich der Erfassung künftiger empirischer Ergebnisse empfehlen die AutorInnen die Erstellung eines unabhängigen Evaluierungsregisters (DuBois et al. 2011, S. 80), das m.E. mittlerweile in den *Mentoring Program Reviews* des ‚National Mentoring Resource Center‘ verwirklicht wurde (vgl. Bowers et al. 2016a). Neben Ergebnissen finden sich hier Ressourcen zur Evaluierungsdurchführung wie z.B. das *Measurement Guidance Toolkit* (Bowers et al. 2016d), auf das in Kapitel 5.1.3.1.2 näher eingegangen wird.

Als methodologische Verbesserung kann die Einbindung eines zweiten Messpunktes erachtet werden, wie sie etwa bei Herrera et al. (2013) erfolgte. Diese Studie beschäftigt sich mit *Higher-risk*-Jugendlichen, die bis dahin in vielen Maßnahmen vernachlässigt wurden oder nur in geringer Anzahl unterstützt werden konnten. Es sind Jugendliche, die aufgrund ihrer sozialen und familiären Situation und ihres Umfelds, oft in Verbindung mit Armut, Herausforderungen entgegengestellt sind, die Probleme wie Schulabbruch, Kriminalität und Drogenmissbrauch begünstigen. Herrera et al. führten eine zweijährige Studie durch, bei der 1.310 Mentees im Alter zwischen 8 bis 15 Jahren involviert waren. Die Hälfte der Jugendlichen wurde randomisiert ausgewählt und einem/r MentorIn zugeordnet. Nach dem Follow-

Up (13 Monate später) wurden Tandems mit der zweiten Hälfte der Jugendlichen als Vergleichsgruppe gebildet. Von den Jugendlichen wurden Fragebögen zu Beginn- und Follow-Up-Zeitpunkt eingeholt, wobei die Eltern, MentorInnen und die ProgrammleiterInnen ebenfalls in die Befragung integriert wurden. Die Ergebnisse waren durchwegs positiv, die Zielgruppe erreichte Verbesserungen in der Schule, ein höheres Selbstwertgefühl sowie positive Auswirkungen auf die Gesundheit (Verringerung der Anzeichen einer Depression). Obwohl diese großangelegte Studie im Gegensatz zu vielen anderen Evaluierungen einen zweiten Messpunkt integrierte, wurde auch hier die Langfristigkeit der Wirkungen vernachlässigt. Zum Zeitpunkt des *Follow-Up* befanden sich 60 Prozent der Jugendlichen noch in einer Tandembeziehung. Eine dritte Datenerhebung nach Beendigung des Programmes wurde nicht in Erwägung gezogen.

2.2.2 Forschungsstand ‚Nightingale‘

2006 führte die ‚Nightingale‘-Gründerin Carina Sild Lönroth eine *Follow-Up*-Studie zum Projekt in Malmö durch, die langfristige Effekte retrospektiv untersuchen sollte. Ziel war es aufzuklären, ob das Projekt zu bedeutsamen Veränderungen im Leben der Mentees und ihrer Zukunftsperspektive beigetragen hatte. Von 27 kontaktierten ehemaligen TeilnehmerInnen stellten sich 14 Jugendliche (50 % weiblich, 50 % männlich) im Alter von 17 bis 19 Jahren, die zwischen 1997 und 2000 an ‚Nightingale‘ teilgenommen hatten, für ein Interview zur Verfügung. Die teilstrukturierten qualitativen Interviews enthielten Fragen zur Erinnerung der ehemaligen Mentees. Es wurde gefragt, inwieweit der/die MentorIn via Beziehung zum Mentee einen positiven Einfluss auf die Outcomes hatte (Lönroth 2007, S. 102). Im Ergebnis zeigte sich: „the time they spent with their mentor had led to an improved self confidence, they were more open and dared to do the kind of things they previously would not have dared to do“ (Lönroth 2007, S. 140). Hinsichtlich der Studiendurchführung benennt Lönroth zwei Schwachstellen. So stellt sie sich die Frage, ob sich nur jene zum Interview meldeten, die positive Erfahrungen gemacht haben. Gleichzeitig macht sie auch selbstreflexiv drauf aufmerksam, dass sie selbst alle Interviews durchgeführt hat und ihre positiven Erwartungen die Befragten beeinflusst haben könnten (vgl. Lönroth 2007, S. 103). Überlegungen zu einer Kontrollgruppe sind nicht formuliert.

Aktuelle Evaluierungen und Studien zu ‚Nightingale‘ kommen aus der Schweiz (Leutwyler et al. 2014) und Katalonien (Feu Gelis 2015). Die schweizerische Studie beschäftigt sich mit

den Möglichkeiten der Optimierung des Programmes in Zug (CH), welche mittels einer qualitativen Befragung der MentorInnen (n=7) zu zwei Zeitpunkten (vor und kurz nach dem Programm) erhoben wurden. Eine relevante Erkenntnis ist, dass die MentorInnen eine zielorientiertere Ausrichtung des Programmes wünschen (Leutwyler et al. 2014, S. 16). Bei der katalonischen Studie handelt es sich um ein quantitatives Verfahren, welches mit einer Kontrollgruppe arbeitete (nichtteilnehmende Kinder aus der Klasse, die die Teilnehmenden besuchten). Die Erhebung wurde zu zwei Zeitpunkten durchgeführt (Beginn und Ende des Projektes) und zeigt signifikante Wirkungen. Es wird argumentiert, dass Mentoring-Programme eine kostensparende und wirksame Methode sind, die sich auch im Vergleich mit anderen sozialpolitischen Maßnahmen bewährt. Dennoch wird auch auf die Notwendigkeit einer (weiterführenden) mittel- bis langfristigen Untersuchung verwiesen (vgl. Feu Gelis 2015, S. 148).

Zu ‚Nightingale‘ Wien wurden bisher mehrere Masterarbeiten erstellt. Die erste beschäftigte sich mit dem Beitrag von ‚Nightingale‘ zur Chancengleichheit (vgl. Blaßnig 2012), eine zweite analysierte 219 Monatsberichte von MentorInnen auf die jeweiligen pädagogischen Strategien bei Schwierigkeiten (vgl. Scherer 2016) und zwei weitere Arbeiten führten eine Evaluierung zur Entwicklung der Sprachkenntnisse der Mentees durch (Himmelbauer 2016; Trautsamwieser 2016). Hierbei wurden 22 Kinder und eine Vergleichsgruppe zu zwei Zeitpunkten mittels eines standardisierten Sprachtests geprüft, geringe Differenzen zwischen Ziel- und Vergleichsgruppe wurden festgestellt. Weiter veröffentlichte der Projektleiter des Programmes in Wien Evaluierungsergebnisse des Jahrganges 2013/14 (vgl. Leeb 2014), die in Kapitel 1.1.2 bereits besprochen wurden.

Zum Forschungsstand ist abschließend festzuhalten, dass eine Forschungslücke hinsichtlich der langfristigen Betrachtungsweise von Mentoring-Programmen besteht. Entsprechende Evaluationen werden aufgrund von fehlenden Ressourcen kaum durchgeführt. Dabei ist es für die Projektleitung, für Stakeholder und schließlich auch für die TeilnehmerInnen von Interesse, ob auch längerfristige Ziele erreicht werden. Für ‚Nightingale‘ Wien ist zusätzlich festzustellen, dass keines der in Kapitel 1.1.2 genannten Ziele systematisch untersucht wurde. In vorliegender Arbeit sollen daher Möglichkeiten und Grenzen für eine Wirkungsanalyse von ‚Nightingale‘ Wien erörtert werden. Hierzu wird ein Forschungsdesign entwickelt, das einerseits methodischen Ansprüchen genügt und andererseits dessen Anwendbarkeit und Akzeptanz durch Befragung von Stakeholdern und ExpertInnen eingeschätzt wird.

3 Erkenntnisinteresse und Forschungsfragen

Wie im Forschungsstand dargestellt wurde, werden Mentoring-Programme regelmäßig evaluiert und es findet eine laufende Weiterentwicklung der Methoden der Wirkungsanalyse statt. Das Wissen um die Wirkungen einer sozialen Maßnahme dient nicht nur der Verbesserung der Wirksamkeit selbst, sondern auch der Prävention von negativen Wirkungen. Eine Evaluierung zum Abschluss des Programmes reicht hierfür nicht aus, zum einen, weil zur Feststellung von Veränderungen Vergleichsdaten benötigt werden, zum anderen, weil soziale Interventionen häufig auf Verhaltens- oder Bewusstseinswirkungen hinarbeiten, die möglicherweise direkt nach dem Programmende noch nicht ausreichend erkennbar sind.

Evaluierungen sind bereits heute oft Bestandteil von Förderverträgen und sozialwirtschaftliche Organisationen berichten vermehrt öffentlich über ihre Leistung und die Wirkung ihrer Tätigkeiten (vgl. SRI - Social Reporting Initiative e.V. 2017a). Aufgrund der Implementierung der Wirkungsorientierung im österreichischen Bundeshaushaltsgesetz ist zu erwarten, dass soziale Programme stärker unter Druck geraten, ihre Wirkung nachzuweisen. Dabei bleibt aber die Frage offen, *wie* Wirkungen bestmöglich erhoben werden können. Am Beispiel von ‚Nightingale‘ Wien soll erörtert werden, welche Aspekte ein – langfristiges – Forschungsdesign berücksichtigen sollte, welche Ressourcen benötigt werden und wie eine Umsetzung realisierbar ist. Mit diesen Inhalten beschäftigt sich die vorliegende Arbeit. Es besteht nicht die Absicht, ein vollständiges, d.h. sämtliche Wirkungsziele beinhaltendes, Forschungsdesign zu entwickeln oder eine Wirkungsanalyse durchzuführen, sondern vielmehr durch einen explorativen Zugang die Möglichkeiten und Grenzen bei der Planung einer langfristigen Wirkungsanalyse aufzuzeigen.

In der Wirkungslogik nach Kurz und Kubek (2013) entstehen Wirkungen auf vier unterschiedlichen Ebenen (siehe Kapitel 2.1.1). Auf der letzten Stufe stehen die gesellschaftlichen Wirkungen, die Impacts genannt werden und in dieser Arbeit nicht berücksichtigt werden können. Die verbleibenden drei Stufen beziehen sich auf Outcomes, d.h. auf „Wirkungen auf die Zielgruppe“ (Kurz und Kubek 2013, S. 39). Diese Outcomes können als kausal aufeinander aufbauend gedacht werden. So zeigen sich Wirkungen zunächst a) in der Erlangung neuen Wissens und gefestigter Fähigkeiten, dann b) in Veränderungen im Handeln und schließlich c) in Veränderungen der Lebenslage (Kurz und Kubek 2013, S. 35). Outcomes verschiedener Ebenen sind auch in zeitlicher Hinsicht voneinander zu unterscheiden: Während die Erlangung neuen Wissens direkt am Projektende evaluiert werden kann, dauert eine

Veränderung im Handeln oft länger und eine veränderte Lebenslage kann vermutlich erst langfristig ausgemacht werden. Kurz und Kubek schlagen für die Erhebung von kurz- und mittelfristigen Wirkungen eine Vielzahl an Methoden an, wobei sich langfristige Wirkungen (Stufe 6) von den kurz- und mittelfristigen Wirkungen (Stufen 4–5) ableiten:

„Wenn während des Projekts festgestellt wurde, dass bei den Teilnehmern Wirkungen auf den Outcome-Stufen 4 (Veränderung im Wissen) und 5 (Veränderungen im Verhalten und Handeln) eingetreten sind, so kann mit einer gewissen Sicherheit angenommen werden, dass das Projekt auch Einfluss auf die langfristigen Wirkungen im Bereich der Lebenssituation (Outcome-Stufe 6) hatte“ (Kurz und Kubek 2013, S. 82).

Für die langfristige Betrachtung werden hier also keine spezifischen Erhebungsmittel oder Methoden zur Verfügung gestellt. Diesen Umstand berücksichtigend lassen sich folgende Fragen formulieren, die an Melling und Gurjee (2013, S. 1) angelehnt ist.

Wie kann festgestellt werden, ob die Teilnahme als Mentee bei ‚Nightingale‘ Einfluss auf die persönliche Entwicklung hat? Wie können insbesondere langfristige Wirkungen (Veränderungen im Handeln und in der Lebenslage) in Bezug auf ehemalige Mentees auf das Programm zurückgeführt werden?

Da diese Arbeit nicht nur den Stand der Forschung, sondern mittels Einbeziehung von Stakeholdern auch die aktuelle Praxis von ‚Nightingale‘ berücksichtigen soll, gliedert sich die Forschungsfrage in zwei Teilaspekte, wobei der erste Bezug zur Forschung, der zweite Bezug zur Praxis nimmt:

- 1. Unter Verwendung a) welches Forschungsdesigns, b) welcher Methode und c) welcher Indikatoren können die kurz-, mittel- und langfristigen Outcomes eines Mentoring-Programms bestmöglich nachgewiesen werden?**
- 2. Welche praktischen Möglichkeiten und Grenzen bestehen in der Planung und Umsetzung einer so angelegten Wirkungsanalyse?**

Um ein Forschungsdesign für eine Wirkungsanalyse entwickeln zu können, bedarf es konkreter Wirkungsziele. Hierzu wurden die Ziele für Mentees aus dem Informationsblatt (Leeb 2016) – siehe auch Kapitel 1.1.2 – in zwei Kernbereiche zusammengefasst. Um den Rahmen der vorliegenden Arbeit nicht zu sprengen, wurde aus jedem Kernbereich ein Hauptziel gewählt und zur weiteren Beantwortung herangezogen.

Tabelle 2: Eigene Kategorisierung der Ziele für Mentees nach dem Informationsblatt von ‚Nightingale‘ Wien für den Jahrgang 2016 (Leeb 2016).

Kompetenzerwerb	Bildung
Selbsteinschätzung	Kennenlernen der Bildungs- und Freizeitmöglichkeiten und kulturellen Einrichtungen
Selbstwert stärken	Bildungsmotivation stärken durch den Kontakt mit einer/einem Studenten/Studentin
Mobilität	Unterstützung der Familie bei Lern- und Bildungsfragen
Orientierung	

Die Steigerung der Bildungsmotivation war seit jeher ein wichtiges Ziel von ‚Nightingale‘. Die Maßnahme, eine/n StudentIn als MentorIn zur Verfügung zu stellen, der/die als Bildungsvorbild agieren soll, unterstreicht dies. In der vorliegenden Arbeit wird dieses Ziel als erstes Hauptziel des Programmes betrachtet. Als zweites Hauptziel von ‚Nightingale‘ wird die Steigerung des Selbstwertes aufgefasst. Einerseits ist dies ein (bereits beforschtes) Wirkungsziel von vielen Mentoring-Programmen im Kinder- und Jugendbereich (vgl. DuBois et al. 2011; Rhodes 2004; Rhodes et al. 2005) und andererseits hat es angesichts der Auswahl der Mentees („geringer Selbstwert, wenig Selbstvertrauen“ [Leeb 2014, S. 16]) eine wichtige Stellung im Gesamtzusammenhang des Programms inne. Kompetenzerwerb im Bereich Mobilität und Orientierung soll in dieser Arbeit vernachlässigt werden, ebenso das Kennenlernen diverser Einrichtungen, genauso wie die „Unterstützung der Familie bei Lern- und Bildungsfragen“ (vgl. ebd.).

Die zuvor genannten Forschungsfragen beziehen sich damit ausschließlich auf kurz-, mittel- und besonders auf langfristige Wirkungen in den Bereichen **Selbstwert** und **Bildungsmotivation**. Das folgende Kapitel erläutert die methodischen Überlegungen sowohl zum forschungs- als auch zum praxisorientierten Teil der Forschungsfrage.

4 Methode

Die vorliegende Arbeit beschäftigt sich mit den Möglichkeiten und Grenzen einer Wirkungsanalyse für das Mentoring-Programm ‚Nightingale‘ in Wien. Dabei liegt der Schwerpunkt auf den langfristigen Wirkungen auf die Zielgruppe (Mentees) in den Bereichen Bildungsmotivation und Selbstwert. Die aufgeworfene Forschungsfrage, wie festgestellt werden kann, dass die Teilnahme bei ‚Nightingale‘ Wien einen Einfluss auf die persönliche Entwicklung der Kinder hat, wurde in folgenden Schritten beantwortet.

- a) **Forschungsorientierter Teilaspekt:** Erstellung eines Forschungsdesigns unter Anwendung des SMART-Ansatzes (vgl. Kurz und Kubek 2013, S. 30) für eine langfristige Wirkungsanalyse (Kapitel 4.1).
 1. Theoriebasierte Exploration aufgrund des aktuellen Forschungsstandes zur Wirkungsanalyse sowie Literaturrecherche zu Erhebungsinstrumenten.
 2. Selektion und sinnvolles Aufeinanderbeziehen von Versuchsanordnung, Methode und Erhebungsinstrumenten bzw. Indikatoren für den gewählten Forschungszweck.
- b) **Praxisorientierter Teilaspekt:** Empirische Beforschung der aktuellen Praxis des Mentoring-Programms sowie der Einschätzung zur Durchführbarkeit des entwickelten Forschungsdesigns durch Einbindung von Stakeholdern (Kapitel 4.2).
 1. Durchführung einer empirisch-qualitativen Erhebung (leitfadengestützte ExpertInnen-Interviews).
Qualitative inhaltsanalytische Bearbeitung der Interviews sowie relevanter *Items* aus dem 2016/17 von MentorInnen beantworteten Onlinefragebogen.
 2. Quantitativ-deskriptive Darstellung relevanter *Items* des Onlinefragebogens.
- c) **Zusammenführung:** Aktualisierung und Diskussion des Forschungsdesigns vor dem Hintergrund der gewonnenen Ergebnisse.

Zusätzlich zu der Datenerhebung in Wien war durch die Teilnahme an der offiziellen ‚Nightingale‘-Konferenz in Barcelona am 15.6.2017 ein erster Kontakt mit dem gesamten ‚Nightingale‘-Netzwerk möglich. Die Teilnahme war aufgrund von zwei Vorträgen, die sich mit den Wirkungen des Mentoring-Programmes auseinandersetzen, für die vorliegende Arbeit relevant. Zusätzlich konnte an der Konferenz der persönliche Kontakt zur ‚Nightingale‘-Gründerin Carina Sild Lönroth geschaffen werden, um diese für die empirische Befragung zu gewinnen.

4.1 Erstellung eines Forschungsdesigns für eine langfristige Wirkungsanalyse

Zur Erstellung des Wirkungsanalyse-Konzepts wurde der SMART-Ansatz von Kurz und Kubek (vgl. 2013, S. 30) angewandt. Dieser Ansatz musste durch Überlegungen zum Ablauf der Erhebung mit besonderer Berücksichtigung der Langfristigkeit der Wirkungen einerseits sowie der unterschiedlichen Voraussetzungen der Mentees andererseits ergänzt werden. Anhaltspunkte konnten dem aktuellen Forschungsstand entnommen werden, woraus die Festlegung von Versuchsanordnung und Methode hervorging. Schließlich musste eine Selektion von getesteten, reliablen und validen Instrumenten erfolgen, die in der Lage sind, Bildungsmotivation und Selbstwert zu messen. Geeignete Instrumente wurden durch Literaturrecherche ausfindig gemacht (siehe hierfür Tabelle 5).

4.1.1 Operationalisierung der Ziele nach dem SMART-Ansatz

Konkrete Zielsetzungen bilden das Fundament einer Wirkungsanalyse. Nur mit operationalisierten Zielen kann überprüft werden, ob die gewünschte Wirkung eingetreten ist. Die Formulierung der Wirkungsziele beinhaltet eine Reihe an Entscheidungen: „Was und von wem etwas als wirksam erachtet wird, [ist] durchaus auch eine politische [Entscheidung]“ (Reidinger 2015, S. 175). Umso relevanter ist es, die Wirkungsziele aus dem Programm heraus zu entwickeln (vgl. Polutta 2010, S. 57). Kurz und Kubek (vgl. 2013, S. 30) empfehlen zur klaren Definition von Wirkungszielen eine Orientierung an SMART-Kriterien. Das Akronym SMART steht dabei für spezifisch (S), messbar (M), akzeptiert (A), realistisch (R) und terminierbar (T). Ein Ziel soll also möglichst eindeutig formuliert sowie messbar und daher quantitativer Natur sein (oder zumindest in irgendeiner Weise mit einem Referenzmaß vergleichbar). Es soll von den Stakeholdern geteilt werden und die mögliche Erreichbarkeit gegeben sein. Der Zeitpunkt der Zielerreichung sollte bereits im Voraus feststehen. Die Akzeptanz der Wirkungsziele ist stark davon abhängig, wie (und von wem) die Ziele formuliert und operationalisiert werden. Diese Zielformulierung wird jedoch nicht nur von der generellen Ausrichtung des Programmes beeinflusst, sondern auch von der Auswahl des Forschungsansatzes, da dieser das Wesen der Ziele maßgeblich mitbestimmt. Um die oben genannten Ziele von ‚Nightingale‘ zu konkretisieren, wurde – immer die angestrebte Akzeptanz berücksichtigend – über die Versuchsanordnung, die Methode und die Erhebungsinstrumente nachgedacht. Insofern hängt die Formulierung von Wirkungszielen anhand des SMART-Ansatzes direkt mit der Erstellung des Forschungsdesigns zusammen. Hierbei war

keine strikte Abfolge möglich, vielmehr handelte es sich um einen iterativen wechselseitigen Entwicklungsprozess zwischen Zielformulierung und -operationalisierung.

Die Erstellung eines Wirkungsanalyse-Konzepts hängt mit einem Wirkungsmodell (Hypothesen zur Kausalwirkung der Intervention) zusammen, das die Verbindungen und Wirkungspfade abbildet (vgl. Plé 2015, S. 28). Ein solches Wirkungsmodell, auch Wirkungslogik oder Wirkungstreppe (vgl. Kurz und Kubek 2013, S. 5) genannt, beinhaltet die Übersicht über das geplante bzw. durchgeführte Projekt/Programm. Hierbei werden die bereits vorgestellten Stufen Inputs, Outputs, Outcomes und Impact detailliert dargestellt und die Wirkungspfade festgelegt (vgl. Kurz und Kubek 2013, S. 33 ff.). Darauf aufbauend sind Versuchsordnung und Methode festzulegen. Hierbei war zu berücksichtigen, dass a) Mentees über unterschiedliche *Incomes*, d.h. über unterschiedliche Voraussetzungen vor der Teilnahme am Programm verfügen (vgl. Kapitel 2.1.1) und b) auch langfristige Wirkungen untersucht werden sollten. Es war eine Methode zu identifizieren, die es erlaubte, Wirkungen mit möglichst hoher Plausibilität auf die Intervention zurückzuführen. Eine Literaturrecherche zeigte Erhebungsinstrumente auf, die Bildungsmotivation und Selbstwert messen können (siehe hierfür Tabelle 5). Dabei wurden getestete objektive, valide und reliable Instrumente bevorzugt. Die Instrumente sollen dem Zweck der Wirkungsanalyse entsprechen und auf die Zielgruppe von ‚Nightingale‘ anwendbar sein. Mit den so gewonnenen Informationen konnte geprüft werden, inwieweit das Forschungsdesign die SMART-Kriterien erfüllte.

Das theoriegeleitete Forschungsdesign einer langfristigen Wirkungsanalyse mit dem Schwerpunkt der Steigerung der Bildungsmotivation wurde im Rahmen einer Seminararbeit erstellt und für die Teilnahme an der Fachhochschulkonferenz Krems zu einem *Extended Abstract* ausgearbeitet (vgl. Kofler 2017). Das Paper wurde durch zwei unabhängige GutachterInnen freigegeben und an der Konferenz vorgestellt. Da das Forschungsdesign mit dem Ziel der Steigerung des Selbstwertes vervollständigt wurde und anhand der erhobenen Daten auf seine Umsetzbarkeit und Akzeptanz hin überprüft wurde, wird es in der Ergebnisbesprechung unter Kapitel 5 dargestellt.

4.2 Datenerhebung durch die qualitative Befragung von AkteurInnen

In die Entwicklung des Forschungsdesigns sollten unterschiedliche Sichtweisen von AkteurInnen einfließen: Bevor standardisierte Instrumente zur Outcome-Messung generiert werden können, muss in „offenen, qualitativen Interviews mit verschiedenen Gruppen von Beteiligten“ herausgefunden werden, „was als Outcome bzw. Erfolg der Intervention aus Sicht

der Betroffenen und deren Bezugspersonen gelten kann“ (Plé 2015, S. 31). Dafür wurden Einzelinterviews mit unterschiedlichen AkteurInnen durchgeführt. Durch die Berücksichtigung der unterschiedlichen und sicherlich teilweise divergierenden AkteurInnen-Perspektiven konnte eine holistische Herangehensweise gefördert werden (siehe auch Plé 2015, S. 28).

4.3 Zugang zum Feld und Beschreibung des Feldes

Der Zugang und die Kontaktmöglichkeiten zu ‚Nightingale‘ Wien waren durch meine eigene Teilnahme als Mentorin im Programmjahr 2014/15 bereits gegeben. Ich führte am 14. September 2016 ein informelles Gespräch mit dem Projektleiter Günther Leeb, um ihm meinen Ansatz vorzustellen und sein Interesse an dem Vorhaben zu klären, bzw. herauszufinden, in welchem Ausmaß Unterstützung hinsichtlich der Zurverfügungstellung von Daten oder Kontakten seinerseits möglich war. Dieses Gespräch wurde nicht aufgezeichnet, sondern im Nachhinein in Form eines Memos festgehalten. Günther Leeb sicherte mir sein Interesse und seine Unterstützung zu. Er stellte mir die anonymisierten Daten der Onlinebefragung der MentorInnen des Programmjahres 2014/15 zur Verfügung, um mir einen ersten Einblick verschaffen zu können. Anhand dieses Gespräches und einer vertiefenden Recherche wurde die Forschungsfrage verändert und konkretisiert. Zunächst richtete sich der Fokus auf die Outcomes der MentorInnen, wobei sich im weiteren Verlauf die Wirkungen des Programmes auf die teilnehmenden Kinder als besser geeigneten Gegenstand für diese Arbeit erwiesen. Weiter ermöglichte er mir den Zugang zu der aktuellen Onlinebefragung der MentorInnen für den Jahrgang 2016/17, die ich auf ausgewählte Aspekte hin analysierte.

Zunächst war es notwendig, zu klären, welche Gruppen von Beteiligten als AkteurInnen in die Erhebung miteinzubeziehen sind und in welcher Form dies zu geschehen hat. Aufgrund der Problemzentriertheit der Fragestellung bot sich das ExpertInnen-Interview als Erhebungsmethode an. Dabei stellte sich die Frage, wessen ExpertInnen-Wissen Relevanz für die Entwicklung des Forschungsdesigns hat. Was zeichnet diese ExpertInnen aus? Bogner et al. (2014, S. 9 ff.) diskutieren die Frage, unter welchen Bedingungen Akteure in einem empirischen Sozialforschungszusammenhang als Experte/Expertin zu betrachten sind. Entscheidender Punkt ist, dass ExpertInnen nicht nur über ein bestimmtes Wissen verfügen, sondern, „dass dieses Wissen in besonderer Weise praxiswirksam und damit orientierungs- und handlungsleitend für andere Akteure wird“ (Bogner et al. 2014, S. 14). Dies trifft jedenfalls auf die ProgrammleiterInnen (Leeb, Leskowa) und die Mitarbeiterin Nadejda Sabeva (in weitere

Folge wird diese Gruppe als OrganisatorInnen bezeichnet) zu, die das Programm in seiner Gesamtheit überblicken. Sie verfügen über relevantes Praxiswissen und „Prozesswissen“ (Bogner et al. 2014, S. 18), also über Kenntnisse des Mentoring-Netzwerkes, der Organisation und der AkteurInnen. Bezüglich der orientierungs- und handlungsleitenden Perspektive ist Günther Leeb verantwortlich für die Planung, Umsetzung und Evaluierung des Programmes, das ihm die charakteristische „Macht-Wissen-Konfiguration“ (ebd.) erlaubt. Zur Vervollständigung wurde die Gründerin und Managerin des ‚Nightingale‘-Netzwerkes, Carina Sild Lönroth, befragt um die länderübergreifenden Perspektive miteinzubringen.

Im Sinne der Definition von Bogner wurden auch andere AkteurInnen als ExpertInnen aufgefasst, die nicht an der Organisation des Programms beteiligt sind, deren Wissen jedoch ebenfalls auf andere AkteurInnen des Programms rückwirken kann. Konkret handelte es sich dabei um Mentees, MentorInnen, Lehrpersonen der teilnehmenden Klassen und die Eltern der teilnehmenden Kinder, die je eine spezifische Beziehung zu ‚Nightingale‘ aufweisen. Sie wurden als ExpertInnen dieser Beziehung betrachtet und in diesem Sinne befragt. Jede/r, der direkt mit dem Programm in Verbindung steht und die Veränderungen durch das Programm abschätzen, bzw. beobachten kann, bringt seine/ihre eigene spezifische Sichtweise und damit seine/ihre Expertise. Dies ist im Rahmen der Konzeptualisierung einer Wirkungsanalyse eine notwendige Basis, um gegebenenfalls weitere Hinweise auf intendierte und vor allem auf nichtintendierte Wirkungen aufnehmen zu können, aber auch, um die Umsetzbarkeit und Akzeptanz aus der je eigenen Perspektive zu prüfen. Als Stakeholder können des Weiteren die FördergeberInnen betrachtet werden. In der vorliegenden Arbeit wurden diese nicht berücksichtigt, da der Schwerpunkt auf der Entwicklung der Wirkungsanalyse aus dem Programm heraus liegt.

Günther Leeb gilt im gesamten Setting als „Türwächter“ oder „Gatekeeper“ (Helfferich 2011, S. 175), er ist die Schlüsselperson für den Zugang zu ‚Nightingale‘. Das bietet den Vorteil einer eindeutigen Ansprechperson, was den Kontakt erleichtert. Gleichzeitig übt die Gatekeeper-Position eine gewisse Kontrolle darüber aus, welche Informationen geteilt werden und für das weitere Vorgehen als relevant eingestuft werden. Es war daher wichtig, zunächst die Zustimmung des Projektleiters von ‚Nightingale‘ Wien für eine Forschung über das Mentoring-Programm einzuholen. Aufgrund seiner Zustimmung war es möglich, Dokumente und weitere Kontaktdaten zu erhalten. So konnte er mir auf Anfrage die Kontaktdaten

der beiden Mitorganisatorinnen zur Verfügung stellen, namentlich die seiner direkten Mitarbeiterin bei der Betreuung der Studierenden der Universitäten, Nadejda Sabeva, und jene der Programmleiterin an der Pädagogischen Hochschule Wien, Elisabeth Leskowa.

Mit beiden Expertinnen führte ich ein Interview durch. Durch das Gespräch mit Elisabeth Leskowa ergab sich im Sinne eines „Schneeballsystems“ (Helfferich 2011, S. 176) die Möglichkeit der Befragung weiterer AkteurInnen. Leskowa bot an, Gespräche mit Mentees, MentorInnen und Eltern zu organisieren. Ich erhielt von ihr die Einladung, beim ‚End-Day‘ an der Pädagogischen Hochschule (Wien, 10.) teilzunehmen und dort die Interviews durchzuführen. Dabei wurde Leskowa ihrerseits zur Gatekeeperin, indem sie mit den AkteurInnen sprach und diese um ein Interview mit mir anfragte. Sie wählte also die Personen aus, mit denen ich sprechen konnte. Im Rahmen des End-Days führte ich sechs Interviews mit jeweils zwei Mentees, zwei MentorInnen und zwei Müttern durch. Zu einem späteren Zeitpunkt organisierte die Programmleiterin der Pädagogischen Hochschule zwei weitere Interviews mit Lehrerinnen, deren Klassen am Projekt teilgenommen hatten. Diese Personen wurden ebenfalls von Leskowa ausgewählt.

Zudem kontaktierte ich die Gründerin, Carina Sild Lönroth, für ein ExpertInnen-Interview in Bezug auf die Organisation der unterschiedlichen Länder und auf den rezenten Beitritt zum ‚European Center for Evidence-Based Mentoring‘. Aufgrund der Teilnahme an der ‚Nightingale‘-Konferenz in Barcelona konnte ich sie durch ein persönliches Kennenlernen auf die Anfrage ansprechen. Ein persönliches Interview war zeitbedingt nicht möglich, Lönroth bot mir an, den Fragebogen schriftlich zu beantworten. Den ausgefüllten Fragebogen erhielt ich via E-Mail.

4.3.1 Durchführung Interviews

Die GesprächspartnerInnen wurden – bis auf die AkteurInnen am ‚End-Day‘ – vorab kontaktiert (E-Mail oder telefonisch) und über das Vorhaben informiert. Bis auf das Interview mit Frau Lönroth fanden alle Interviews persönlich und mündlich statt. Die Durchführung unterschied sich örtlich je nach Verfügbarkeit der befragten Person und fand in den jeweiligen Büros (Leeb, Leskowa, Lehrerin G), bei der Person zu Hause (Leeb, Sabeva) oder in einem Raum der Pädagogischen Hochschule Wien statt (Mutter A und B, Kind C und D, Mentorin E, Mentor F, Lehrerin H). Abgesehen vom ersten Gespräch mit Günther Leeb wurden alle persönlichen Gespräche mit einem Diktiergerät aufgezeichnet. Alle Personen wurden befragt, ob sie namentlich genannt werden möchten. Für Eltern, Kinder, MentorInnen

und Lehrerinnen wurde die Anonymisierung von Beginn an vorausgesetzt. Die OrganisatorInnen zeigten sich mit der namentlichen Nennung einverstanden. Die Einverständniserklärungen der Eltern zur Befragung ihrer Kinder wurden von Elisabeth Leskova mündlich eingeholt.

Tabelle 3 zeigt eine übersichtliche Darstellung der befragten Personen. Die anonymisierten Namen wurden mit der jeweiligen ExpertInnen-Rolle ersetzt, da deren Kenntnis für die Auswertung der Inhalte relevant war. Da jeweils zwei AkteurInnen pro Rolle befragt wurden, wurden zusätzlich Identifikationsbuchstaben hinzugefügt. Die Reihenfolge der Buchstaben entspricht der Reihenfolge der Interviews.

Tabelle 3: Übersicht über die durchgeführten ExpertInnen-Interviews einschließlich Durchführungsdatum und Dauer.

Nr.	Name <i>Quellenverweis</i>	Rolle des Experten/der Expertin	Interview Datum	Dauer des Interviews
1	Günther Leeb	Projektleiter Wien	14. 09.2016	Memo
2	Günther Leeb <i>Int. Leeb</i>	Projektleiter Wien	09.03.2017	01:19:21
3	Elisabeth Leskova <i>Int. Leskova</i>	Projektleitung an der PH Wien und Betreuung der Lehramtsstudierenden	17.05.2017	00:42:03
4	Nadejda Sabeva <i>Int. Sabeva</i>	Mitarbeiterin und Betreuung der Studierenden an den Universitäten	02.06.2017	00:45:50
5	Mutter A <i>Int. Mutter A</i>	Mutter eines teilnehmenden Kindes Jahrgang 2016/17	31.05.2017	00:05:26
6	Mutter B <i>Int. Mutter B</i>	Mutter eines teilnehmenden Kindes JG 2016/17	31.05.2017	00:04:55
7	Kind C <i>Int. Kind C</i>	Mentee (m) JG 2016/17	31.05.2017	00:05:26
8	Kind D <i>Int. Kind D</i>	Mentee (m) JG 2016/17	31.05.2017	00:10:05
9	Mentorin E <i>Int. Mentorin E</i>	Mentorin JG 2016/17	31.05.2017	00:08:47

10	Mentor F <i>Int. Mentor F</i>	Mentor JG 2016/17	31.05.2017	00:24:46
11	Lehrerin G <i>Int. Lehrerin G</i>	Lehrerin an der Mittelschule (PH Wien)	29.06.2017	00:36:19
12	Lehrerin H <i>Int. Lehrerin H</i>	Lehrerin an der Volksschule (PH Wien)	29.06.2017	00:35:24
13	Carina Sild Lönroth <i>Int. Lönroth</i>	Gründerin und Managerin ‚Nightingale‘	28.06.2017	schriftlich

4.3.2 Leitfadengestützte ExpertInnen-Interviews

Die Interviews wurden leitfadengestützt durchgeführt und vorab an die jeweilige ExpertInnen-Rolle angepasst (vgl. hierfür ab Seite 105). Die Gespräche wurden offen durchgeführt, wobei der Leitfaden als Vorbereitung für die Durchführung der Interviews diente. Es wurde besonders auf die Gliederung und Darstellung geachtet, um den Leitfaden als Stütze im Gespräch verwenden zu können, ohne den Kontakt zum/zur InterviewpartnerIn zu verlieren (Bogner et al. 2014, S. 27 ff.). Dies war wichtig, um eine rasche „Vertrauensbeziehung“ (Mayring 2002, S. 69) mit dem/der GesprächspartnerIn aufzubauen. Die Leitfäden für die OrganisatorInnen (siehe hierfür ab Seite 107) sahen eine erzählgenerierende Frage zur eigenen Position im Prozess des Programmes vor. Es stellte sich heraus, dass dieser Einstieg in allen drei Fällen situativ nicht geeignet war, sondern von mir zunächst der Forschungsschwerpunkt und der aktuelle Stand vorgestellt wurde. In diesen Gesprächen fungierten die GesprächspartnerInnen als ExpertInnen bezüglich des Ablaufs des Projektes, weniger hingegen als ExpertInnen von Evaluationsaspekten. So waren nicht alle Befragten mit den (meist psychologischen und quantitativen) Forschungstools vertraut, die ich in das Forschungsdesign eingebaut hatte. Dadurch ergab sich die Schwierigkeit, intuitiv abzuschätzen, wie detailliert ich die GesprächspartnerInnen auf das Thema hinführen sollte und welche Fragen ich in Bezug auf das gewählte Forschungsdesign stellen konnte. Die konkrete Darstellung der Forschungsfrage und des Forschungsdesigns ermöglichte schließlich dennoch ein problemzentriertes Gespräch, das an die konkreten Probleme und offenen Fragen anknüpfen konnte und den Fokus damit auf die Fragestellung fixierte.

Die Leitfäden für Mentees, MentorInnen, Lehrpersonen und Eltern (siehe hierfür ab Seite 115) zielten eher auf die subjektive Meinung und persönliche Erfahrungen bzw. auf das eigene Erleben der Wirkungen des Programms. Das Forschungsdesign wurde im Gespräch weitgehend vernachlässigt. Die Leitfäden für die Interviews mit den Mentees (siehe hierfür Seite 117) überarbeitete ich mit einer Kollegin, welche Erfahrung im Umgang mit Kindern im Volksschulalter hat. Der Leitfaden wurde sprachlich angepasst und es wurden unterschiedliche Antwortmöglichkeiten angeboten. Für Lönroth wurde ein Fragebogen in englischer Sprache ausgearbeitet, der offene, problemzentrierte Fragen beinhaltete. Diesen konnte ich ihr an der Konferenz persönlich übermitteln, sendete das Dokument auf ihre Anfrage auch per E-Mail.

In den teilstrukturierten Gesprächen wurde den AkteurInnen immer die Möglichkeit gegeben, frei zu erzählen. Auch erfolgte eine situative Anpassung der Reihenfolge der Leitfadenfragen und es wurden Rückfragen und ad-hoc-Fragen gestellt.

4.3.3 Transkription

Die mit einem Diktiergerät aufgezeichneten Gespräche wurden im Anschluss transkribiert. Dabei wurde weitestgehend wörtlich transkribiert. Die Transkriptionsregeln wurden an Dresing und Pehl (vgl. 2015, S. 21 ff.) angelehnt und erweitert. So wurden Verständnissignale und Fülllaute (,ähm‘, ,mhm‘) transkribiert und mit einem Bedeutungshinweis (z.B. bejahend, verneinend) versehen, wenn es im Kontext relevant war. Emotionen und Unterbrechungen wurden in Klammern angegeben, ebenso Pausen. Unverständliche Passagen wurden mit Klammern kenntlich gemacht. Besonders betonte Wörter wurden in Großbuchstaben angegeben.

4.3.4 Analyse der Interviews – Qualitative Inhaltsanalyse

Es gibt derzeit kein allgemein gültiges Instrument zur Auswertung von ExpertInnen-Interviews. Grundsätzlich können unterschiedliche qualitative Auswertungsmethoden herangezogen und bei Bedarf kombiniert werden. Die gewählte Methode ist abhängig vom Forschungsdesign und der jeweiligen Fragestellung (Bogner et al. 2014, S. 71–72). Die für diese Arbeit durchgeführten ExpertInnen-Interviews dienten der Gewinnung von Informationen, etwa zu den Abläufen und Prozessen des Programmes sowie zur Möglichkeit der Implementierung einer Wirkungsanalyse. Ferner wurde nach wahrgenommenen Wirkungen gefragt. Hierbei handelt es sich primär um manifestes, konkret formuliertes und direkt zugängliches

Wissen, welches mit der qualitativen Inhaltsanalyse bearbeitet werden kann. Im Gegensatz zu anderen Methoden, die auf die Exploration latenter Sinnstrukturen ausgerichtet sind, akzentuiert diese die Informationen im Text: „[...] das Wissen der Experten wird als eine Ansammlung von Informationen konzeptualisiert“ (Bogner et al. 2014, S. 72).

Die transkribierten Interviews bringen eine große Menge an Textmaterial mit sich. Die qualitative Inhaltsanalyse bezweckt, die Fülle an Material durch Einordnung einzelner Aussagen in ein Kategoriensystem überblickbar zu machen. Diese Einordnung bietet nicht nur eine inhaltliche Bewertung und sichert die Qualität des Materials, sondern ermöglicht eine systematische Reduktion des Textes (vgl. Mayring 2002, S. 114 ff.; vgl. Kaiser 2014, S. 99 ff.; vgl. Gläser und Laudel 2013, Rn. 24 ff.) und kann in weiterer Folge genutzt werden, um Bereiche zu vergleichen, zusammenzufassen, zu generalisieren, oder Verbindungen zwischen den Kategorien herzustellen (vgl. Gläser und Laudel 2013, Rn. 55 ff.). Die Kategorienbildung und das Kodieren sind Techniken, die mit der *Grounded Theory* bekannt wurden und mittlerweile auch außerhalb dieses Methodologienstranges angewandt werden.

In der qualitativen Inhaltsanalyse können zwei Ansätze unterschieden werden: Mayring (2002) und Gläser und Laudel (2013). Während Mayring ein Kategoriensystem aus dem Text heraus entwickelt und dieses dann (nach nochmaliger Prüfung) theorie- und regelgeleitet stringent am restlichen Material anwendet (vgl. Mayring 2002, S. 114 ff.), gehen Gläser und Laudel mit einem aus der Theorie entwickelten Kategorienset an das Material heran: „Sie gehen davon aus, dass gerade durch die Nutzung konzeptioneller Kategorien die theoretischen Vorüberlegungen des Forschungsprojektes während des gesamten Analyseverfahrens Berücksichtigung finden“ (Kaiser 2014, S. 101). Die Kategorien, auch *codes* genannt (*codes* sind „short strings that represent categories and that briefly state what the information is about“ [Gläser und Laudel 2013, Rn. 28]), bleiben in der weiteren Analyse offen und können am Material weiterentwickelt werden. Die Zugänge unterscheiden sich damit hauptsächlich in Art und Weise, wie das Kategoriensystem etabliert wird. Aufgrund der starken Problemzentriertheit des Erkenntnisinteresses und seiner Verbindung zur Theorie bietet sich für diese Arbeit der Ansatz nach Gläser und Laudel zur Analyse der Daten an.

Somit wurde vor der Bearbeitung der Transkripte ein Codesystem angelegt, das sich in erster Linie am Erkenntnisinteresse – an der Forschungsfrage – orientiert (siehe hierfür auf Seite 102). Diese Erstversion bestand neben den unterschiedlichen Codes in Stichwortform aus einem Kürzel und einer oder mehreren Fragen, die an die Forschungsfrage angelehnt waren.

Diese Fragen sollten anhand der Inhalte der Transkriptionen beantwortet werden. Die Transkripte wurden anhand der Kategorien zeilenweise bearbeitet, zugleich wurden bei Bedarf neue Codes gebildet oder bestehende erweitert – diese Offenheit im Prozess war wichtig, um relevante Informationen aus den Inhalten zu extrahieren, die vor der Bearbeitung nicht bedacht wurden (siehe hierfür Seite 103). Die Überschriften 5.2.1 bis 5.4.4 (Ergebnisse) repräsentieren das entwickelte Kategoriensystem in seiner endgültigen Form. Im Anschluss wurden die Aussagen innerhalb der einzelnen Kategorien synthetisiert sowie einzelne Passagen als Originalzitate beispielhaft hervorgehoben.

4.3.5 Kodierung mittels MAXQDA

Die meisten Software-Programme für die Analyse von qualitativen Inhalten, wie das hier verwendete MAXQDA 12, arbeiten mit Codes (vgl. Gläser und Laudel 2013, Rn. 39 ff.). Die Codierung – also angewandte Codes im Text – können dann für die Auswertung des Datenmaterials genutzt werden. MAXQDA wurde für die Analyse qualitativer Daten entwickelt und bietet die Möglichkeit, entweder direkt im Programm zu transkribieren oder Texte zu importieren. Ein Codesystem kann individuell angelegt werden. Die Texte können im Programm gelesen und einzelne Passagen mit ausgewählten Codes verknüpft werden. Dies ermöglicht eine strukturierte, systematische Übersicht. Aus den Codes können im Anschluss *Summary-Grids* erstellt werden, welche eine übersichtliche Möglichkeit für die Paraphrasierung der einzelnen Textstellen darstellen. Aus diesen Zusammenfassungen können später individuelle Summary-Tabellen erzeugt werden, die entweder eine weitere Reduktion des Textes zulassen oder als Übersicht dienen.

4.3.6 Schwerpunktmäßige Auswertung des Onlinefragebogens von ‚Nightingale‘

Wie bereits beschrieben, wurden mir die Daten der Onlinebefragung der MentorInnen des aktuellen Jahrganges 2016/17 zur Verfügung gestellt. Die Angaben aus den qualitativen *Items* des Fragebogens (offene Fragen) wurden in MAXQDA importiert und analog zu den ExpertInnen-Interviews mit Absatznummern dargestellt. Dadurch konnten Teile der Onlinebefragung mittels desselben Verfahrens inhaltsanalytisch ausgewertet und als Ergänzung zu den persönlichen Interviews verwendet werden. Für quantitative *Items* wurden deskriptive Häufigkeitsdarstellungen erzeugt. Der Onlinefragebogen wurde insgesamt von 95 Personen anonymisiert beantwortet, die Zielgruppe umfasste alle ‚Nightingale‘-MentorInnen in Österreich (Wien, Salzburg, Linz und Graz). Konkret bezog sich die Analyse und Auswertung auf folgende *Items* des Onlinefragebogens:

- F6: „Welche der folgenden Kompetenzen hat Dein Mentee durch das Projekt erworben?“ Hierbei wurden nur die Bereiche ‚Mehr Selbstvertrauen‘ und ‚Mehr Motivation zum Lernen‘ beachtet. Dazu gab es jeweils die Antwortmöglichkeiten auf einer vierstufigen Skala von ‚trifft zu‘ bis ‚trifft nicht zu‘, mit der Option ‚weiß nicht‘ (beantwortet von 92, übersprungen von 3 Personen).
- F10: „Von welcher der durchgeführten Aktivitäten hat Dein Mentee am meisten profitiert? Und warum?“ (offene Frage, beantwortet von 90, übersprungen von 5 Personen).
- F29: „Hast du Verbesserungsvorschläge (z.B. zur Organisation und zu den Inhalten der Treffen, zu Berichten, zur Vernetzung etc.)?“ (offene Frage, beantwortet von 68, übersprungen von 27 Personen).
- F31: „Hast du Änderungsvorschläge bezüglich der Organisation (z.B. Training und Information vor Beginn der Treffen, Start-Day etc.)?“ (offene Frage, beantwortet von 60, übersprungen von 35 Personen).

Diese Items sind in Tabelle 4 übersichtlich dargestellt und mit einer Bezeichnung versehen, die im weiteren Verlauf als Quelle angegeben wird. Für die Darstellung aller Items siehe Seite 105.

Tabelle 4: Übersichtliche Darstellung der ausgewählten Items aus den Daten der Onlinebefragung der MentorInnen für das Mentoring-Programm ‚Nightingale‘, Jahrgang 2016/17 (eigene Darstellung).

Nr.	Bezeichnung	Frage	Beantwortet von:	Übersprungen von:
F6	F6 - Kompetenzen	„Welche der folgenden Kompetenzen hat Dein Mentee durch das Projekt erworben?“	92	3
F10	F10 – Aktivität	„Von welcher der durchgeführten Aktivitäten hat Dein Mentee am meisten profitiert? Und warum?“	90	5
F29	F29 – Verbesserung	„Hast du Verbesserungsvorschläge (z.B. zur Organisation und zu den Inhalten der Treffen, zu Berichten, zur Vernetzung etc.)?“	68	27
F31	F31 – Änderung	„Hast du Änderungsvorschläge bezüglich der Organisation (z.B. Training und Information vor Beginn der Treffen, Start-Day etc.)?“	60	35

5 Ergebnisse

Die bereits aufgezeigte Diskussion um Wirkungsmessung und -steuerung in der Sozialen Arbeit/Sozialpädagogik ist bedeutend für das methodische Vorhaben. So gibt es große Divergenzen, ob Soziale Arbeit „standardisierbar, im manageriellen Sinne steuerbar und methodisierbar“ (Polutta 2010, S. 55) sein kann und ob dies zu einer Professionalisierung führen kann oder diese eher beendet. Diese Diskussion steht stark im Zusammenhang mit politischer (und damit zusammenhängend, oft auch finanzieller) Steuerung, beispielsweise im *New Public Management*. Dazu wird hier festgehalten, dass in der geplanten Arbeit keine Veränderungen oder Einschnitte im Konzept des Mentoring-Programmes durchgeführt werden, sowie das Bewusstsein vorhanden ist, dass weit nicht alle Ziele in der Sozialen Arbeit/Sozialpädagogik quantifizierbar und messbar sind. Zudem soll es keinesfalls um ökonomische Zielsetzungen gehen, die dazu führen, dass Creamingeffekte entstehen können, die dazu führen, dass nur noch Kinder angenommen werden, die am ehesten die Erreichung bestimmter Ziele versprechen.

Für eine Auseinandersetzung mit der praktischen Planung einer Wirkungsanalyse wurde als erster Schritt ein literaturbasiertes hypothetisches Forschungsdesign für die langfristige Wirkungsmessung der Stärkung der Bildungsmotivation und des Selbstwertes durch die Teilnahme beim Mentoring-Programm ‚Nightingale‘ entworfen. Die Entwicklung des Designs dient als Basis für die empirische Befragung und Datenauswertung. Es bietet einen Eindruck, was bei einer Wirkungsanalyse bedacht werden muss, darauf aufbauend wurden die Leitfäden (siehe Anhang) und somit die Fragen an die ExpertInnen entwickelt.

5.1 Entwicklung des hypothetischen Forschungsdesigns

Wie schon im Kapitel 4.1.1 beschrieben, wurde der Erstentwurf auf der Fachhochschul-Konferenz in Krems vorgestellt und im Tagungsband publiziert (vgl. Kofler 2017). Da das Forschungsdesign konstitutiv für die weitere Arbeit und die Beantwortung der Forschungsfragen ist, soll hier dennoch näher darauf eingegangen werden. Zudem wurde das Design mit dem Faktor Selbstwert vervollständigt.

Um ein Forschungsdesign für eine Wirkungsanalyse zu erstellen bedarf es konkreter Ziele. Bei der Entwicklung eines Programmes oder einer Maßnahme sollen die Wirkungsziele so gestaltet sein, dass sie messbar und überprüfbar sind. Im vorliegenden Fall gibt es bereits formulierte Ziele, die retrospektiv operationalisiert werden müssen.

5.1.1 Wirkungsziele als SMART-Ziele

Als Basis eines Wirkungsziels stehen zunächst die Fragen im Raum, welche Zielgruppe erreicht werden soll und was das Programm bei der Zielgruppe erreichen oder verändern soll (vgl. Kurz und Kubek 2013, S. 28). Diese Zielsetzungen sind von ‚Nightingale‘ bereits formuliert. Mit dem Fokus auf Bildungsmotivation und Selbstwert sind folgende Ziele hervorzuheben:

- „Bildungsmotivation stärken durch den Kontakt mit einer Studentin [sic!], die den in Österreich höchstmöglichen Bildungsabschluss anstrebt.
- Selbstwert stärken“ (Leeb 2016)

Diese Ziele betreffen die Zielsetzungen für die SchülerInnen (Mentees) und haben somit eine klare Zielgruppe.

5.1.2 Auswahl der Versuchsanordnung und Methode

Um Veränderungen feststellen zu können, bedarf es Messungen zu verschiedenen Zeitpunkten. Wie bereits angedeutet, werden Evaluierungen häufig am Ende eines Programmes durchgeführt, wodurch keine Veränderung durch das Programm erkenntlich gemacht werden kann. Erst eine Datenaufnahme zu mindestens zwei Zeitpunkten ermöglicht es, eine Wirkung nachvollziehen zu können. Darüber hinaus gilt es, das Kontrafaktische (vgl. Plé 2015, S. 28) zu berücksichtigen, d.h. es muss herausgehoben werden, was ohne Intervention bzw. Maßnahme passiert wäre. Daraus geht die Anforderung einer Kontroll- oder Vergleichsgruppe hervor.

Die Arbeit von Plé (2015) bietet einen Überblick über unterschiedliche Forschungsdesigns für eine quantitative Wirkungsanalyse, die Kontroll- bzw. Vergleichsgruppen vorsehen. Dabei ist das experimentelle Design der *Gold-Standard*, das Design mit höchster Reliabilität. Hierfür ist eine randomisierte Kontrollgruppe notwendig, d.h. eine Gruppe aus zufällig gewählten TeilnehmerInnen, die zum Vergleich mit der Zielgruppe herangezogen wird. Da Mentees auf Basis bestimmter Kriterien in das Programm aufgenommen werden, ist es nicht sinnvoll, auf eine randomisierte Kontrollgruppe zurückzugreifen. Vielmehr sollen die Kriterien, auf deren Basis die Mentees ausgewählt worden sind, auch auf die Mitglieder der Kontrollgruppe zutreffen. Das Quasi-Experiment wird von Plé als zweitbestes Design für eine Wirkungsanalyse vorgeschlagen: „Der Unterschied zum Experiment besteht darin, dass anstelle einer Randomisierung, also einer per Zufallsauswahl erfolgende Aufteilung der Ziel-

und Kontrollgruppe, eine statistisch kontrollierte Vergleichsgruppe konstruiert wird“ (Plé 2015, S. 29). Eine Vergleichsgruppe wird also derart gebildet, dass sie die Kriterien der teilnehmenden Gruppe widerspiegelt, wofür die Methode der *Matching on Observables* angewandt werden kann. Dies bedeutet, dass die Vergleichsgruppe aufgrund von bestimmten Kriterien der Zielgruppe gebildet wird, was bereits ab wenigen Vergleichspaaren durchgeführt werden kann (vgl. Plé 2015, S. 28 ff.).

Das Forschungsdesign des Quasi-Experiments kann somit für die Entwicklung einer Wirkungsanalyse für ‚Nightingale‘ verwendet werden, da die Mentees nach vorgeschriebenen Kriterien ausgewählt werden und es somit einer Vergleichsgruppe mit ähnlicher Merkmalsverteilung bedarf. Zur Eruiierung der Wirkung (und damit zum Vergleich zwischen den beiden Gruppen) kann die *Double-Difference-Methode* (vgl. Caspari und Barbu 2008, S. 28 ff.) herangezogen werden. Diese Methode misst die Differenz der zeitlichen Veränderung zwischen Ziel- und Vergleichsgruppe. Da die zeitliche Veränderung die Differenz zwischen zwei Zeitpunkten ist, kann von einer *Double-Difference* gesprochen werden. Somit sind die Voraussetzungen der hier geplanten Wirkungsanalyse erfüllt: eine Aufnahme der Daten zu zwei Zeitpunkten in zwei Gruppen, wobei die Zielgruppe die Maßnahme erfährt, die Vergleichsgruppe nicht. Um der Bedingung der Langfristigkeit gerecht zu werden, wird ein dritter Messpunkt hinzugefügt. Das Forschungsdesign wird damit über drei Erhebungszeitpunkte verfügen. Eine Datenerhebung soll vor dem Programmstart stattfinden, um die *Incomes* (vgl. Beywl 2009) zu erheben, d.h. zu eruiieren, mit welchen Voraussetzungen die Kinder in das Programm starten (M_0). Ein zweiter Messpunkt soll am Programmende liegen, um die unmittelbare Wirkung der Intervention abschätzen zu können (M_1). Die Beobachtung der langfristigen Wirkung macht einen dritten Messpunkt notwendig (M_2). Dafür ist ein – noch nicht definierter – Zeitpunkt nach Programmende geplant.

Die Wirkungsanalyse für die Feststellung der Outcomes von ‚Nightingale‘ wird somit anhand eines Quasi-Experiment-Designs geplant, das eine statistisch gebildete Vergleichsgruppe beinhaltet. Ziel- und Vergleichsgruppe werden zu drei Messpunkten evaluiert und mit der *Double-Difference-Methode* verglichen. Offen bleibt die Frage nach entsprechenden Instrumenten, die eine Messbarkeit der Ziele ermöglichen.

5.1.3 Auswahl der Instrumente

Um Wirkungsziele zu operationalisieren, bedarf es konkreter Instrumente, die eine Messbarkeit möglich machen. Um geeignete Instrumente für die Messung von Bildungsmotivation und Selbstwert zu finden, wurde eine Literaturrecherche (siehe hierfür Tabelle 5) durchgeführt. Ziel war es, bereits bestehende und den wissenschaftlichen Gütekriterien (Objektivität, Reliabilität, Validität) entsprechende Instrumente zu finden, die in einer Wirkungsanalyse angewandt werden können.

5.1.3.1 Bildungsmotivation

Die Bildungsmotivation wird definiert als „persistente Motivation und Wertschätzung für Bildung, Lernen und damit Weiterentwicklung“ (Schober et al., S. 126). Diese Motivation ist nicht (nur) in klassischen Bildungs-Indikatoren wie Schulnoten oder das Erreichen einer höheren Schulstufe zu messen, welche erst in großer zeitlicher Distanz zur Intervention Wirkungen aufzeigen können. Bildungsmotivation ist eine persönliche Einstellung, die sich bereits vor einer Sichtbarwerdung in Noten oder Schulstufe entwickeln und auch in außerschulischen Aktivitäten zeigen kann. Es wird somit ein Instrument benötigt, das Persönlichkeitsveränderung in einer Weise anzeigen kann, die Rückschlüsse auf die Bildungsmotivation zulässt. Hierfür eignet sich das Fünf-Faktoren-Modell (engl. *Big-Five-Model*), eines der prominentesten Instrumente aus der psychologischen Forschung (vgl. Bleidorn und Ostendorf 2009, S. 160). Es wurde ausgewählt, weil es bereits in zahlreichen, auch internationalen Studien angewandt wurde und als robust, valide und reliabel gilt (vgl. Rammstedt und John 2014, S. 35). Auf das Fünf-Faktoren-Modell und dessen Anwendungsmöglichkeiten im Rahmen der Wirkungsanalyse soll nun eingegangen werden.

5.1.3.1.1 Fünf-Faktoren-Modell

Das Fünf-Faktoren-Modell (FFM) wurde anhand eines lexikalischen Ansatzes zunächst im deutschsprachigen Raum entwickelt und durch die Umsetzung in englischer Sprache in den 1960er Jahren bekannt (vgl. Goldberg 1992, S. 26; vgl. Digman 1990, S. 418). Die fünf Faktoren des Modells beruhen auf der Analyse der persönlichkeitsbeschreibenden Wörter einer Sprache. Diese wurden systematisch zusammengefasst (vgl. Rammstedt und John 2014, S. 36; vgl. Rauthmann 2016a) und auf folgende Faktoren reduziert (vgl. Rammstedt und John 2014, S. 35; vgl. Bleidorn und Ostendorf 2009, S. 162) :

- Neurotizismus (emotionale Stabilität)

- Extraversion
- Offenheit/Imagination
- Gutmütigkeit (Verträglichkeit)
- Gewissenhaftigkeit

Das Big-Five-Inventroy-10 von Rammstedt und John (2014) ist in englischer und deutscher Sprache verfügbar, aber erst für junge Erwachsene ab 16 Jahren einsetzbar. Bleidorn und Ostendorf (2009) entwickelten einen FFM-Fragebogen für Kinder und Jugendliche in deutscher Sprache, aufbauend auf ein Instrument zur Persönlichkeitsabbildung von Mervielde und Fruyt (1999): Das *Hierarchical Personality Inventory for Children* (HiPIC) wurde aus 9.000 Beschreibungen des Kindes durch die Eltern entwickelt und ist bei Kinder von sechs bis zwölf Jahren im Fremdbeurteilungsverfahren und ab 12 Jahren im Selbstbeurteilungsverfahren einsetzbar. Für den Fremdbeuricht eignen sich die Eltern und/oder die Lehrpersonen, welche die Kinder der Ziel- und Vergleichsgruppe einschätzen. Der Fragebogen besteht aus 144 *Items* zu 18 Facetten. Die *Items* sind kurz, leicht lesbar und verständlich gebildet (vgl. Bleidorn und Ostendorf 2009, S. 161), wie folgende Beispielitems³ zeigen:

Neurotizismus: Beinhaltet die Skalen Ängstlichkeit (*Item*: „Macht sich schnell Sorgen über etwas“) und Selbstvertrauen („Zweifelt an sich selbst“).

Extraversion: Enthalten sind hier die Skalen Energie („Ist ständig in Bewegung“), Expressivität („Behält Gedanken und Gefühle für sich“, negativ gepolt), Optimismus („Sieht die Dinge von der sonnigen Seite“) und Schüchternheit („Knüpft schwer Kontakt“).

Imagination (Offenheit): Dieser Faktor kombiniert mehrere Faktoren aus vorherigen FFM-Modellen (Intellekt und Offenheit) und beinhaltet Kreativität („Hat originelle Einfälle“), Intellekt („Braucht Zeit um etwas zu verstehen“, negativ gepolt) und Neugierde („Lernt gerne etwas dazu“).

Gutmütigkeit: Der ursprüngliche Faktor Verträglichkeit wurde weiterentwickelt und besteht nun aus Altruismus („Nimmt Rücksicht auf andere“), Dominanz („Spielt den Boss“), Egozentrismus („Kann nur schwer mit anderen teilen“), Entgegenkommen („Hält sich nicht an Vereinbarungen“) und Reizbarkeit („Gerät schnell mit jemandem in Streit“).

³ Diese Beispielitems werden bei Bleidorn und Ostendorf (2009) genannt, eine vollständige Auflistung der Items steht nicht zur Verfügung.

Gewissenhaftigkeit: Der letzte Faktor besteht aus Konzentration, Beharrlichkeit („Hält durch, bis das Ziel erreicht ist“), Ordentlichkeit und Leistungsmotivation („Will in allem herausragen“) (Bleidorn und Ostendorf 2009, S. 162; zit. in Kofler 2017, S. 431 f.).

Zur Durchführungszeit gibt es bei Bleidorn und Ostendorf keine Angaben, der vergleichbare Big-Five-Test von Rammstedt und John mit 10 Skalen ist mit einer Bearbeitungszeit von unter einer Minute angegeben (vgl. Rammstedt und John 2014, S. 35).

Das HiPIC oder Big-Five-Inventar für Kinder und Jugendliche spiegelt sämtliche Faktoren der Persönlichkeitsentwicklung wieder. Dies bietet im Rahmen einer Wirkungsanalyse nun mehrere Vorteile und Möglichkeiten. Zunächst kann die Bildungsmotivation durch die Big-Five-Faktoren erkannt werden. Dafür wird angenommen, dass die Bildungsmotivation eine Voraussetzung für Bildungserfolg, konkreter Schulerfolg, darstellt. Zwei Faktoren korrelieren mit schulischem Erfolg: Offenheit (Imagination) und Gewissenhaftigkeit (vgl. ETS 2012, S. 1 ff.). Diese können somit als Indikatoren für die Bildungsmotivation herangezogen werden. Unter der Annahme, dass die Steigerung der Bildungsmotivation als Persönlichkeitsveränderung bereits vor der Verbesserung der schulischen Leistung erkannt werden kann, lässt sich die intendierte Wirkung anhand dieser Faktoren bereits frühzeitig ablesen (vgl. Kofler 2017, S. 432). Zugleich können mittels der restlichen Faktoren (Neurotizismus, Gutmütigkeit, Extraversion) auch nicht-intendierte Veränderungen respektive Wirkungen gemessen werden, was einen bedeutenden Vorteil gegenüber anderen Instrumenten darstellt.

Das HiPIC bzw. Big-Five-Inventar für Kinder und Jugendliche in deutscher Sprache wird zusammengefasst als reliables Instrument für die Persönlichkeitsbeurteilung von Kindern und Jugendlichen betrachtet und kann somit für die Messung von Persönlichkeitsveränderungen herangezogen werden. Es bietet den Vorteil, dass es auch nicht-intendierte persönlichkeitsbezogene Wirkungen erfassen kann. Wirkungen, die außerhalb der Persönlichkeit liegen, werden vom Fünf-Faktoren-Modell nicht erfasst. So werden die Folgen einer gesteigerten Bildungsmotivation (beispielsweise verbesserte Schulnoten) vernachlässigt. Um diese Indikatoren ebenfalls in die Wirkungsanalyse mitaufzunehmen, wird das FFM mit ausgewählten Instrumenten des *Measurement Guidance Toolkit* ergänzt.

5.1.3.1.2 *Measurement Guidance Toolkit*

Das *Measurement Guidance Toolkit*⁴ (MGT) (Bowers et al. 2016d) ist eine Plattform für Evaluationsinstrumente von Mentoring-Programmen. Namhafte ExpertInnen im Bereich der Evaluierung und Wirkungsmessung von Mentoring-Programmen (David DuBois, Carla Herrera, Ed Bowers und weitere) erkannten, dass zur Evaluierung von Mentoring-Programmen oft die Ressourcen fehlen. Zur Reduktion des Aufwandes bieten sie mit dem MGT Empfehlungen für geprüfte Instrumente, die in der Evaluierung von Mentoring-Programmen einsetzbar sind. Das MGT gründet m.E. in der Forderung von DuBois et al. (2011, S. 80) nach einem unabhängigen Register für Evaluierungen. Es umfasst Evaluationsinstrumente für aktuell sieben Bereiche (die Bereiche werden regelmäßig erweitert), in denen Mentoring-Programme häufig Ziele setzen. Die empfohlenen Instrumente sind erprobt und konnten jeweils bereits positive Outcomes nachweisen. Einer der Bereiche ist als *academics* betitelt und schließt Outcomes zum Verhalten und zur Performanz in der Schule ein (vgl. Bowers et al. 2016b). Darin enthalten sind auch Einstellungen und Verhaltensweisen, die einen Schulerfolg begünstigen.

Das Set zum Bereich *academics* besteht aus drei einzelnen Erhebungsinstrumenten: *academic performance*, *growth mindset for intelligence* und *school connectedness*, auf die im Folgenden kurz eingegangen werden soll. Die *academic performance* (Bowers et al. 2016b) stellt die Beobachtung der Schulnoten dar, weshalb ich den Titel als Schulleistung übersetzen möchte. Bowers et al. weisen darauf hin, dass die Erhebung von Noten sehr oft von politischen Stakeholdern und Fördergebern gefordert wird. Die AutorInnen sehen die Angabe von Schulnoten bzw. deren Selbsteinschätzung bei TeilnehmerInnen von 8 bis 18 Jahren als durchführbar an. Somit ist dieses Instrument an unterschiedlichen Messzeitpunkten, auch nach dem Programmende, anwendbar. Das „growth mindset“ (Bowers et al. 2016c) erhebt den Glauben an sich selbst und an die Möglichkeit, die eigene Intelligenz bzw. die eigenen Fähigkeiten zu steigern. SchülerInnen mit einem ausgeprägten *growth mindset* erkennen die Möglichkeiten ihrer eigenen Entwicklung besser und zeigen Ausdauer und Durchhaltevermögen in der Erreichung ihrer (schulischen) Ziele. Der Fragebogen enthält acht *Items* mit sechs Antwortmöglichkeiten (Beispielfrage: „I don’t think I personally can do much to increase my intelligence“) und ist über die Homepage für nicht kommerzielle

⁴ Das *Measurement Guidance Toolkit* wurde von der Abteilung National Mentoring Resource Center (NMRC) des Office of Juvenile Justice and Delinquency Preventions erstellt. Verwaltet wird es von derselben Stelle.

Zwecke downloadbar. Dieses Instrument ist ab 12 Jahren anwendbar, womit ein Teil der Zielgruppe von ‚Nightingale‘ nicht befragt werden könnte. Zudem ist der Fragebogen in englischer Sprache verfasst. Das dritte Instrument ist die *school connectedness* (Bowers et al. 2016e), womit die Verbundenheit zur Schule abgefragt wird. Der Fragebogen enthält sechs *Items* und fünf Antwortmöglichkeiten, die von den Kindern im Selbstbericht beantwortet werden sollen (Beispielfragen: „I work hard at school“, „I enjoy being at school“). Auch dieser Fragebogen steht zum Download zur Verfügung. Auch hier ist eine Limitation durch die Alterseinschätzung gegeben, so wird er für 11 bis 18-Jährige anwendbar eingestuft und kann somit ebenfalls nicht für alle teilnehmenden Kinder von ‚Nightingale‘ erhoben werden.

Zusammenfassend ermöglicht das Big-Five-Inventar von Bleidorn und Ostendorf ein etabliertes, explizit an Kinder und Jugendliche adaptiertes Instrument für die Erfassung von Persönlichkeitsveränderungen, bei denen zwei der fünf Faktoren mit Bildungsmotivation korrelieren. Darüber hinaus kann das Instrument auch nicht-intendierte Wirkungen in den restlichen drei Bereichen aufzeigen. Die Fragen sind in deutscher Sprache verfügbar und im Selbst- und Fremdbereich bewertbar. Die Instrumente aus dem *Measurement Guidance Toolkit* bieten eine zusätzliche Perspektive auf die (Folgen einer veränderten) Bildungsmotivation mit Schwerpunkt im schulischen Bereich, der für ‚Nightingale‘ eine relevante Rolle spielt. Der Vorteil dieser in Mentoring-Programmen häufig verwendeten Instrumente liegt in der direkten Verfügbarkeit der Fragebögen mit weiterführender Literatur und Anleitung zur Auswertung. Von Nachteil sind die Begrenzung des Anwendungsalters sowie die ausschließliche Verfügbarkeit in englischer Sprache. Dennoch wurden die Instrumente in das Forschungsdesign mitaufgenommen, um die Möglichkeiten und Grenzen der Anwendung in der qualitativen Befragung (ExpertInnen-Interviews) zu thematisieren.

Für die Entwicklung des Forschungsdesigns wurden zwei Wirkungsziele von ‚Nightingale‘ ausgewählt. Hinsichtlich der Stärkung der Bildungsmotivation wurden nun mehrere ausgewählte Indikatoren und Instrumente vorgestellt und besprochen. Im Folgenden soll auf das Ziel der Stärkung des Selbstwertes eingegangen werden und ein adäquates Erhebungsinstrument vorgestellt werden.

5.1.3.2 Selbstwert

Der Selbstwert oder das Selbstwertgefühl wird als „subjektive Bewertung der eigenen Persönlichkeit, die Zufriedenheit mit sich selbst“ (Asendorpf und Neyer 2012, S. 208) angesehen und wird als Persönlichkeitseigenschaft betrachtet. Obwohl die Stärkung des Selbstwerts (engl. *self-esteem*) ein Wirkungsziel zahlreicher Mentoring-Programme darstellt, gibt es keine empfohlenen Instrumente oder Verweise auf der Homepage des MGT. Die Messung des Selbstwertes benötigt daher eigenständige Instrumente außerhalb des MGT-Repositoriums. Diese wurden mittels Literaturrecherche (siehe hierfür Tabelle 5) ausfindig gemacht und sollen in aller Kürze dargestellt werden.

5.1.3.2.1 Selbstwertskalen

Eine verbreitetes Instrument zur Messung des Selbstwertes ist die Selbstwertskala nach Rosenberg, die auch in die deutsche Sprache übersetzt wurde (vgl. Collani und Herzberg 2003). Sie beinhaltet 10 *Items* mit einer vierstufigen Antwortskala (Beispielitem: „Alles in allem bin ich mit mir selbst zufrieden“; „Ich fürchte, es gibt gar nicht so viel, worauf ich stolz sein kann“). Angaben zur Dauer der Durchführung und Altersbeschränkungen konnten nicht ausfindig gemacht werden. Daher wird ein weiteres Instrument zur Messung des Selbstwertes vorgestellt.

Die Multidimensionale Selbstwertskala (MSWS) nach Schütz und Selling (vgl. Daig et al. 2008) wurde für den deutschsprachigen Raum weiterentwickelt. Die MSWS enthält 32 *Items*, die in 10–15 Minuten (inkl. lesen und verstehen) bearbeitet werden können. Dabei sind die *Items* in Frageform entwickelt und werden mit einer 7-stufigen Ratingskala beantwortet (Beispielfrage: „Haben Sie das Gefühl, dass es keinen Bereich gibt, in dem Sie ‚gut‘ sind?“). Die Selbstwertschätzung wird als „zeitlich relativ stabiles Merkmal angesehen“ (Daig et al. 2008, S. 167), dennoch kann es bei Misserfolgen zu kurzfristigen Schwankungen kommen. Die MSWS ist ab 14 Jahren einsetzbar und entspricht den Gütekriterien. Dabei beinhaltet das Testset Normtabellen (geschlechter- aber nicht altersspezifische Unterscheidungen) und Auswertungsschablonen, die eine objektive Auswertung ermöglichen. Die Auswertungszeit wird auf 5–10 Minuten geschätzt (vgl. Daig et al. 2008, S. 166 ff.). Somit ist bei den Indikatoren und Instrumenten für die Erfassung des Selbstwertes zwar die Übersetzung in die deutsche Sprache gegeben, dennoch gibt es auch hier altersspezifische Einschränkungen. Zur Diskussion der Grenzen und Limitationen werden die Selbstwertskalen gleichwohl in das Forschungsdesign integriert. Ergänzend sei angemerkt, dass einige Studie

Korrelationen zwischen einigen Dimensionen des FFM und Selbstwert gefunden haben (vgl. beispielsweise Amirazodi und Amirazodi 2011).

5.1.3.3 Literaturrecherche

Ausgangspunkt der Recherche war jeweils *u:search*, eine Suchmaschine der Universitätsbibliothek Wien. Tabelle 5 zeigt den Prozess der Identifikation von Instrumenten. In einem ersten Schritt wurde ‚Sekundärliteratur‘ recherchiert, die auf die schlussendlich ausgewählten Instrumente hinführte.

Tabelle 5: Übersicht zur Literaturrecherche der Erhebungsinstrumente (eigene Darstellung).

Suchbegriffe (<i>u:search</i>)	Ergebnisse
‚mentoring‘, ‚assessment‘, ‚meta‘	<p>-----▶ Metaanalyse zur Wirkungsmessung im Mentoring (DuBois et al. 2011)</p> <p>↓</p> <p>MGT academics (Bowers et al. 2016d)</p>
‚Psychologie‘, ‚Persönlichkeit‘	<p>-----▶ Lehrbuch Persönlichkeitspsychologie (Asendorpf und Neyer 2012)</p>
	<p>----- Big Five</p> <p>----- Selbstwertskalen</p>
▶ ‚big five‘, ‚Kinder‘	-----▶ FFM (Bleidorn und Ostendorf 2009)
▶ ‚Selbstwertskalen‘	-----▶ Selbstwertskalen (Daig et al. 2008)

5.1.4 Darstellung des hypothetischen Forschungsdesigns

Zur Beantwortung des forschungsorientierten Teilaspekts der Forschungsfrage, unter Verwendung von welchem a) Forschungsdesign, b) welcher Methoden und c) welcher Indikatoren die langfristigen Wirkungen bestmöglich nachgewiesen werden können, wurde nach Instrumenten und Tests zur Messung von Bildungsmotivation und Selbstwert gesucht, die den wissenschaftlichen Gütekriterien entsprechen und bereits bei Kindern und Jugendlichen angewandt wurden. Innerhalb eines a) Quasi-Experiment-Designs, das als bestgeeignetes Design für die Rahmenbedingungen von ‚Nightingale‘ vorzuschlagen ist, sollen Ziel- und

Vergleichsgruppe zu drei Zeitpunkten befragt werden (M_0 , M_1 , M_2). Die drei Zeitpunkte tragen sowohl der Anforderung des Aufzeigens von Veränderungen als auch der langfristigen Beobachtung Rechnung. Die Daten werden anhand der b) *Double-Difference*-Methode ausgewertet, welche die relative Veränderung von Bildungsmotivation und Selbstwert der Zielgruppe im Untersuchungszeitraum aufzeigt. Die c) zu erhebenden Indikatoren sind in Abbildung 1 zusammenfassend dargestellt. MGT und SWS werden in dieser Darstellung als Orientierungshilfe mit jenen Faktoren in Verbindung gebracht, für die Korrelationen zu bestehen scheinen.

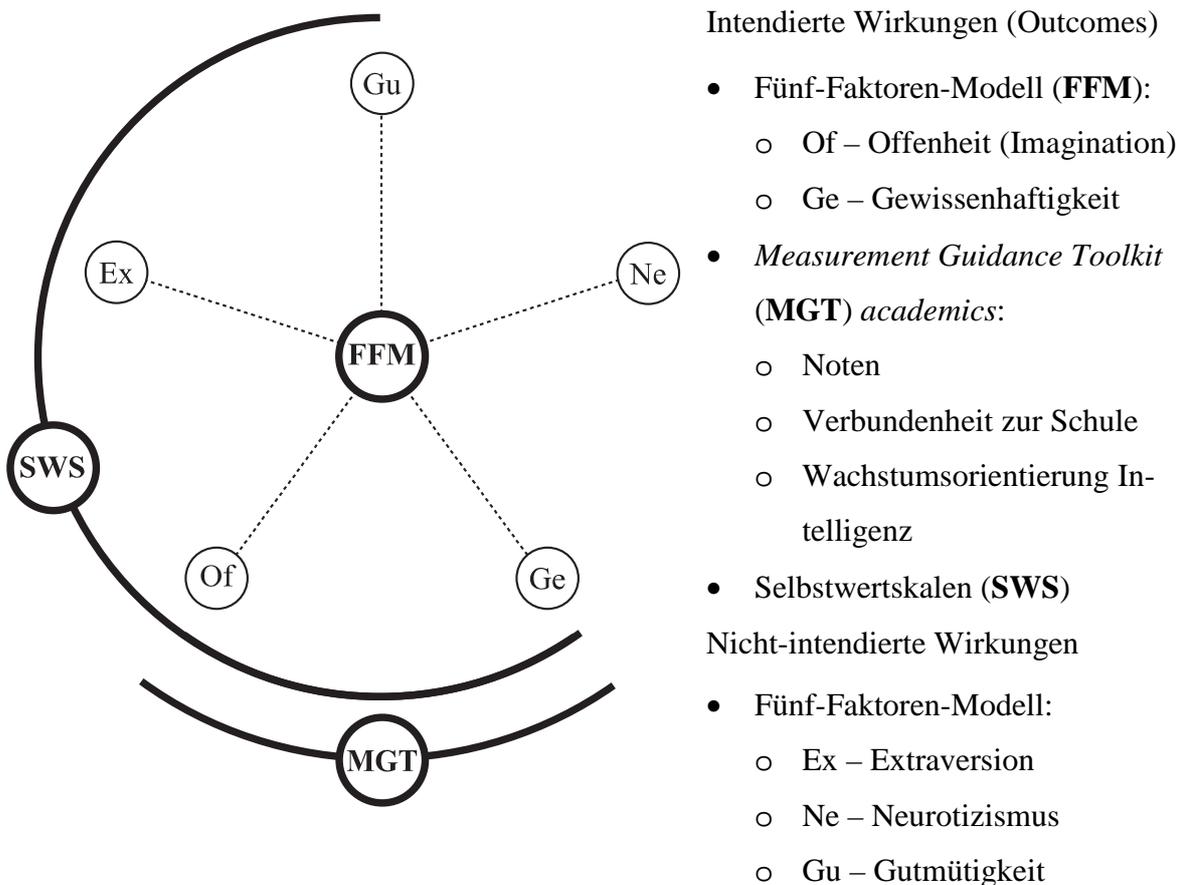


Abbildung 1: Instrumente zur Messung von Bildungsmotivation und Selbstwert (eigene Darstellung, erweitert nach Kofler 2017, S. 443).

Der Messpunkt M_0 liegt kurz vor dem Beginn des Programmes. Die Erhebung zu diesem Zeitpunkt ermöglicht nicht nur eine grundsätzlich bessere Vergleichbarkeit, sondern berücksichtigt auch die *Incomes* (vgl. Beywl 2009). Der zweite Messpunkt ist direkt nach Programmende geplant und als M_1 benannt. Um eine langfristige Wirkung in den Bereichen Bildungsmotivation und Selbstwert feststellen zu können, wurde ein dritter Messpunkt geplant (M_2), dessen Zeitpunkt nicht festgelegt werden konnte. Die in Abbildung 1 dargestellten Indikatoren werden für Ziel- sowie Vergleichsgruppe zu allen drei Messpunkten erhoben

anhand der *Double-Difference-Methode* miteinander verglichen. Abbildung 2 zeigt die doppelte Differenz, die sich durch Gruppen und Zeitpunkte ergeben. Diese doppelte Differenz entspricht dem Outcome, der als Veränderung direkt nach der Intervention ($M_1 - M_0$) bzw. langfristig ($M_2 - M_1$ oder $M_2 - M_0$) sichtbar wird. Aufgrund des Einsatzes einer Vergleichsgruppe kann das Kontrafaktische (vgl. Plé 2015, S. 28) berücksichtigt und damit die Wirkungen „isoliert“ werden, die direkt auf das Programm zurückführbar sind. Mit dem FFM können im selben Verfahren nicht-intendierte persönlichkeitsbezogene Wirkungen festgestellt werden.

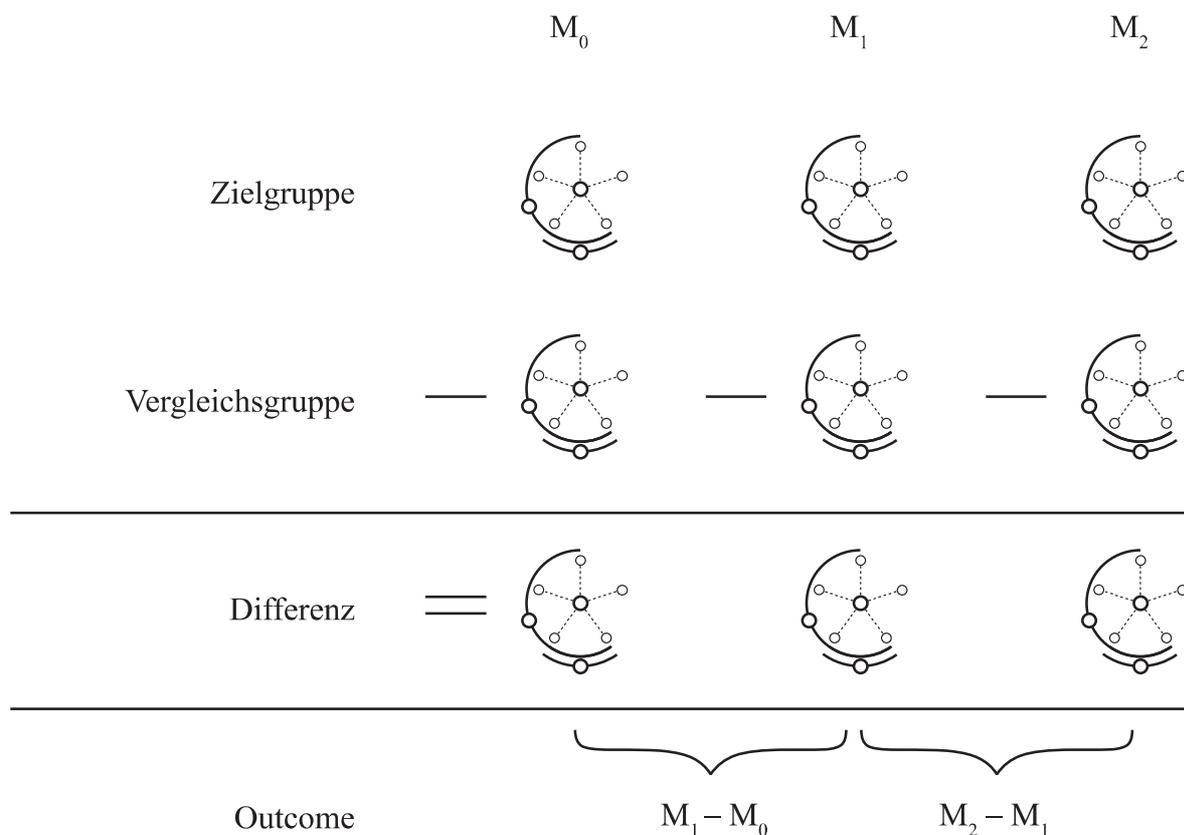


Abbildung 2: Berechnung der Outcomes anhand der *Double-Difference-Methode* (eigene Darstellung, erweitert nach Kofler 2017, S.443)

Anhand dieses Zwischenergebnisses kann überprüft werden, inwieweit die SMART-Kriterien (spezifisch, messbar, akzeptiert, realistisch, terminierbar) bereits erfüllt sind. Die ersten beiden Kriterien scheinen erfüllt. Die Wirkungsziele der Steigerung von Bildungsmotivation und Selbstwert lassen sich vor dem Hintergrund der gewählten Instrumente und Methode recht eindeutig definieren: *Eine intendierte Wirkung ist eingetreten, wenn das teilnehmende Kind eine Steigerung in den jeweiligen Items (bzw. eine Senkung bei Negativ-Polung) ab M_1*

erfährt (vgl. Kofler 2017). In dieser Formulierung sind keine Mindest- oder Höchstwerte bezüglich Steigerung bzw. Senkung angegeben. Auf die Definition von solchen Werten wurde verzichtet, um Creamingeffekte zu vermeiden. Da das Programm darauf abzielt, eine positive Wirkung in beiden Dimensionen zu erreichen, wird davon ausgegangen, dass diese einer Erhöhung (bzw. einer Senkung, wenn negativ gepolte *Items* verwendet werden) der jeweiligen Werte gleichkommt. Die Zielerreichung kann als realistisch (viertes Kriterium) angesehen werden, die angewandten Instrumente bzw. Tests bereits in anderen Kontexten – und ebenfalls mit Kindern – positive Outcomes nachweisen konnten. Offen sind die verbleibenden Kriterien der Akzeptanz und Terminierbarkeit. Ersteres wird nach eigener Definition als Kriterium bezüglich des Wirkungszieles an sich aber auch in Bezug auf die Art der Messung (bzw. auf das Forschungsdesign insgesamt) verstanden. Hierfür ist relevant, mit den Stakeholdern bzw. AkteurInnen des Programmes einen gewissen Konsens zur Sinnhaftigkeit und zu den Möglichkeiten der Umsetzung zu erreichen. Die Terminierbarkeit soll ebenfalls im Rahmen der ExpertInnen-Interviews diskutiert werden, wobei ein Schwerpunkt auf die Langfristigkeit und somit auf den dritten Messpunkt M_2 gelegt wird. Neben den Kriterien der Akzeptanz und der Terminierbarkeit waren noch weitere Fragen und Ideen offen, die in der qualitativen Befragung beantwortet werden sollten:

- Einflussfaktoren auf die Mentoring-Beziehung:
 - Welche Einflüsse sind relevant für die Wirkungsanalyse?
- Vergleichsgruppe:
 - Für das methodische Design muss eine Vergleichsgruppe gebildet werden, wie ist dies umsetzbar?
- Psychologische Instrumente:
 - Ein Teil der ausgewählten Tests ist nicht in der deutschen Sprache verfügbar, dürfen diese übersetzt werden?
 - Die Instrumente sind teilweise altersbegrenzt, die Anwendung im Selbstbericht somit nur zum Teil möglich. Wie ist damit umzugehen? Und damit zusammenhängend:
 - Könnten – auch um eine holistische Sichtweise zu fördern – verschiedene andere AkteurInnen auf die Entwicklung des Kindes befragt werden? Dafür eignen würden sich die MentorInnen, die Eltern und die Lehrpersonen. Wie schätzen diese den dadurch entstehenden Mehraufwand ein?
- Terminierbarkeit:

- Die Langfristigkeit der Beobachtung macht es notwendig, den Kontakt zu den ehemaligen TeilnehmerInnen aufrecht zu erhalten. Wie kann dies gewährleistet werden?
- Wie wird die Möglichkeit eines dritten, verpflichtenden Treffens im Programm gesehen? Wie wird dessen Umsetzbarkeit eingeschätzt?
- Akzeptanz:
 - Wie stehen die OrganisatorInnen zur Wirkungsorientierung?

All diese Aspekte und Fragen wurden in die qualitative Befragung der AkteurInnen miteinbezogen und mit den Angaben der MentorInnen des Jahrganges 2016/17 ergänzt. Kapitel 5.2 bis 5.4 präsentieren die Ergebnisse. Zunächst wurden die Informationen über die Organisation von ‚Nightingale‘ und dessen Evaluierung aktualisiert. Daraufhin wurden die wichtigsten Einflüsse auf die Mentoring-Beziehung, angelehnt an das Rhodes-Model (vgl. Rhodes 2005), thematisiert. Abschließend wird in Kapitel 5.4 die Umsetzbarkeit des Forschungsdesigns kritisch betrachtet und die ausstehenden Kriterien der Terminierbarkeit und der Akzeptanz aufgrund der Aussagen der AkteurInnen beleuchtet.

5.2 Organisation

Dieses Kapitel soll die Angaben aus Kapitel 1.1.2 aktualisieren und vervollständigen. In den Gesprächen mit den OrganisatorInnen (Leeb, Leskowa, Sabeva) wurden einige Veränderungen seit 2014 (vgl. Leeb 2014) festgestellt, die relevant für die Planung und Umsetzung einer Wirkungsanalyse sind. Diese sollen nun vorgestellt werden.

5.2.1 Dauer des Programmes

‚Nightingale‘ Wien dauert pro Durchlauf sieben Monate. Seit der Europäische Integrationsfond 2014 als Fördergeber gefunden wurde und dadurch die finanziellen Ressourcen erhöht wurden, können anstelle der vorherigen 30 Tandems mehr als 70 Tandems organisiert werden (vgl. Int. Leeb, 393). Diese Anzahl ist in einem Durchgang organisatorisch nicht zu bewältigen, daher wird ‚Nightingale‘ seit 2014 in zwei Durchgängen ausgeführt, einmal von Mai bis November und einmal von November bis Ende Mai (vgl. Int. Leeb, 365). Im ersten Durchgang liegen die Sommerferien dazwischen. Im Herbst knüpfen die Tandems dann wieder an und führen ihre Treffen fort (vgl. Int. Leeb, 393). Im Programmjahr 2016/17 fanden Kooperationen mit acht Schulen in Wien statt, die insgesamt 100 Tandems ermöglichten.

5.2.2 Auswahl der Schulen

Die Auswahl der Schulen erfolgt anhand der bereits in der Publikation von Günther Leeb (2014) dargestellten Kriterien, er konkretisiert jedoch weiter: „das Profil sind typische Halbtagsschulen, [...] Schulen in denen Migration, Armut und wenig Bildungskapital zusammenfallen. [...] Das gibt es in Wien [...] überall, außer im 1. oder im 19. oder in Hietzing“ (Int. Leeb, 385).

Tabelle 6 zeigt die Häufigkeit der teilnehmenden Schulen nach Bezirk. Unter den Schulen im 10. Bezirk befindet sich die einzige Mittelschule, die Praxismittelschule der PH Wien, auf die nachfolgend detaillierter eingegangen werden soll.

Tabelle 6: Bezirke und Anzahl der Schulen, die an ‚Nightingale‘ in Wien teilnehmen.

Bezirk	Anzahl der teilnehmenden Schulen
Wien, 2.	3 Volksschulen
Wien, 15.	2 Volksschulen
Wien, 22.	1 Volksschule
Wien, 10.	1 Volksschule 1 Mittelschule

5.2.3 ‚Nightingale‘ an der Pädagogischen Hochschule

‚Nightingale‘ wird seit zwei Jahren an der Pädagogischen Hochschule umgesetzt. Die Kapazitäten sind auf 20 Tandems pro Jahrgang beschränkt (vgl. Int. Leskowa, 124). An der PH Wien herrscht die Besonderheit, dass sich unter den teilnehmenden Schulen eine Mittelschule befindet – alle anderen Schulen sind Volksschulen. Zudem sind die beiden teilnehmenden Schulen der PH Wien, die Praxisvolks- und Praxismittelschule ausschließlich mit angehenden LehrerInnen als MentorInnen besetzt. Die Studierenden haben hier die Möglichkeit, einen Teil ihrer Praxis über ‚Nightingale‘ zu absolvieren (vgl. Int. Leskowa, 124 und Int. Leeb, 215).

Dennoch verläuft das Programm auch an der PH nach den gleichen organisatorischen Abläufen. Einmal in der Woche unternehmen die MentorInnen für ca. drei Stunden eine Aktivität mit der/dem Mentee, die Reflexionstreffen der MentorInnen finden mit Frau Leskowa statt. Elisabeth Leskowa hat den Vorteil, dass sich sowohl die Kinder als auch die Studierenden regelmäßig im Haus befinden und sie diese zum Teil persönlich kennt. So kann rasche Hilfestellung bei Problemen geleistet werden, zudem besteht direkter Kontakt zu den Eltern und Lehrpersonen der Kinder (vgl. Int. Leskowa, 128 ff.). Die Studierenden kennen die Kinder vorab nicht (vgl. ebd., 140).

5.2.4 Auswahl der Mentees

Wie bereits im Kapitel 1.1.2 beschrieben, bestehen Kriterien zur Auswahl der Kinder, die zur Teilnahme eingeladen werden. Hierbei zeigten sich bei den befragten Lehrerinnen an der Pädagogischen Hochschule markante Unterschiede.

Lehrerin H sieht die Auswahl der Kinder als ihre Aufgabe, ohne jedoch konkrete Vorgaben erfüllen zu müssen (vgl. Int. Lehrerin H, 49). Lehrerin H schlägt die Kinder nach eigenen Kriterien vor, wobei ein wichtiges Kriterium

„Bedürftigkeit [ist], Kinder, wo ich weiß, dass der Fernseher zu Hause den ganzen Tag läuft, das erfährt man durch viele verschiedene Interventionen. Kinder, deren Selbstwertgefühl leidet. Kinder, die sehr, sehr interessiert sind am Unterricht aber zuhause den Background dafür nicht haben. Das sind so meine Kriterien.“ (Int. Lehrerin H, 43 ff.).

Hierbei zeigt sich ein deutlicher Unterschied zu Lehrerin G, welche die Auswahl der Kinder gewissermaßen als nicht in ihrer Verantwortung liegend erachtet. Lehrerin G kann aus diesem Grund das Auswahlverfahren bzw. die Kriterien nicht beschreiben: „Naja irgendeinen Grund gibt es natürlich schon, wo ich mir denk, das, aber das habe ich nicht bewertet oder gewichtet“ (Int. Lehrerin G, 44). Der Projektleiter hingegen sieht die Kompetenz zur Auswahl der Kinder – trotz Vorschlag von Kriterien – bei den Lehrpersonen: „Kinder, um die sich die LehrerInnen Sorgen machen [...] die Idee ist, dass sich die LehrerInnen die förderwürdigsten Kinder aussuchen“ (Int. Leeb, Z. 114 ff.).

In der Onlinebefragung der MentorInnen wird mehrfach erwähnt, dass das Auswahlverfahren der Kinder unklar ist oder überarbeitet werden sollte: „Ich bin mir [...] nicht sicher, ob immer die Kinder ins Projekt kommen, die das wirklich brauchen“ (F32, 35). „Ich hatte teilweise das Gefühl dass ich [mein Mentee] zwangsbeglücke“ (F32, 123). Es gäbe genug Kinder, denen die Teilnahme am Projekt mehr bringen würde (vgl. F32, 121), auch höre man

immer wieder von Kindern, „die sehr gerne mitgemacht hätten und für deren Entwicklung das Projekt auch sehr gut gewesen wäre, die aber nicht teilnehmen durften/konnten“ (F32, 53).

Dass die ausgewählten bzw. vorgeschlagenen Kinder teilnehmen können, hängt vom Einverständnis der Eltern ab. Dies stellt sich, wie Günther Leeb bereits in der Evaluierung von 2014 aufzeigt, häufig als großes Hindernis dar (vgl. Int. Leeb 2014, S. 19).

5.2.5 Einverständnis der Eltern

Die Eltern von der Teilnahme bei ‚Nightingale‘ zu überzeugen, zeigt sich als große Herausforderung: „Ich glaube einfach, man lässt sich nicht so gerne ins Wohnzimmer schauen“ (Int. Lehrerin H, 23). Um die Zustimmung der Eltern zu erhalten, ist es wichtig, dass sie das Programm verstehen. Dafür gibt es Informationsmaterial in unterschiedlichen Sprachen und Infogespräche mit DolmetscherInnen (vgl. Int. Leeb 2014, S. 19). Dennoch werden diese Angebot selten genutzt: „Also das Verständnis könnte man sich ja durch die Informationen holen, aber die Lesefaulheit, auch egal in welcher Sprache [...]“ (Int. Lehrerin H, 27). Nach Herrn Leeb spielt es auch eine Rolle, wie viel Zeit und Engagement von Seiten der LehrerInnen aufgebracht wird um Kinder oder Eltern zu überzeugen (vgl. Int. Leeb, 116). Häufig glauben die Eltern nicht, dass das Programm unentgeltlich stattfindet. Erst nachdem sie oder ihre Kinder von den teilnehmenden Mentees erfahren wie das Programm abläuft, wollen sie mitmachen, bzw. sind bereit ihr Einverständnis zu geben. An diesem Punkt ist eine Teilnahme jedoch erst im darauffolgenden Durchgang wieder möglich.

Dennoch gibt es trotz des Wissens um das Programm auch weiterhin Schwierigkeiten, die vorgeschlagenen Kinder bei ‚Nightingale‘ aufzunehmen. Im Jahrgang 2016/17 wurden von Lehrerin H drei Kindern in das Programm eingeladen, diese Kinder erhielten eine Einverständniserklärung für Eltern – nur eine unterschriebene Bestätigung kam zurück. Für die Eltern wäre die Teilnahme am Programm ein weiterer Termin in der meist schon dicht befüllten Woche, zudem stufen die Eltern die Priorität des Programm oft nicht hoch ein, meint der Projektleiter (vgl. Int. Leeb, 116). Auch spielen kulturelle und religiöse Gründe eine wichtige Rolle. Die Eltern wollen ihr Kind nicht mit einer fremden Person alleine durch die Stadt ziehen lassen, wollen nicht, dass das Kind mehr Selbstwert oder Selbstbewusstsein erlangt. Leeb vermutet die Angst, dass das Kind „quer zu ihrem Erziehungsstil beeinflusst wird“ (Int. Leeb, 116).

Die Auswahl der MentorInnen erfolgt zunächst über die freiwillige Bewerbung, dabei kooperiert ‚Nightingale‘ Wien mit diversen Universitäten und Hochschulen.

5.2.6 MentorInnen und Kooperationen mit Universitäten und Hochschulen

Die MentorInnen, die gemäß des Konzeptes von ‚Nightingale‘ Studierende sind, werden über Universitäten und Hochschulen angeworben. ‚Nightingale‘ möchte nicht nur bei den Kindern, sondern auch bei den MentorInnen unterschiedliche Wirkungen erzielen. Die größte Wirkung für die Studierenden bei ‚Nightingale‘ liegt in der Steigerung der interkulturellen und pädagogischen Kompetenz, der Abbau von Vorurteilen spielt eine wichtige Rolle. Vor allem für MentorInnen, die später mit Kindern oder mit dem Thema Migration arbeiten, finden hier wichtige Lerneffekte statt (vgl. Int. Leeb, 195).

Um die MentorInnen auf das Programm aufmerksam zu machen, gibt es Kooperationen mit unterschiedlichen Departements der Universität Wien, mit der Pädagogischen Hochschule und der Universität für angewandte Künste. Günther Leeb baut diese Kooperationen weiter aus – eine zukünftige Zusammenarbeit mit der Fachhochschule Campus Wien ist angedacht (vgl. Int. Leeb, 215).

Besonderheiten gibt es, wie schon beschrieben, bei der Pädagogischen Hochschule, bei der ausschließlich zukünftige Lehrpersonen als MentorInnen tätig sind. Zudem wird über das Departement der Entwicklungspsychologie der Universität Wien ein Masterpraktikum angeboten, bei dem die Teilnahme als MentorIn bei ‚Nightingale‘ ermöglicht wird. Diese Masterstudierenden haben intensivere und konkretere Aufgabenstellung und werden dabei von Nadejda Sabeva betreut. Die Psychologie-MentorInnen treffen ihre/n Mentee häufiger und erarbeiten genauere Zielsetzungen kombiniert mit konkreter Recherchearbeit. Dabei werden die SMART-Kriterien verwendet und nach dem individuellen Fall formuliert. Sabeva gibt als Beispiel „mehr Bewegungsfreude entwickeln“ (Int. Sabeva, 120) an, was bedeutet, dass hier nicht an den sechs formulierten Zielen festgehalten wird, die in Leeb (2016) veröffentlicht wurden. Die von den Psychologiestudierenden erarbeiteten Wirkungsziele sollen im Zeitraum des Programmes erreicht werden. Sabeva meint dazu: „die Ziele waren bis jetzt auch oft so formuliert, dass sie [die MentorInnen] es [...] bei jedem Kind [...] einfach direkt wissen und nicht extra überprüfen müssen“ (Int. Sabeva, 122). Es werden keine diagnostischen Verfahren angewandt, vielmehr wird das Ziel mit Beobachtung (Protokollierung des Verhaltens) festgemacht. Es geht dabei um konkrete Einzelfälle, die detailliert betrachtet werden können.

Die Abteilung Entwicklungsdiagnostik der Universität Wien betreut seit 2015 Masterarbeiten zu ‚Nightingale‘. Zuletzt wurde ein Sprachtest durchgeführt, ein standardisierter Test mit einer Kontrollgruppe aus der gleichen Klasse (Himmelbauer 2016), auf den noch näher eingegangen wird. Weitere Arbeiten behandelten die Problemlösungskompetenzen der MentorInnen (Scherer 2016; Trautsamwieser 2016).

Von den Fördergebern werden aber keine Wirkungsnachweise im Sinne einer Wirkungsmessung verlangt, sondern eine Projektevaluierung mit Finanzbericht.

5.2.7 Fördergeber und verpflichtende Evaluierung

Von den derzeitigen Fördergebern, die aus dem ‚Bundesministerium für Bildung‘, dem ‚Bundesministerium für Europa, Integration und Äußeres‘ und der ‚Kahane Foundation‘ bestehen, wird keine Wirkungsmessung respektive keine Wirkungsanalyse verlangt. Sie verlangen nach Finanzberichte, die aufzeigen wofür das Geld eingesetzt wurde und Sachberichte, in denen beschrieben wird, welche Zielgruppen erreicht wurden, welche Ziele erreicht wurden und welche Schwierigkeiten wie beseitigt wurden. Dafür werden die monatlichen Protokolle der MentorInnen, sowie die Evaluierung erstellt. Teilweise kommen die Fördermittel über Förderungen der ‚Kinderfreunde‘ – durch die finanziellen Mittel des Bundesjugendförderplan und des Familienförderplanes wird ‚Nightingale‘ mitfinanziert (Int. Leeb, 231). Die Pädagogische Hochschule erwartet keine eigene Evaluierung, diese wird für ganz Wien von Günther Leeb durchgeführt.

Um den Fördergebern gerecht zu werden, wurde der „Deutsch-Spracherwerbseffekt“ (Int. Leeb, 373) mitaufgenommen, weil das Integrationsministerium einen Schwerpunkt auf diesen Bereich legt. Dazu meint Sabeva (vgl. 83), dass es sicherlich darauf ankommt, auf was die Fördergeber Wert legen, welche Kriterien diese haben und wie es in der Evaluierung oder weiterführenden Forschung berücksichtigt wird. Durch den Gewinn von Fördergelder des Bundesministeriums für Europa, Integration und Äußeres im Jahre 2014 konnten die finanziellen Ressourcen von ‚Nightingale‘ verdoppelt werden, wodurch, wie bereits beschrieben, doppelt so viele Tandems gebildet werden können. Dennoch reicht das Budget nicht für eine Intensivierung – beispielsweise eine Verlängerung des Programmes – oder weiterführende bzw. langfristige Evaluierung aus:

„Aber wir, also Günther und ich, haben außer der Betreuung des Projekts wenig Ressourcen das Projekt selber zu erforschen, er macht es trotzdem, recht viel, er hat auch viele Daten

gesammelt und wir arbeiten ständig an der Verbesserung, aber das was du meinst, mit langfristig bei dem Effekt der Kinder, das können wir nicht umsetzen“ (Int. Sabeva, 103).

Im nächsten Punkt soll hier nun auf die Evaluierung als status-quo eingegangen werden.

5.2.8 Evaluierung des ‚Nightingale‘-Netzwerkes

Im Netzwerk gibt es standardisierte Angaben, die jedes Mitglied jährlich einreichen muss. Dieser *Annual Report* beinhaltet hauptsächlich die Anzahl der Tandems und teilnehmenden Schulen, oder wie viele und welche Arten von Trainings und Schulungen als Leistung erbracht wurden, „to be able to compare and have experience exchange as well safety-proof the ‚Nightingale‘ concept“ (Int. Lönroth, 2). Für das nächste Jahr ist geplant einen zusätzlichen qualitativen Teil der Evaluierung hinzuzufügen, „to be able to compare the effects and develop new tools for evaluation“ (ebd.). Zusätzlich zu dem aktuellen Jahresbericht unternimmt jedes Mitglied eigenen Evaluierungen. Dies ist notwendig, da zwar alle Länder mit Studierenden arbeiten, aber das Alter oder die Auswahl der Kinder nicht durchgängig gleich sind (vgl. Int. Lönroth, 3).

Auf die Frage, ob es Mitglieder gibt, die langfristige Evaluierungen bereits durchführen, gibt Lönroth (vgl. 4) an, dass ‚Nightingale‘ in Spanien, Norwegen und der Schweiz die Möglichkeit hatten, Wirkungsstudien durchzuführen. Zudem wurde kürzlich bestätigt, dass ‚Nightingale‘ Malmö Forschungsgelder für eine Wirkungsforschung erhalten wird, die ggf. auch in eine vergleichende Studie zwischen den Mitgliedern ausgeführt werden könnte. In allen Fällen ist es abhängig von den finanziellen Möglichkeiten, solche Studien durchzuführen (vgl. Int. Lönroth, 4 f.).

Bei der Teilnahme an der ‚Nightingale‘-Konferenz in Barcelona konnten Informationen zu der aktuellen Wirkungsstudie eingeholt werden. Im Vortrag eines Vertreters des Forschungszentrum ‚Spora Sinergies‘ (2017) wurde die aktuellste Evaluierung von *Rossinyol*⁵ vorgestellt. Mittels qualitativem Ansatz wurden verschiedenste AkteuerInnen (*agents*) individuell und in Fokusgruppen befragt und in einer Inhaltsanalyse ausgewertet (vgl. Spora Sinergies 2017, S. 4). Der Fokus lag dabei auf die Outcomes der Mentees, die im Vortrag vorgestellt wurden. Bezogen auf die hier betrachteten Ziele, zeigt die katalanische Studie auf, dass es eine Stärkung des Selbstwertes bei den Kindern gibt, wobei das Ein-zu-eins-

⁵ *Rossinyol* ist die Bezeichnung für ‚Nightingale‘ in Barcelona.

Setting eine große Rolle spielt. Hinsichtlich der Bildungsmotivation betrachtete Spora Sinergies die *academic performance*, welche sich durch emotionale und soziale Inklusion verbessert (vgl. ebd. 2017, S. 20 ff.). Auch äußert sich Forschungsinstitut zu einer langfristigen Wirkungsanalyse: „Impact evaluation: the big challenge! Some aspects of the programme [sic!] are considered difficult to measure since they are intangibles and produce effects in the long term“ (Spora Sinergies 2017, S. 16).

Lönroth wäre offen für eine vergleichende Wirkungsanalyse zwischen den Ländern:

„I really would like to do a comparison between our different Nightingales/Rossinyols...I think it could strengthened the idea about mentoring, Nightingale especially, and show a broader evidence based results. But of course also show weakness, or potentials that can be developed“ (Int. Lönroth, 5).

5.2.9 Schwerpunkt MentorInnen

Der „Kern der Evaluierung“ (Int. Leeb, 163) von ‚Nightingale‘ Wien ist der Onlinefragebogen der MentorInnen. Dieser liefere vor allem Zahlen – wie viele Kinder besuchten welche Institutionen, wie viele Kinder haben welche Fähigkeiten erlernt (Int. Leeb, 167). Die Monatsprotokolle der MentorInnen analysiert der Projektleiter zusätzlich, auch spielen die Reflexionen mit den MentorInnen, die aufgetretenen Probleme und deren Lösungen eine Rolle: „Die Quelle der Erkenntnis sind die StudentInnen, denen man ja auch zumuten kann, dass sie ganz gut reflektieren können, aus ihrer eigenen Perspektive“ (Int. Leeb, 195). Die MentorInnen beschreiben dabei, was sie gelernt haben und was ihrer Meinung nach der/die Mentee dabei gelernt hat.

Bei den Daten der Onlinebefragung der MentorInnen des Jahrgangs 2016/17 soll kurz auf die Frage 10 eingegangen werden. Die Frage: „Von welcher der durchgeführten Aktivitäten hat Dein Mentee am meisten profitiert? Und warum?“, wurde von 90 Personen beantwortet und von 5 Personen übersprungen. Die vorliegenden Aussagen unterstreichen die Annahme, dass der/die Mentee an Selbstwert dazugewonnen hat. Der Begriff ‚Selbstvertrauen‘ wird acht Mal verwendet, der Begriff ‚Selbstbewusstsein‘ drei Mal. Der Begriff der ‚Bildungsmotivation‘ kommt in den Aussagen der MentorInnen nicht vor, aber es gibt einige Aussagen, dass die Kinder durch Museumsbesuche oder Besuche der Bibliothek an Interesse und Lust auf Museum und Lesen dazugewonnen haben. Dazu erklärt Günther Leeb, dass sich die Studierenden in *Soft-Skills* wie ‚Bildungsmotivation‘ meist eher zurückhaltend äußern, was

laut dem Projektleiter zeigt, dass dieser Bereich nicht von den Studierenden beobachtet wird (vgl. Int. Leeb, 238).

Vergleicht man diese Aussage mit den Werten der Onlinebefragung, bei der Frage F6-Kompetenzen („Welche der folgenden Kompetenzen hat Dein Mentee durch das Projekt erworben“) kann eine Übereinstimmung mit der Einschätzung von Herrn Leeb festgestellt werden. Wie in Abbildung 3 und Abbildung 4 ersichtlich, unterscheiden sich die Angaben über die Einschätzung der erlangten Kompetenzen in ‚Mehr Selbstvertrauen‘ (das als Äquivalent zu Selbstwert betrachtet wird) und ‚Mehr Motivation zum Lernen‘ (das als Äquivalent zu Bildungsmotivation betrachtet wird).

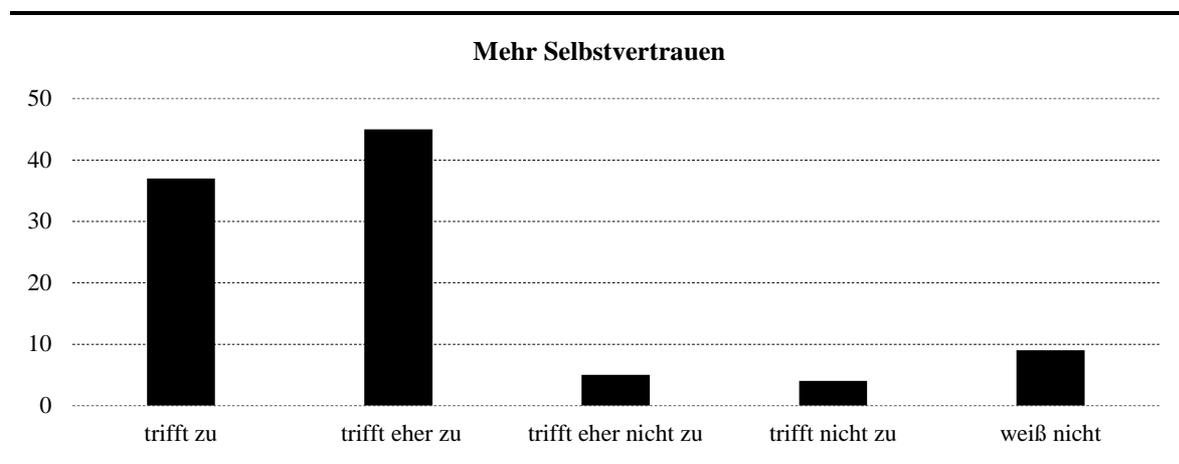


Abbildung 3: Angaben der MentorInnen (n=92) in der Onlinebefragung für ganz Österreich. Frage: Welche Kompetenzen hat dein Mentee durch das Projekt erworben. Antwortoption: Mehr Selbstvertrauen. Eigene Darstellung.

Bei der Frage F6-Kompetenzen antworteten 92 MentorInnen, drei übersprangen die Frage. 37 Prozent der MentorInnen geben an, dass es zutrifft, dass ihr Mentee mehr Selbstvertrauen (ohne genauere Definition des Begriffs) durch das Projekt erworben hat, 45 Prozent geben ‚trifft eher zu‘ an, neun Prozent ‚weiß nicht‘. Bei der Bildungsmotivation sinkt die eindeutige Angabe von ‚trifft zu‘ auf 12 Prozent, ‚trifft eher zu‘ geben weitere 33 Prozent der MentorInnen an. Wesentlich höher ist die Angabenhäufigkeit bei ‚weiß nicht‘ mit 22 Prozent.

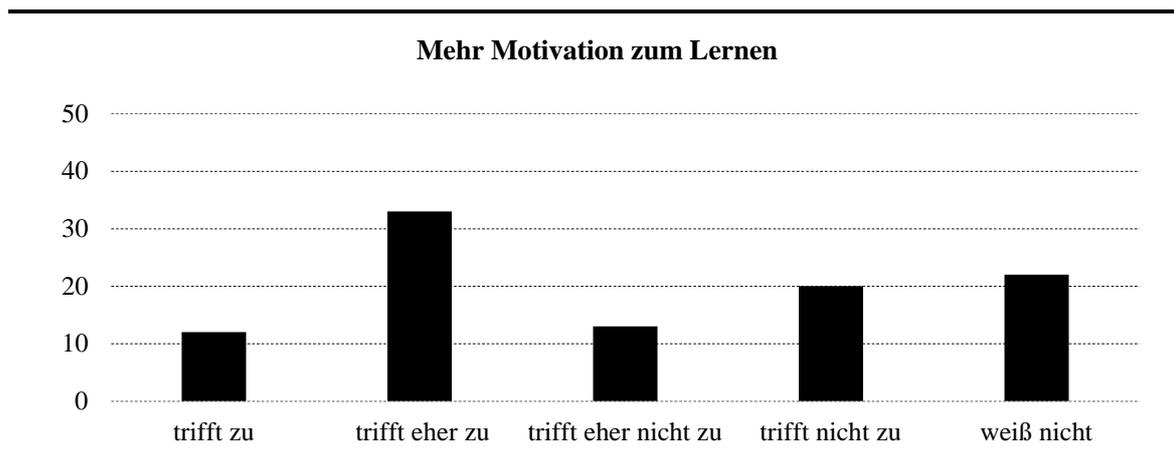


Abbildung 4: Angaben der MentorInnen (n=92) in der Onlinebefragung für ganz Österreich. Frage: Welche Kompetenzen hat dein Mentee durch das Projekt erworben? Antwortoption: ‚Mehr Motivation zum Lernen‘. Eigene Darstellung.

Einige MentorInnen kommentieren bei F10-Aktivität, dass sie nicht einschätzen können, wie ihr Mentee von den Aktivitäten profitiert hat und verweisen teilweise darauf, dass das Kind dies am besten einschätzen könne (vgl. F10-Aktivität, 80).

5.2.10 Befragung von Kinder/Mentees

Bei der Integration der Kinder in die Evaluierung zeigt sich Günther Leeb kritisch. Er betont, dass emotionale Faktoren nicht messbar sind, weil es hierfür reflexive Aussagen der Kinder bräuchte:

„Dinge, die Kinder wahrscheinlich gar nicht wahrnehmen, [...] bewusst, und vielleicht auch gar nicht beschreiben können, die spielen wahrscheinlich eine wichtige Rolle.as ist glaub ich im Mentoring generell so. Da muss man sozusagen schon sehr reflektiert an die Sache herangehen als Mentee, dass man [...] dann beschreiben kann, was hat sich [...] verbessert dadurch dass ich den und den kennengelernt hab“ (Int. Leeb, 119).

Die Mentees zu befragen ist „mühsam“ (Int. Leeb, 171), Leeb beschreibt, dass die Kinder in einem Gruppensetting die Fragen nach dem beantworten, was die Erwachsenen hören wollen, und dass diese immer erwarten, dass etwas gut war und gefallen hat, schon allein, weil das Programm kostenlos ist. Die Kinder antworten meistens eher bejahend, wenn sie erzählen, dann eher faktenorientiert, erzählen dann von ihren Erlebnissen. Aber auch in der Kleingruppe sind die Kinder manchmal zu schüchtern um etwas zu erzählen oder sie verstehen nicht, worum es geht. Die Herausforderung ist, aus den Kindern in dieser kurzen Zeit zu erfahren, wo sie profitiert haben. Da bräuchte es eine Vertrauensposition, man müsste sich Zeit nehmen können, öfter treffen. Sie vielleicht mit bestimmten Situationen konfrontieren,

schauen woran sie sich erinnern. Das ist aufwendig und kann daher nicht durchgeführt werden (vgl. Int. Leeb, 195).

Zudem gibt es sprachliche Barrieren. Die teilnehmenden Kinder sind oft – u.a. durch die Sprache – in ihrer Wahrnehmung und Reflexion eingeschränkt, was auch Lehrerin H anspricht:

„Das Bewusstsein bei den Kindern, die wir auswählen, das sind oft Wahrnehmungen des eigenen Lebens so reduziert. Oder der eigenen Körperlichkeit. Oder des eigenen Vermögens. Sehr reduziert und die Sprache nicht vorhanden um das zu beschreiben, wie es ihnen geht. Und dann gibt es nur die Wörter ‚gut‘, ‚nicht gut‘, ‚ja‘ und ‚nein‘.“ (Int. Lehrerin H, 147)

Die Kinder erhalten am Ende des Programmes einen kurzen Fragebogen, den sie mit den MentorInnen ausfüllen. Dieser Fragebogen enthält Fragen zu den Aktivitäten und zur eigenen Einschätzung, eine der Fragen lautet „Was hast du bei oder durch die Treffen mit deiner Studentin gelernt?“ (Leeb 2014, S. 18). Dieser Fragebogen wird in den meisten Fällen – häufig aufgrund der sprachlichen Barrieren – gemeinsam mit der/dem MentorIn ausgefüllt, wodurch die MentorInnen wiederum Einfluss auf die Beantwortung der Fragen haben (vgl. Int. Leeb, 191).

Auch die Eltern erhalten einen kurzen Fragebogen, den sie am Abschlussfest ausfüllen. Dieser kann in jeder Sprache ausgefüllt werden. Dazu schreibt Leeb (vgl. 2014, S. 19), dass die Feedbacks aus Dankbarkeit meist sehr positiv ausfallen, die Eltern aber teilweise auch konkrete Beobachtungen beschreiben.

Eine regelmäßige Reflexion mit den Kindern findet nicht statt. Während die MentorInnen einmal monatlich zu Reflexionstreffen verpflichtet sind und so durch das Programm begleitet werden, gibt es für die teilnehmenden Kinder keine unterstützende Begleitung während des Programmes (vgl. Int. Leskova, 141). Auch in der Onlinebefragung der MentorInnen wird in einem Kommentar auf die fehlende Reflexion mit den Mentees hingewiesen (vgl. F29-Verbesserung, 12).

Wie im Evaluierungsbericht von Leeb (2014) beschrieben, wurden bereits mehrfach Fokusgruppen mit den Lehrpersonen und eigene Gruppengespräche mit den Kindern durchgeführt. Die Lehrpersonen in die Evaluierung miteinzubinden brachten bereits interessante Erkenntnisse, doch sind die Lehrpersonen aufgrund ihres dichten Terminplans nicht immer bereit, zusätzliche in Gruppendiskussionen oder -gesprächen zur Verfügung zu stehen. Viele Lehrpersonen sind davon überzeugt, dass die Teilnahme am Programm gut für die Kinder ist,

aber können auch Lehrpersonen nicht abschätzen, welche anderen Umstände zur Weiterentwicklung des Kindes beigetragen haben (vgl. Int. Leeb, 203 f.).

Die Befragung der OrganisatorInnen und AkteurInnen, in diesem Fall die Lehrerinnen, bieten somit zusammengefasst ein konkreteres Bild der Prozesse und Organisation von ‚Nightingale‘. Seit 2014 hat sich die Kapazität des Programmes in Wien verdoppelt, es nahmen im Projektjahr 2016/17 einhundert Tandems teil. Dabei wird ‚Nightingale‘ an sieben Volks- und einer Mittelschule in ganz Wien angeboten. Besonderheiten zeigen sich in der Pädagogischen Hochschule, die einerseits die einzige Mittelschule im Programm miteinbringt und andererseits die Teilnahme als MentorIn als Praxisteil der LehrerInnenausbildung angerechnet wird. Des Weiteren besteht eine Kooperation mit dem Masterstudiengang der Psychologie, bei der die MentorInnen sich intensiver und mit eigenen Zielsetzungen den Mentees widmen.

Die Auswahl der Kinder wurde durch die Gespräche mit den ExpertInnen transparenter. Dabei gibt es keine starren Kriterien, die erfüllt werden müssen. Eine Lehrperson sah sich nicht mit der endgültigen Auswahl der Kinder beauftragt, eine weitere Lehrperson entwickelte eigenen Kriterien und erkennt die Auswahl als ihre Aufgabe und Verantwortung. Nach der Auswahl der Kinder stellt sich eine wesentliche Hürde in der Programmorganisation – die Einverständniserklärung der Eltern aufgrund unterschiedlicher Gründe nicht einholbar und damit bleibt den vorgeschlagenen Kindern die Teilnahme am Programm verwehrt. Die Fördergeber verlangen derzeit eine Beschreibung der Tätigkeiten, einen Finanzbericht und eine Evaluierung – diese wird mittels Onlinefragebogen bei den MentorInnen durchgeführt. Die Kinder und Eltern werden in die Evaluierung mittels eines kurzen Fragebogens am Ende des Programmes eingebunden, eine Reflexion mit den Kindern findet nicht statt. Es gab Ansätze, die Kinder stärker in die Befragung miteinzubringen, doch wird den Kindern aufgrund von sprachlichen Barrieren und gesellschaftlichen Erwartungen die Reflexionskompetenz eher abgesprochen.

Bevor auf weitere Ergebnisse eingegangen wird, soll hier noch ein relevanter Punkt in der derzeitigen Evaluierung aufgenommen werden. Seit Frühjahr 2017 ist das ‚Nightingale‘-Netzwerk ein Mitglied des ‚European Center for Evidence-Based Mentoring‘ (ECEB-Mentoring). Zu dieser direkten Verbindung zur Wirkungsorientierung wurde die Managerin Lönroth befragt.

5.2.11 Das ‚Nightingale‘-Netzwerk als Mitglied im ECEB-Mentoring

Wie bereits im Forschungsstand angesprochen, wurde 2016 das erste Zentrum für Wirkungsforschung in Mentoring-Programmen für Europa gegründet. Das ‚European Center for Evidence-Based Mentoring‘ (ECEB-Mentoring) ist in enger Kooperation des ‚Center for Evidence-Based Mentoring‘ in Boston, USA (University of Massachusetts, Boston), das unter Leitung von Jean Rhodes steht. Ziel dieses Zentrums ist die Wirkungsforschung für Mentoring-Programme in Europa zu stärken und zu verbinden: „we seek to the advance in production, dissemination, and uptake of evidence-based practices in ways that improve the effectiveness of practice and, ultimately, create stronger, more enduring mentoring relationships“ (ECEB Mentoring 2017, S. 1). Weiter ist es das Ziel, ein europaweites Netzwerk aufzubauen. Die Wirkungsforschung in Mentoring-Programmen soll gestärkt werden, mit dem Ziel die Qualität von Mentoring-Programmen zu optimieren und deren Wirkung gezielt aufzeigen zu können (vgl. ECEB Mentoring).

Im ‚Nightingale‘-Newsletter vom April 2017 berichtete die ‚Nightingale‘-Gründerin Carina Sild Lönroth, dass das ‚Nightingale‘-Netzwerk nun Mitglied des ECEB-Mentoring ist. Durch die Teilnahme an der ‚Nightingale‘-Konferenz in Juni 2017, konnte ich mit Frau Lönroth in Kontakt treten und sie beantwortete mir einige Fragen zu der Mitgliedschaft. So erklärte sie, dass sie direkt vom Zentrum kontaktiert wurde. Von der Mitgliedschaft erwarte sie sich mehr Kontakt zu der Mentoring-Forschung und zudem direkten Kontakt zu Jean Rhodes:

„I wanted to be a part of this network is to learn more and get insight into new research made in the field of mentoring. But hopefully also get experience exchange with Boston, US where one of the ‚Nightingale‘ partners (the Rossinyol, Girona) now have some research going on with Jean Rhodes“ (Int. Lönroth, 6).

Zugleich wünscht sie sich vom ECEB-Netzwerk mehr über andere Mentoring-Programme in Europa zu erfahren. Lönroth betont, dass Forschung zu Mentoring-Programmen sehr wichtig ist, um Wissen auszutauschen und aufzuzeigen, was im Mentoring funktioniert oder wo es Probleme gibt. Dabei gehe es auch darum, das eigenen Konzept zu verbessern und die hohe Qualität zu optimieren (vgl. ebd. 7 f.). Als Risiko in der EBP sieht Lönroth, dass die Forschung nicht genug Einfluss haben wird, weil die OrganisatorInnen der Mentoring-Programme zu sehr mit dem Programm selbst beschäftigt sind. Sie werden nicht die nötigen Ressourcen haben, sich im ECEB-Netzwerk einzubringen: „The risk is since people do not

know each other, and will maybe not come on every conference, it will fade away and not have the impact it could have had“ (Int. Lönroth, 9).

5.3 Einflussfaktoren auf die Mentoring-Beziehung

Angelehnt an das Mentoring-Model von Jean Rhodes (Rhodes 2005), auf das bereits im Kapitel 5.1. eingegangen wurde, wurden die möglichen Einflüsse auf die Mentoring-Beziehung für ‚Nightingale‘ ausgewertet und erweitert. Im Folgenden wird auf die *Incomes* der MentorInnen und Mentees, die Dauer des Programmes und der Faktor ‚Familie und Umgebung‘ eingegangen.

5.3.1 *Incomes* MentorInnen

Mit welchen Voraussetzungen und Interesse die MentorInnen in das Programm kommen, ist äußerst relevant für die Mentoring-Beziehung, denn die Studierenden bringen einen eigenen Hintergrund mit, durch den sie unterschiedlichen Einfluss auf die Kinder haben können (vgl. Int. Lehrerin H, 153). Die erzielte Wirkung kommt u.a. auf die MentorInnen an, die sehr unterschiedlich sind und unterschiedliche Voraussetzungen und Erfahrungen mitbringen. Auch in Bezug auf deren eigenen Einschätzung und das, was sie dem Kind zutrauen, kommt es zu qualitativen Unterschieden. Dabei gibt es MentorInnen, die „ein bisschen versinken in Hoffnungslosigkeit [...] und da kann ich mir vorstellen, dass die Effekte sehr viel kleiner sind auf die Kinder“ (Int. Sabeva, 213), zugleich gibt es gibt bestimmt MentorInnen, die zu einem Thema ganz viel beitragen und demnach größere Wirkungen erzielen können (vgl. ebd.).

5.3.2 *Incomes* Mentees

Die *Incomes* der Mentees spielen ebenfalls eine relevante Rolle im Mentoring-Prozess. Diese werden mit der Auswahl der Kinder beachtet und protokolliert. Dabei sind die Entwicklungsstufe und Entwicklung des Kindes relevant, diese wurde in mehreren Fällen angesprochen. Mentor F beschreibt: „Ich konnte [...] beobachten, am Anfang des Programms, war er sozusagen Kind und jetzt [...] geht er schon Richtung Jugendlicher“ (Int. Mentor F, 27). Hierbei ist es sicherlich relevant, welches Alter das Kind bei Beginn des Programmes hat. Durch die Kooperation mit einer Mittelschule erhöht sich das Alter der Kinder auf 11 bis 13 Jahren. Dieser Sprung in der Entwicklung des Kindes wirkt sich auch auf die Mentoring-Beziehung aus: „Je älter Kinder werden, quasi von 11 bis 13, diese Altersgruppe, [desto] schwierig[er] [sind sie] zu begleiten“ (Int. Leskowa, 168).

Nach Leskowa erfolgen Abbrüche von Seiten der Mentees eher in der Mittelschule, weil die Zeit lieber mit den Freunden oder Familie verbracht wird, bzw. der Abbruch von Seiten der Familie unterstützt wird (vgl. ebd. 263). Wie alt das Kind ist und in welcher Entwicklungsstufe es sich befindet, beeinflusst die Mentoring-Beziehung. Um die natürliche Entwicklung des Kindes in der Wirkungsanalyse berücksichtigen zu können, bedarf es daher konkreter Aufzeichnungen bzw. Einschätzungen der Entwicklungsstufe des teilnehmenden Kindes. Diese könnten den zusätzlichen Mehrwert schaffen, eine konkrete Basis für das *Matching* darzustellen, „wenn man die Zuteilung von Mentee zu Student nicht blind macht, sondern durch einen Fragebogen beispielsweise nach Gemeinsamkeiten sucht [...]“ (F31-Änderung, 29).

5.3.3 Dauer des Programmes

Die Dauer des Mentoring-Programmes stellte einen wichtigen Faktor für die Wirkungsanalyse dar. Insgesamt liegt der Tenor der OrganisatorInnen und AkteurInnen einstimmig darauf, dass das Programm länger als die derzeit sieben Monate dauern sollte. Der Projektleiter meint dazu: „Der Zeitraum der Intervention ist ja nicht besonders lang“ (Int. Leeb, 100) und „[um] bestimmte verhaltenspsychologische Wirkungen erzielen zu können müsste es länger dauern, ein, zwei Jahre“ (Int. Leeb, 207). Auch die Lehrpersonen sprechen die kurze Dauer des Programms an, Lehrerin H (vgl. Int. Lehrerin H, 155) schätzt die Nachhaltigkeit des Programmes besser ein, wenn es bis zu vier Jahre andauern würde, Lehrerin G (vgl. Int. Lehrerin G, 94) spricht sich ähnlich dessen für zwei oder drei Jahre Begleitung durch ‚Nightingale‘ aus. Entsprechend relevant ist, ob die Tandems auch nach Beendigung des offiziellen Programmes in Kontakt bleiben und die Mentoring-Beziehung weiter aufrechterhalten. Da dieser Punkt eine Evaluierung am Ende des Programmes übersteigt, sind diese Perspektive und die Berücksichtigung dieses Faktors nur in einer langfristigen Betrachtung möglich.

5.3.4 Familie und Umgebung

In der Organisation des Programmes und während der Mentoring-Beziehung haben die Familie und das Umfeld großen Einfluss, wie mit der oftmals fehlenden Einverständniserklärung der Eltern bereits aufgezeigt werden konnte. Doch auch während des Programmes ist die Akzeptanz und Mitarbeit der Eltern konstitutiv für eine potentielle Wirkung.

„Beim Mädchen [...] hat es auch irgendwie nicht ganz gut funktioniert mit dem wir holen ab oder wir treffen uns. Das hat aber weniger mit ihr persönlich zu tun, sondern einfach mit der

Situation, die dieses Mädchen zu Hause hat. [...] Aber es ist natürlich sehr schwierig, Kinder, wie soll ich sagen, einzuladen oder abzuholen, wenn sie nicht vor Ort sind. Auch wenn sie das noch so gerne möchten, aber wenn es da Eltern sind, die da nicht mit dabei sind, wird es schwierig“ (Int. Lehrerin G, 56).

Wie bereits beschrieben, hängen auch Abbrüche häufig mit der Familie und dem Umfeld zusammen: „Fußball, Training, Bruder, Familie ist wichtiger als fortzugehen mit einem, den man nicht mal so gut kenn“ (Int. Leskova, 236) sind gerade bei älteren Mentees häufige Gründe.

Die genannten Einflussfaktoren ‚Familie und Umgebung‘, die Dauer des Programmes und die *Incomes* der MentorInnen und Mentees sind elementare Einflüsse auf die Outcomes eines Programmes.

5.4 Umsetzbarkeit

Das Forschungsdesign, wie es in Kapitel 5.1.3.2 vorgestellt wurde, ist eine Antwort auf die Forschungsfrage, wie die Wirkungen von ‚Nightingale‘ in der Stärkung der Bildungsmotivation und des Selbstwertes nachvollzogen werden könnten. Allen voran benötigt es hierbei konkrete Wirkungsziele, die nach SMART-Kriterien formuliert sind. Da diese Kriterien bereits die Operationalisierung der Ziele beinhalten, musste zunächst auf die Methodik eingegangen werden. Dabei wurde nach einem Forschungsdesign (Quasi-Experiment), nach einer Methode (Double-Difference) und nach passenden Instrumenten gesucht, die den Gütekriterien entsprechen (Big-Five-Inventar, *Measurement Guidance Toolkit*, Selbstwertskalen). Diese Fragebögen sind dem Ansatz der Langfristigkeit entsprechend an drei Zeitpunkten abzufragen, vor dem Programm, nach dem Programm und einen unbestimmten Zeitraum nach Beendigung des Programmes. Dabei sollen unterschiedliche Sichtweisen miteinbezogen werden: Mentees, MentorInnen, Lehrpersonen, Eltern. Offen bleibt die Frage, ob dies in der Realität so umsetzbar wäre. Wo sind die Möglichkeiten und die Grenzen eine langfristige Wirkungsanalyse bei ‚Nightingale‘ Wien durchzuführen? Darauf wird nun eingegangen.

5.4.1 Vergleichsgruppe

Für das Quasi-Experiment ist eine Vergleichsgruppe notwendig. Eine solche Vergleichsgruppe zu erstellen und an mehreren Zeitpunkten zu befragen, wurde bereits im Sprachtest der Universität Wien durchgeführt. Der Sprachtest erstellte eine Vergleichsgruppe aus Kindern derselben Klasse. Lehrerin H konnte hierbei erste Erfahrungen sammeln, sie wählte im

Jahrgang 2016/17 ein Kind für die Vergleichsgruppe und erzählt von ihren Unsicherheiten: „Das war so meine größte Schwierigkeit, weil die Kinder ja so unterschiedlich sind und jetzt soll ich ein Vergleichskind [auswählen], [das eine] Sprachschwäche genauso wie [das teilnehmende Kind] hat“ (Int. Lehrerin H, 68). Lehrerin H fokussierte sich in der Auswahl des Kindes für die Vergleichsgruppe auf die Leistungsfähigkeit und die sprachlichen Fähigkeiten, betonte aber, dass es in den Charakterzügen der Kinder große Unterschiede gab (vgl. ebd.).

Günther Leeb bemerkt, dass es relevant ist, die Familienverhältnisse zu berücksichtigen, dass dies aber nahezu unmöglich ist (vgl. Int. Leeb, 100). Auch Lehrerin G spricht dies an. Die familiären Verhältnisse seien nie richtig vergleichbar. Auch wenn die Schulen im 10. Bezirk hauptsächlich Familien aus dem türkischsprachigen Raum besuchen, unterscheiden sich diese enorm (vgl. Int. Lehrerin G, 126).

Übereinstimmend antworteten die OrganisatorInnen und die LehrerInnen, dass die Vergleichsgruppe in derselben Klasse gebildet werden sollte, „weil dann die handelnden Personen gleich sind“ (Int. Lehrerin G, 124) und da diese Kinder „noch am ehesten Gemeinsamkeiten“ (Int. Sabeva, 143) haben.

Ob die Bildung und Befragung einer Vergleichsgruppe an den Schulen grundsätzlich möglich ist, und in der Praxis von den Eltern ermöglicht wird, bestätigt Lehrerin G für die Praxisschulen: „Also in der Praxisschule ist eigentlich beforschen kein Thema. Also dass Sie das irgendwie vorher einholen müssen, dass Sie das dürfen oder wie auch immer, das, ja.“ (Int. Lehrerin G, 122). Sabeva vermutet in der Einholung der Erlaubnis der Eltern keine Schwierigkeiten, deutet aber auf den erhöhten Aufwand hin: Doppelt so viel Kinder müssen ausgewählt werden, doppelt so viel Einverständnisse eingeholt werden: „Manche Familien sind da überfordert mit dem und sind sehr unzuverlässig. Das ist eher das Problem. Das ist ein richtiger Aufwand für eine Lehrerin“ (Int. Sabeva, 159).

5.4.2 Psychologische Instrumente

„Nightingale“ und vergleichbare Mentoring-Programme zielen oft auf Persönlichkeitsveränderung ihrer Zielgruppe ab. Die Beobachtung, Messung und Vergleichbarkeit von Veränderungen in diesem Bereich stellt eine der größten Herausforderungen dar. Während die Veränderung von Noten, der Übergang in eine andere Schulstufe oder das Kennenlernen von

Einrichtungen quantifizierbar und messbar ist, stellt sich bei der Veränderung der Bildungsmotivation und des Selbstwertes die Frage, wie diese Faktoren operationalisierbar gemacht werden können.

In der Entwicklung des Forschungsdesigns wurde nach Instrumenten gesucht, die für eine Feststellung von Veränderungen in den Bereichen Bildungsmotivation und Selbstwert anwendbar sind. Dabei war eine Auseinandersetzung mit psychologischen Instrumenten unumgänglich. Hierbei stellte sich bereits die Frage, ob diese Instrumente ohne psychologische Fachkenntnisse in einer Wirkungsanalyse angewandt werden können. Diese Frage und weitere Fragen zur Planung, dem Einsatz und der Auswertung von psychologischen Test wurden mit der diplomierten Psychologin und Mitarbeiterin bei ‚Nightingale‘, Nadejda Sabeva besprochen.

Psychologische Tests dürfen nach Sabevas Einschätzung – je nach Test und bei gesunden ProbandInnen – auch von nicht psychologisch ausgebildeten Personen durchgeführt werden. Ausgeschlossen sind dabei Intelligenztests, die einer besonderen Schulung bedürfen (vgl. Int. Sabeva, 69). Auswertung und Interpretation der Ergebnisse bedürfen Expertise und Erfahrung. Hierfür würde sie empfehlen, eine/n ExpertIn hinzuzuholen (vgl. Int. Sabeva, 63). Dennoch ist Sabeva der Meinung, dass es durchaus möglich wäre, im Rahmen eines Forschungsdesigns für eine Wirkungsanalyse die Tests ungeschult durchzuführen, auszuwerten und zu interpretieren.

Zu achten ist bei der Auswahl der Tests auf das Alter der Normierung, dieses sollte nicht älter als 7 bis 8 Jahre sein. Sollte der Test älter sein, können die Fragen dennoch angewandt werden, aber sie empfiehlt, die Normgruppe dann nicht zu beachten, sondern eher auf die qualitativen Aussagen der Kinder einzugehen (vgl. ebd. 65).

Bezugnehmend auf die teilweise nur in englische Sprache zur Verfügung stehenden Instrumente, erklärt Sabeva, dass diese nicht anwendbar wären, bzw. deren Übersetzung enorme Ressourcen benötigen würde. Eine Adaptierung von Tests in andere Sprachen kommt einer eigenen Forschung gleich, bedarf eines professionellen Dolmetschens und muss in mehreren Durchläufen in der neuen Sprache getestet werden. Der Fragebogen ist genauestens auf die Dimensionen und *Items* in der jeweiligen Sprache aufgebaut, eine einfache Übersetzung würde die Ergebnisse verzerren und die Messbarkeit wäre nicht mehr garantiert (vgl. Int. Sabeva, 77 ff.). Dies ist nicht die einzige sprachliche Hürde. Der Projektleiter berichtet, dass

er bereits mit den Professorinnen der Universität Wien, Abteilung Entwicklungspsychologie über die Evaluierung von beispielsweise dem Selbstwert der Kinder gesprochen hat:

„da hab ich erfahren, es gibt keinen Test dafür, für diese Altersgruppe, der nicht ganz massiv auf den Sprachkenntnissen, also Deutschsprachkenntnissen, aufbaut. Da musst du [...] selbstanalytische Aussagen treffen und das ist [...] für Erwachsene auf Dauer schon gar nicht mehr so einfach, für ein Kind [mit] Zweitsprache [Deutsch] sozusagen nicht möglich, da kommt nichts Seriöses raus“ (Int. Leeb, 100).

Bezüglich der Vorgaben zum Mindestalter der Befragten, unterscheidet Sabeva zwischen einem streng wissenschaftlichen Ansatz und einer Evaluierung – je nach eigenem Ansatz und Zweck der Erhebung. Für ersten müssen alle wissenschaftlichen Gütekriterien erfüllt sein, ist dies aber nicht das Ziel der Untersuchung, können die Kriterien aufgeweicht sein. Es ist dabei wichtig, sich zu überlegen, was für die eigenen Befragung Sinn macht, dann ist das Alter, wenn es nun ein oder wenige Jahre Unterschied ausmacht, eher vernachlässigbar (vgl. Int. Sabeva, 81). Es ist zu überlegen, welche Kriterien und welche Vorgehensweisen den Zweck erfüllen und ob eine Vergleichbarkeit angestrebt wird, beispielsweise bei Langzeitstudien. Ist dies der Fall ändert sich natürlich die Herangehensweise (vgl. ebd. 83).

Bezüglich der Selbst- und Fremdevaluierung meint Sabeva, dass verschiedene Sichtweisen immer von Vorteil sind, aber es auch immer eine Ressourcenfrage ist. Bei der Fremdevaluierung durch die Eltern verweist Sabeva auf ihr bekannte Studien (bezugnehmend auf ihre Diplomarbeit, siehe hierfür Sabeva 2012), die aufzeigen, dass bei Befragungen zu der Persönlichkeit des eigenen Kindes, Eltern häufig ihre eigenen Persönlichkeitseigenschaften angeben, bzw. diese im Kind erkennen. Gerade aus diesen Gründen eignet sich eine Einschätzung von verschiedenen AkteurInnen (vgl. Int. Sabeva, 167).

5.4.3 Terminierbarkeit – dritter Messpunkt

Wie im Forschungsdesign (Kapitel 5.1) bereits dargestellt, braucht es für die Wirkungsanalyse klar formulierte Zielsetzungen, die operationalisierbar und vergleichbar sind und den SMART-Kriterien entsprechen. Hierfür wurden bereits einige Kriterien erfüllt, es wurde darauf eingegangen, dass die Zielsetzung bei der Stärkung der Bildungsmotivation und des Selbstwertes durch die Auswahl der Methoden und Instrumente spezifisch (S), messbar (M) und realistisch (R) wird. Offen bleiben die Akzeptanz (A) und die Terminierbarkeit (T) der Wirkungsziele. Auf diese soll nun mittels Auswertung der Befragung eingegangen werden.

In der Entwicklung des Forschungsdesigns für die Wirkungsanalyse wurde der Zeitpunkt des dritten Messpunktes für die langfristige Betrachtung nicht definiert. Hierfür sollte in der qualitativen Befragung von Seiten der unterschiedlichen ExpertInnen eingeschätzt werden, wann eine dritte Befragung – nach Programmende – realistisch und möglich wäre. Bei der theoretischen Planung für eine langfristige Wirkungsanalyse ergeben sich einige Möglichkeiten und Herausforderungen, auf die zusätzlich eingegangen werden soll.

Bei der Befragung ergaben sich unterschiedliche Ansichten bezüglich eines dritten Messpunktes. Beide Lehrpersonen und Frau Leskowa sehen eine neuerliche Befragung der Kinder wenige Monate nach den Sommerferien als sinnvoll. Lehrerin H spricht von vier Wochen nach den Sommerferien (vgl. Int. Lehrerin H, 139 ff.), Lehrerin G von ca. drei Monaten (November, Dezember) nach den Ferien (vgl. Int. Lehrerin G, 110 ff.) und auch Frau Leskowa sieht eine neuerliche Befragung der Kinder und Eltern im Zeitraum von drei Monaten bis spätestens sechs Monaten nach den Ferien als realistisch (vgl. Int. Leskowa, 226). Die drei Expertinnen beziehen sich dabei vor allem auf die Erinnerungsfähigkeit der Kinder. Im neuen Schuljahr (vier Wochen nach den Ferien) oder kurz nach den ersten Schularbeiten (drei Monate nach den Ferien), könnten die Kinder einen Vergleich ziehen, wie es ohne den/die Mentorin ist:

„Vor Weihnachten ist der zweite Schwung an Schularbeiten schon erledigt und da können Kinder sehr gut abschätzen, wie tue ich mich jetzt, wie geht es mir jetzt. Und da passiert der Vergleich vielleicht noch, wie ist es mir voriges Jahr gegangen“ (Int. Lehrerin G, 114).

Sabeva erklärt dazu, dass die Erinnerung bei vielen psychologischen Tests nicht vorhanden sein muss:

„Wenn man weiß, das sind drei [Messpunkte], dann muss man das so konzipieren, dass man nicht erfragt, woran sich die Kinder erinnern, sondern wie geht es ihnen [...] wie selbstsicher sind sie, wie gut kennen sie sich aus, wie viele Kulturangebote nehmen sie in Anspruch. Wie fühlen sie sich in der Schule. Wie viel trauen sie sich zu oder verfolgen sie eine eigene Stärke oder Talent, oder solche Sachen. Und da brauchen sie sich an gar nichts erinnern“ (Int. Sabeva, 190).

Sabeva selbst sieht einen geeigneten Zeitpunkt dann, wenn die Kinder 14 Jahre alt sind, „weil sich [dann] im österreichischen Bildungssystem gezeigt hat, sind sie wo angekommen oder nicht“ (Int. Sabeva, 185). Je nachdem, welche Klasse und Schule der/die ehemalige Mentee zu seinem/ihren vierzehnten Lebensjahr besucht, ergäbe sich die Herausforderung, den Kontakt zu den Kindern respektive Jugendlichen aufrechtzuerhalten. Die vorhandenen

Adressdaten der Kinder, die während des Programmes aufgenommen werden, sind unzuverlässig, die Telefonnummern der Eltern ebenso, zudem wird die Schule gewechselt. Nach Sabeva wäre es gegebenenfalls möglich, die Eltern, die ohnehin im Vorherein ihr Einverständnis für eine Langzeituntersuchung abgeben müssten, nach einer Übermittlung der Informationen zu bitten, wenn ihr Kind die Schule wechselt. An der Schule könnte man eine Ansprechperson einrichten, die den Kontakt ermöglicht (vgl. Int. Sabeva, 195 ff.). Dennoch schließt Sabeva:

„Grundsätzlich denke ich, dass gerade diese Kinder, die so viele Probleme mitbringen und zu wenig Förderung erfahren, [...] mehrere solche Maßnahmen [bräuchten]. [...] Und deswegen ist der Anspruch, dass sie mit 16 noch was davon spüren, möglicherweise zu hoch für diese Maßnahme [angesetzt]“ (Int. Sabeva, 185).

Je später der dritte Messpunkt also liegt, desto mehr Herausforderungen ergeben sich. Einerseits birgt dieses Vorgehen einen enormen organisatorischen und finanziellen Aufwand, der viele Ressourcen beansprucht. Andererseits wird es immer schwieriger, die intendierten oder nicht intendierten Wirkungen auf das Programm rückführen zu können: „Je länger der Zeitraum ist, desto mehr andere Faktoren kommen im Leben der Kinder dazu“ (Int. Sabeva, 185). Daher meint Sabeva, dass auch ein früherer dritter Zeitpunkt sinnvoll wäre, der mit oder kurz nach der Beendigung der Volksschule angesetzt werden könnte. Dennoch ist auch hier eine Reihe an Informationen zu erheben und der Kontakt aufrechtzuerhalten.

Eine Idee der Umsetzung wäre ein weiterer Fixtermin im Programm, nach offiziellem Programmende, in dessen Rahmen die AkteurInnen ein drittes Mal befragt werden. Ein solcher dritter Programmpunkt wäre grundsätzlich umsetzbar, „es ist alles eine Kosten- und Ressourcenfrage“ (Int. Leeb, 121). Der Projektleiter hat bereits über ein Treffen mit Ehemaligen nachgedacht: „Eine Möglichkeiten eines Wiedersehens wäre in Form von einem ‚Nightingale‘-Treffen, einem Fest oder ähnlichem, wo man sich wieder trifft. Es ist aber ein großer Aufwand und muss finanziert werden“ (Int. Leeb, 131). Auch die meisten AkteurInnen zeigen sich einem Wiedersehen, das ähnlich dem Abschlussfest gestaltet sein könnte, als offen. Die Mütter und MentorInnen würden gerne an einer weiteren ‚Nightingale‘-Veranstaltung teilnehmen, die Lehrerinnen zeigen sich ebenso interessiert. Allein die Mentees sind einem weiteren Fixtermin gegenüber kritisch eingestellt. Kind D lässt erkennen, dass er viele Termine hat, eine Veranstaltung wie das Abschlussfest sorgen dabei für Stress: er möchte ein weiteres Treffen nicht, er hatte am ‚End-Day‘ noch Hausaufgaben zu erledigen und hatte erste um halb eins Schule aus (vgl. Int. Kind D, 127). Auch Kind C ist nicht begeistert über

ein weiteres Treffen mit allen Beteiligten, es sind zu viele Personen (Int. Kind C, 119). Außerdem wäre er lieber schwimmen gegangen, konnte dies aber nicht, weil er am Abschlussfest teilnehmen musste (Int. Kind C, 82).

Auch wenn ein dritter gemeinsamer Termin im Programm mitaufgenommen werden könnte, bleibt der Zeitpunkt einer solchen Veranstaltung noch offen. Denn auch hier gilt, dass die Aufrechterhaltung der Kontakt mit allen AkteurInnen die wohl größte Herausforderung darstellt, die – neben den zusätzlichen Ressourcen für die Veranstaltung – einige finanzielle und organisatorische Ressourcen beanspruchen würde. Demnach würde sich hier ein Termin anbieten, der sich noch innerhalb der Volksschulzeit der ehemaligen Mentees befindet.

5.4.4 Akzeptanz

Akzeptanz wird bei Kurz und Kubek wie folgt beschrieben: „Die Wirkungsziele müssen von den Stakeholdern akzeptiert werden. Das bedeutet, dass ein gemeinsames Verständnis über die Wirkungsziele besteht und dass diese von allen Beteiligten mitgetragen werden.“ (Kurz und Kubek 2013, S. 30). Hiermit beziehen sich Kurz und Kubek (2013) auf die Festlegung von Wirkungszielen bei der Entwicklung einer sozialen Intervention. Retrospektiv sind die Ziele – die Bildungsmotivation und den Selbstwert zu stärken, aber auch alle anderen intendierten Zielsetzungen – als akzeptiert anzusehen, da die Ziele kommuniziert werden und das Programm von Stakeholdern und Fördergebern unterstützt und bereits seit Jahren umgesetzt wird. Offen bleibt die Frage, ob eine Wirkungsanalyse und der hypothetisch damit verbundene Mehraufwand für die OrganisatorInnen und AkteurInnen akzeptiert würden.

Günther Leeb erklärte den Aufwand, vor allem für Lehrpersonen, als nicht groß, aber in der Summe der Dinge doch als anstrengend (vgl. Int. Leeb, 119). Die Lehrpersonen in die Wirkungsanalyse miteinzubinden könnte damit eine Herausforderung darstellen. Bezüglich des jetzigen Aufwandes für ‚Nightingale‘ bestätigt Lehrerin H (vgl. 113 und 123), dass es nicht viele Ressourcen von ihr verlange. Im Jahrgang 2016/17 nahm ein Kind ihrer Klasse an ‚Nightingale‘ teil – wenn sie hierbei rechtzeitig ihre Aufgaben erledigt, ist es kein großer Zusatzaufwand für die Lehrerin. Sie beschreibt zudem, dass Lehrpersonen oft mit Evaluierungen und ähnlichem beschäftigt sind, „aber wenn mir etwas wichtig ist, wenn ich merke, [...] da ist jetzt was dahinter, wo man das [Ergebnis] haben will“ (Int. Lehrerin H, 111), dann wäre sie bereit, Mehraufwand zu betreiben.

Während sich Mentorin E für die gezielte Beobachtung bereit erklären würde, sieht Mentor F eine gemeinsame Reflexion mit dem Kind kritisch: „Das geht dann in ein Ausmaß, wo ich

sag, ich möchte ja Beziehungsarbeit leisten und in dem weiter kommen und sehe da den Mehrwert sozusagen im Sein, wenn ich wirke, im Moment.“ (Int. Mentor F, 63). Der Mehraufwand einer weiteren oder erweiterten Befragung der MentorInnen schätzt Mentor F abhängig vom Kalendermonat ein, gegen Ende des Semesters wäre ihm dies zu viel Aufwand. Es wäre auch abhängig davon, wie viele Fragen zu beantworten wären und ob diese standardisiert oder offen zu beantworten wären (vgl. Int. Mentor F, 61).

Während die Mutter A sich für die Ausarbeitung eines Fragebogens und der Beobachtung ihres Kindes gerne zur Verfügung stellt (vgl. Int. Mutter A, 58), zeigt Mutter B auf, dass diese zu viel für sie sei. Dies könnte auf die Sprachkenntnisse der Mutter zurückzuführen sein. Eine Erleichterung durch die Übersetzung der Tests in ihre Erstsprache lehnt die Mutter jedoch ab (vgl. Int. Mutter B, 64 ff.).

Frau Leskowa (vgl. Int. Leskowa 184) zeigt sich schon länger interessiert an einer Dokumentation der Entwicklung der teilnehmenden Kinder, interpretierbar als begleitende Wirkungsanalyse. Würde eine solche Wirkungsorientierung vorgeschrieben sein, wäre es der Programmleiterin an der Pädagogischen Hochschule wichtig, dass diese von den Personen, die im Programm arbeiten, mitgestaltet wird, da diese die relevanten Erfahrungen mitbringen.

„Ich der Überzeugung bin, dass das nachhaltiger und effektiver wirkt, weil wir einfach die Grunderfahrung [haben] und die Grundentwicklung sehr wichtig ist, [...] egal welche Messungen man aufstellen möchte, [...] das ist mal das wichtige von den PraktikerInnen aber das andere ist natürlich schon die Theorie auch, [...] ich glaube es braucht immer eine Gruppe die aus allen besteht. [...] Aber wenn von oben gestartet, braucht man [...] die Menschen, die beteiligt sind, um wirklich ein effektives Programm zu entwickeln“ (Int. Leskowa, 112)

Wie in dieser Aussage ersichtlich, ist die Frage um die Akzeptanz eng verbunden mit der Vor- und Nachteilen einer Wirkungsorientierung. Die Befragten zeigten sich in ihrer Meinung zu einer Wirkungsorientierung bzw. Wirkungsanalyse zwiespalten und dem im Forschungsstand aufgezeigten Diskurs zur Wirkungsorientierung in der Sozialen Arbeit ähnlich.

Der Projektleiter kam mit den Fragen im Interview ins Überlegen. Er meint, dass eine Messbarkeit schwierig ist, aber mit einem hohen Aufwand möglicherweise realisierbar, „nur wenn man dann versucht wie könnte das praktisch gehen kommt man halt irgendwie ins Grübeln“

(Int. Leeb, 83). Während des Gesprächs zeigte sich Herr Leeb mit einer eher kritischen Haltung gegenüber einer Wirkungsorientierung oder Wirkungsmessung:

„Die Qualität liegt in der momentanen Wertschätzung des Kindes“ (Int. Leeb, 139), „das sind alles, [...] eine ganze Fülle an kleinen Geschenken für die Kinder, die sich durch nichts anderes, als dass sie eine bildungsbenachteiligte Situation haben, ohnehin schon verdient haben zu bekommen und eigentlich in noch viel größeren Ausmaß bekommen müssten. Und das genügt eigentlich [...] als Rechtfertigung für das Programm“ (Int. Leeb, 207).

Eine ständige Abfrage oder Beobachtung lenke vom Wesentlichen ab. Man könne eine Wirkungsanalyse zwar versuchen, er würde sich auch gerne überzeugen lassen, dennoch ist er der Meinung, dass man am Ende keine herausragenden Ergebnisse erhalten würde. Er bezieht sich auf die von ihm erstellten Evaluierungen und meint, dass in 80-85 Prozent der Fälle das Mentoring funktioniere, in denen beide Seiten vom Programm profitieren (vgl. Int. Leeb, 143). Gerade bei „soft skills“ (Int. Leeb, 100) sei es nahezu unmöglich festzustellen, wie sich diese verändern und wie dies auf das Programm zurückzuführen ist. Günther Leeb spricht auch an, dass das Programm für eine Veränderung von verhaltenspsychologischen Faktoren zu kurz sei und die Ziele für eine solche Messung klarer definiert werden müssten (vgl. ebd. 207), und dass es „politisch auch nicht einfach zu beurteilen ist, wie wertvoll das ist“ (Int. Leeb, 143).

Auch Sabeva und Leskowa zeigen kritische Aspekte der Wirkungsorientierung auf, sind dies grundsätzlich aber positiver eingestellt und sehen mehr Vor- als Nachteile.

Sabeva stellt sich ebenfalls die politische Frage: „Wer misst was mit welcher Intention?“ (Int. Sabeva, 171). Vom organisatorischen Aspekt betont sie immer wieder, dass Langzeituntersuchungen einen hohen Mehrwert haben, aber zu selten durchgeführt werden. Der Ressourceneinsatz ist enorm, die Studie müsste lange vorbereitet und mehrfach getestet werden, bis eine reguläre Durchführung möglich sei (vgl. ebd. 211). Dennoch sieht Sabeva Wirkungsmessung als notwendig und plädiert dafür, diese als Standard bei sozialen Interventionen einzuführen. Sie bezieht dies deutlich auf die Sozialwirtschaft in Wien, in die viele Ressourcen und viel Aufwand gesteckt werden und die Outcomes oft fraglich sind. Dabei kommt es drauf an, was mit der Wirkungsmessung erreicht werden will – ob diese als *Drop-out*-Mechanismus dienen soll, um Projekte fallen zu lassen, die nicht gleich eine Wirkung nachweisen können, oder ob man wissen will, welches Projekt was bringt: „Aber selbst das finde ich gerechtfertigt, da die Ressourcen nun mal begrenzt sind, muss man sich anschauen, welche Maßnahmen bringen im Endeffekt etwas“ (Int. Sabeva, 169). Sabeva geht konkreter

auf die Auswahl förderungswürdiger Projekte ein und zeigt auf, dass hier eine Wirkungsorientierung einen Unterschied machen könnte:

„Ich denke mir [...] im Idealfall ist der Geldgeber so kompetent, dass er eben sehr wohl weiß, was will er denn fördern. Und das [...] fände ich grundsätzlich sinnvoller, als er tut so als wüsste er es nicht und fördert alles Mögliche relativ wahllos oder nach Sympathie. Die können besser Berichte schreiben, dann fördern wir halt die. Weil wie wird das jetzt entschieden? Also irgendwie wird es entschieden, es wird nie alles gefördert, und dann ist die Frage nach welchen Kriterien, dass die Wirkungsmessung möglicherweise schon besser, als das, was im Moment gemacht wird“ (Int. Sabeva, 173).

Auch Leskova spricht sich für eine Wirkungsanalyse aus, zeigt aber die Gefahren einer solchen auf. Sie erkennt die Wirkungsmessung als notwendig. Diese unterstützt nicht nur die Beteiligten, sondern auch die zukünftigen TeilnehmerInnen (vgl. Int. Leskova, 118). Es sei hilfreich zu wissen, worauf man schauen kann, schon lange wünsche sich Leskova eine begleitende Planung bzw. die Nachvollziehbarkeit der Entwicklung der Mentees (vgl. 184). Dennoch stellt sich die Organisatorin die Frage, inwieweit gewisse Indikatoren wirklich messbar sind. Häufig können nur subjektive Einschätzungen und Beobachtungen miteinbezogen werden – hier sieht sie die Schwierigkeit, diese Subjektivität transparent darstellen zu können. Dennoch ist es aufgrund des Settings von ‚Nightingale‘ sinnvoll, diese qualitative, subjektive Betrachtungsweise mitaufzunehmen und zu nutzen (vgl. Int. Leskova, 118). Würde die Art der Wirkungsmessung aber, beispielsweise vom Fördergeber, strikt vorgegeben sein, sieht Leskova die Gefahr, dass ebendieses Subjekt vernachlässigt wird. Würde das Programm in ein enges Korsett gedrängt, konzentriere man sich nur noch auf die Einhaltung dessen und verliere den Fokus für den Einzelfall (vgl. ebd. 122): „Mit jeder Art von Druck, von Vorgaben, wird das Intuitive eingeschränkt“ (Int. Leskova, 164).

Leskova und Sabeva sehen beiden einen potentiellen Mehrwert für die Weiterentwicklung des Programmes. Mit den Ergebnissen einer Wirkungsforschung zu ‚Nightingale‘ könnte „die ganze Arbeit verbessert, professionalisiert und mit gleichem Aufwand [...] einfach mehr [...] für die Kinder [rausgebracht werden]“ (Int. Sabeva, 169). In diesem Zitat spricht Sabeva weitere konstitutive Elemente in der Diskussion um die Wirkungsorientierung in der Sozialwirtschaft an: Effizienz und Professionalisierung. Zum Begriff der ‚Effizienz‘ meint Sabeva, dass dieser im sozialen Bereich schnell als „wirtschaftlich abgetan“ (vgl. ebd.) ist. Dabei seien die Ziele im Sozialen meist groß und die Ressourcen gering – der soziale Bereich

stehe somit immer vor der Herausforderung, die Effizienz zu erhöhen, so dass man mit den geringen Mitteln möglichst viel bewirkt (vgl. Int. Sabeva, 169).

6 Diskussion der Ergebnisse

Im Folgenden wird die Diskussion der dargestellten Ergebnisse eine Antwort auf die Forschungsfrage ermöglichen. Das Kapitel soll dazu dienen, den forschungsorientierten und praxisorientierten Teilaspekt der Forschungsfrage zusammenzuführen. Insbesondere wird dargestellt, wie sich das entwickelte Forschungsdesign in die aktuelle Praxis einfügen würde, welche Änderungen des Status Quo zu erwarten wären und welche Möglichkeiten und Grenzen dabei beachtet werden sollten. Nachdem das Forschungsdesign für eine Wirkungsanalyse in den Bereichen Bildungsmotivation und Selbstwert entwickelt und in Kapitel 5.1 dargestellt wurde, ergab sich eine Reihe an offenen Fragen, etwa zur Umsetzung, Terminierbarkeit und Akzeptanz. Diese Fragen wurden in die qualitative Befragung integriert und in deren Auswertung sowie in der Auswertung des Onlinefragebogens 2016/17 besonders berücksichtigt.

In diesem Kapitel wird zunächst das Forschungsdesign kurz umrissen. Ferner soll auf die aktuelle Evaluierung eingegangen werden, um die Unterschiede zum hier entwickelten Forschungsdesign aufzuzeigen. Wie die Ergebnisse gezeigt haben, bestehen unterschiedliche Voraussetzungen für die einzelnen Programm-Durchgänge. Entsprechende Schlussfolgerungen hinsichtlich der Bildung der Vergleichsgruppe werden präsentiert. Daran anschließend wird die Frage diskutiert, ob die Evaluierungen im Fremd- oder Selbstbericht durchgeführt werden sollen. Wie dem Anspruch der Langfristigkeit gerecht werden kann und was bezüglich der Anwendung der Instrumente beachtet werden muss. Schließlich wird die Akzeptanz der Wirkungsanalyse für ‚Nightingale‘ Wien anhand der ExpertInnen-Rolle OrganisatorInnen besprochen. Zusammengefasst sollen anhand dieser Ergebnisse die Möglichkeiten und Grenzen bei der Planung und Umsetzung einer Wirkungsanalyse für ‚Nightingale‘ dargestellt werden.

6.1 *Rekapitulation des Forschungsdesigns*

Das hypothetische Forschungsdesign versucht, durch die Wahl von Quasi-Experiment und *Double-Difference*-Methode sowie geeigneter psychologischer Instrumente Messbarkeit und Vergleichbarkeit für Bildungsmotivation und Selbstwert herzustellen. Dank mehrfacher Messpunkte können Veränderungen ausgemacht werden, mittels einer Vergleichsgruppe können diese auf das Programm rückgeführt werden. Das FFM betrachtet die Persönlichkeitsveränderung auf fünf Ebenen, von denen zwei mit der Bildungsmotivation korrelieren.

Daher kann angenommen werden, dass die restlichen drei Komponenten eine persönlichkeitsbezogene Veränderung in nicht-intendierten Zielen feststellen kann. In Kombination mit den Instrumenten des MGT und den Selbstwertskalen kann ein ganzheitliches Bild geschaffen werden. Anhand eines dritten Messpunkt M₂ könnte die langfristige Wirkung des Programmes festgestellt werden.

6.2 Aktuelle Evaluierung ‚Nightingale‘ Wien

Derzeit wird von den Fördergebern keine Wirkungsanalyse verlangt, obwohl neben dem Finanz- und Sachbericht auch das Erreichen der Ziele aufgezeigt werden muss. Jedes Mitglied des ‚Nightingale‘-Netzwerks gibt einen Jahresbericht ab. Dieser beinhaltet neben Angaben zur TeilnehmerInnen-Anzahl Berichte zu unternommenen Aktivitäten. Hierbei muss betont werden, dass diese Aktivitäten nicht als Wirkungen, sondern eher als Leistungen zu betrachten sind (vgl. Kurz und Kubek 2013, S. 27). Diese Jahresberichte sind nicht mit einer Wirkungsanalyse zu vergleichen, in der die Leistung des Programmes (Output) lediglich die Grundlage der Wirkung (Outcome) darstellen würde. Bei einer Durchführung der Wirkungsanalyse anhand des hier vorgestellten Forschungsdesigns könnten die Jahresberichte jedoch möglicherweise als Informationsquelle zur Anreicherung der Output-Darstellung dienen. Bei ‚Nightingale‘ Wien findet zusätzlich zu den Jahresberichten eine Evaluierung statt, welche die MentorInnen in den Mittelpunkt stellt und auf deren monatlichen Protokollen sowie Teilnahme an einer Onlinebefragung basiert. Die MentorInnen sind dabei aufgefordert, einzuschätzen, wie ihr Mentee vom Programm profitiert hat. Diese Evaluierung geht über eine reine Output-Darstellung hinaus, ist jedoch stark auf die subjektive Einschätzung der MentorInnen angewiesen. Darüber hinaus ist in diesem Setting keine längerfristige Beobachtung der Wirkung möglich. Für die Dauer des Programmes allerdings könnten diese Informationen wertvolle Hinweise geben, wenn es darum geht, unerwartete oder außergewöhnliche Einzelfall-Ergebnisse einer künftigen Wirkungsanalyse zu deuten.

6.3 Vergleichbarkeiten im Double-Difference-Ansatz

Es stellt sich die Frage, inwieweit sich Evaluierungsergebnisse innerhalb einer Wirkungsanalyse, aber auch über verschiedene Evaluierungen hinweg vergleichen lassen. Grundsätzlich muss zwischen zwei Vergleichbarkeiten unterschieden werden. Zunächst gilt es, die zeitliche einzelfallbezogene Vergleichbarkeit unbedingt dadurch zu gewährleisten, dass die Evaluierung zu allen Messpunkten stets von derselben Person durchgeführt wird. Zweitens

muss zwischen Ziel- und Vergleichsgruppe auf eine gewisse Ähnlichkeit geachtet werden, wie anhand folgender Umstände gezeigt werden soll.

Die Ergebnisse der empirischen Befragung zeigen, dass das Programm komplexer ist als angenommen. Durch die erhöhten Förderbeiträge seit 2014 konnte das Programm erweitert werden und hat sich seitens der Anzahl der Tandems verdoppelt. Aufgrund begrenzter organisatorischer Kapazitäten wurden seither zwei Durchgänge durchgeführt, die zwar beide jeweils sieben Monate lang dauern, der Durchgang von Mai bis November jedoch die großen Sommerferien umfasst. Aus diesem Unterschied in der Ausgangslage geht hervor, dass bei der Bildung der Vergleichsgruppe auf eine hohe Ähnlichkeit zur Zielgruppe geachtet werden sollte. In Fällen unterschiedlicher Ausgangslagen bei mehreren Programmdurchgängen ist es daher wünschenswert, stets nur jene Ziel- und Vergleichsgruppen zu heranzuziehen, die dieselbe Ausgangslage teilen.

Eine weitere Komplexitätsstufe ergibt sich durch die Teilnahme einer Mittelschule. Die Kinder in der Mittelschule befinden sich in der nächsthöheren Schulstufe, d.h., sie haben die Volksschule bereits verlassen. Dadurch entstehen unterschiedliche Voraussetzungen: die Kinder sind älter, sind in ihrer Entwicklung fortgeschrittener, kurz vor oder bereits in der Pubertät. Bei den Mentees der Mittelschule ist die Abbruchrate im Programm höher als in der Volksschule. Auch für diesen Fall soll Ähnlichkeit zwischen Ziel- und Vergleichsgruppe angestrebt werden. Praktisch empfehlenswert ist es daher, dass die Vergleichsgruppe aus jener Klasse (und damit jener Schule) gebildet wird, in der sich auch die Personen der Zielgruppe befinden.

Unterschiede ergeben sich jedoch nicht nur in Bezug auf die Mentees, auch die MentorInnen weisen Unterschiede auf. Im Masterpraktikum beschäftigen sie sich intensiver mit ihrem Mentee und erstellen eigene Wirkungsziele. Diese zielgerichtete Betreuung verändert womöglich die Wirkung beim Mentee, ist aber in weiter Folge nicht zu berücksichtigen, da die *Double-Difference*-Methode zwischen der Ziel- und Vergleichsgruppe und den unterschiedlichen Zeiträumen vergleicht, nicht aber zwischen den einzelnen Kindern.

Zusammenfassend zeigt sich, dass bei Beachtung der personellen Kontinuität der Evaluierung sowie bei sorgsamer Bildung der Vergleichsgruppe die nötige Vergleichbarkeit für den *Double-Difference*-Ansatz auch bei unterschiedlichen Voraussetzungen geschaffen werden kann. Zu betonen ist dabei jedoch, dass diese keine einzelfallbezogene Vergleichbarkeit über

die Kinder hinweg (insbesondere nicht über die TeilnehmerInnen verschiedener Durchgänge oder Schultypen) impliziert.

6.4 Bildung von Ziel- und Vergleichsgruppe

Auch die Auswahl der TeilnehmerInnen selbst ist für Wirkungsfeststellung relevant. Grundsätzlich ist die Chance auf Wirkung umso höher, je größer das Wirkungspotential beim Kind ist. Bei der Auswahl der Kinder für die Teilnahme kann festgehalten werden, dass die Lehrpersonen die Gründe für die Entscheidung aufzeichnen. Dabei gibt es jedoch keine stringenten Kriterien, die erfüllt sein müssen. Die *Incomes* (vgl. Beywl 2009) beschreiben, mit welchen Erfahrungen, Vorwissen oder Fähigkeiten die beteiligten Personen in das Programm kommen. Auch die im Rhodes-Modell (2005) aufgezeigten Einflüsse auf die Mentoring-Beziehung können als *Incomes* betrachtet werden. Die Faktoren a) persönliche Geschichte und Erfahrung, b) sozialen Kompetenzen und c) Entwicklungsstufe des Kindes spielen eine entscheidende Rolle für die Persönlichkeitsentwicklung des Kindes. Diese *Incomes* sind jedoch bereits vor dem Programmstart relevant, da sie die Baseline für die Nachvollziehbarkeit der Wirkungen darstellen. Aus diesem Grund sieht das Forschungsdesign einen ersten Messpunkt M_0 vor. Um die Lehrpersonen bei ihrer Entscheidung zu unterstützen, welche Kinder für die Programmteilnahme vorgeschlagen werden sollen, könnte die M_0 -Evaluierung von allen Kindern der Klasse durchgeführt werden. Die Ergebnisse würden den Lehrpersonen dazu dienen, ihre Entscheidung zusätzlich abzusichern. Weiter wären diese Testergebnisse hilfreich bei der Bildung einer Vergleichsgruppe.

Mit der Bildung einer Vergleichsgruppe gibt es bereits Erfahrungen bei teilnehmenden Schulen. Für einen Sprachtest im Zusammenhang mit dem ‚Nightingale‘-Programm Wien wählten Klassenvorstände Kinder aus, die ähnliche Sprachfähigkeiten wie die Kinder der Zielgruppe hatten (vgl. Himmelbauer 2016, S. 24). Im vorliegenden Forschungsdesign bietet der Messpunkt M_0 mit den unterschiedlichen Tests eine objektive Möglichkeit, eine Vergleichsgruppe zu bilden. Aufgrund der Datenerhebung aller Kinder einer Klasse (respektive aller vorgeschlagenen Kinder), kann festgestellt werden, welche nichtteilnehmenden Kinder ein vergleichbares Selbstwertgefühl und eine ähnliche Bildungsmotivation wie die teilnehmenden Kinder aufweisen.

Die Frage bleibt offen, ob die Eltern beider Gruppen ihre Kinder für eine langfristige Befragung zur Verfügung stellen. Außerdem ist es nach Wöhrer (2017) ratsam, auch die Zustimmung der Kinder einzuholen: Forschung mit Kindern bedarf nach der UN-Konvention der

Kinderrechte von 1989 der Zustimmung von Eltern und Kindern (vgl. Wöhrer et al. 2017, S. 45).

6.5 *Selbst- und Fremdevaluierung*

Die Befragung zur organisatorischen Praxis hat ergeben, dass keine der Bezugspersonengruppen MentorInnen und LehrerInnen über den gesamten vorgesehenen Untersuchungszeitraum mit dem Kind in Kontakt steht. So ist eine mögliche Fremdevaluierung durch MentorInnen ausgeschlossen, weil sie ihr Mentee zum Zeitpunkt M_0 noch nicht, oder aber zu wenig gut kennen. LehrerInnen scheiden aus, weil die meisten Kinder kurz nach der Programmteilnahme die Schulstufe wechseln, womit keine Evaluierung durch dieselbe Lehrperson zu den Zeitpunkten M_0 und M_2 bzw. M_1 und M_2 möglich ist. Die verbleibende Bezugsgruppe Eltern, die für eine Fremdevaluierung in Frage käme, weist diesbezüglich zwei Probleme auf: a) Es ist in vielen Fällen mit Sprachbarrieren zu rechnen und b) eine verzerrte Wahrnehmung ihrer Kinder ist nicht auszuschließen (vgl. Int. Sabeva, 167; vgl. Sabeva 2012, S. 23). Wenngleich eine ergänzende Fremdevaluierung durch die Eltern unter gewissen Voraussetzungen möglich erscheint, wird im Folgenden nur mehr die Selbstevaluierung durch die Mentees in Betracht gezogen. Mentees, die aufgrund von Sprachbarrieren nicht an der Evaluierung teilnehmen können, sollen deswegen jedoch keinesfalls von der Teilnahme am Programm ausgeschlossen werden.

Eine Fokusgruppe mit den Kindern, wie in Leeb (2014) beschrieben, finden bei ‚Nightingale‘ Wien nicht mehr regulär statt. In solchen oder ähnlichen Gruppensettings sind die Kinder schüchtern oder beantworten die Fragen höflich und nach Erwartung der Erwachsenen. Hier kann eine Wirkungsanalyse wie in der vorliegenden Arbeit geplant ansetzen, die nicht danach fragt, ob das Programm gefallen hat oder was am meisten Spaß machte. Die Reflexionsfähigkeit von Kindern wird unterschiedlich eingeschätzt. Während Wöhrer (vgl. 2017, S. 99 ff.) diese hoch bewertet, meint Günther Leeb (vgl. Int. Leeb, 119), die Kinder seien damit überfordert, die Konsequenzen des Erlebten für sich selbst einzuschätzen. Wenn Kinder erzählen würden, dann eher faktenorientiert. Das vorgeschlagene Forschungsdesign versucht, einen dritten Weg einzuschlagen, der sowohl Überforderung der Reflexionsfähigkeit der Kinder als auch das Stellen trivial-suggestiver Fragen vermeidet. Stattdessen beobachten die Kinder konkrete, faktische und zeitlich aktuelle ‚Tatbestände‘ an sich selbst (z.B. „Ich habe originelle Einfälle“ oder „Ich lasse alles herumliegen“ [Bleidorn und Ostendorf 2009,

S. 162]), über die sie Auskunft geben können. Das Ziehen der Konsequenzen daraus bezüglich der Wirkungsziele wird hingegen nicht dem Kind überantwortet, sondern geht aus den theoretischen Überlegungen zum Forschungsansatz hervor.

6.6 Langfristigkeit

Das vorgeschlagene Verfahren ist aus einem weiteren Grund von Vorteil: Während die Follow-Up-Studie von Lönroth (2007) auf das Erinnerungsvermögen der Kinder angewiesen war, sieht das entwickelte Forschungsdesign zu jedem Zeitpunkt nur die Erhebung des aktuellen Zustandes vor. Dies ist insofern bedeutend, als die beiden Lehrpersonen G und H sowie Leskowa die Erinnerungsfähigkeit bereits nach wenigen Monaten beeinträchtigt sehen (vgl. Int. Lehrerin G 110 ff.; Int. Lehrerin H, 139 ff.; Int. Leskowa 226 ff.). Sabeva hingegen wies auf die Möglichkeit eines nicht-erinnerungsbasierten Verfahrens hin und sieht dementsprechend die Möglichkeit für eine langfristige Wirkungsbeobachtung gegeben (Int. Sabeva, 190). Der nicht-retrospektive Aufbau des Analysekonzepts ist daher ein wichtiger Aspekt, wenn es um den Nachweis langfristiger Wirkungen geht.

Der konkrete Zeitpunkt des dritten Messpunkts M_2 wurde von den AkteurInnen unterschiedlich festgesetzt. Jene Personen, die Erinnerungsfähigkeit als Voraussetzung der Evaluierung ansahen, setzten den dritten Messpunkt innerhalb von sechs Monaten nach Beendigung des Programmes an. Sabeva, die mit dem nicht-erinnerungsbasierenden Verfahren vertraut war, schlug für den dritten Messpunkt das Alter von 14 Jahren vor (4–6 Jahre nach dem Beginn des Programmes). Da sechs Monate zu wenig sind, um eine langfristige Wirkung beobachten zu können, ist der spätere Zeitpunkt als dritter Messpunkt anzusetzen. Mit dem Messpunkt (M_2) nach 4–6 Jahren reiht sich das Forschungsdesign in bestehende Definitionen von Langfristigkeit in Wirkungslogiken ein (vgl. Kellogg Foundation 2004, S. 2).

Mit diesem Messpunkt ergibt sich die organisatorische Herausforderung der Aufrechterhaltung des Kontaktes sowohl mit den TeilnehmerInnen als auch mit der Vergleichsgruppe. Für eine dritte Befragung müssten die Kontaktdaten aktualisiert vorhanden sein. Offen bleibt die Bereitschaft der AkteurInnen – hierzu zählt in diesem Fall auch die Vergleichsgruppe – an einer dritten Befragung nach 4–6 Jahren teilzunehmen.

6.7 Anwendbarkeit der Instrumente

Eine Zielsetzung der vorliegenden Arbeit war es, ausschließlich Erhebungsinstrumente einzusetzen, die die Hauptgütekriterien Objektivität, Reliabilität und Validität erfüllen. Von den

ausgewählten Verfahren sind das FFM und die Selbstwertskalen in deutscher Übersetzung verfügbar. Die Testungen zu *growth mindset* und *school connectedness* sind in englischer Sprache verfasst, wobei die ExpertInnen-Befragung ergeben hat, dass von einer eigenen Übersetzung abgeraten wird (Int. Sabeva, 77). In der vorliegenden Fremdsprache können die Tests nicht angewandt werden, daher ist auf den Einsatz dieser Instrumente zu verzichten oder aber nach einem Ersatz in deutscher Sprache zu suchen. Die Noten können unabhängig von der Sprache aufgenommen werden, wenn Vergleichbarkeit gegeben ist, was innerhalb von Wien respektive Österreich der Fall ist.

6.7.1 Normierung und Altersbegrenzung

Die Normierung ist ein „Nebengütekriterium für diagnostische Verfahren“ (Rauthmann 2016b), das eine Einordnung der erhobenen Werte durch den Vergleich mit den Werten einer Normgruppe ermöglicht. Dies ist notwendig, wenn ein Erhebungsergebnis auf seine Bedeutung – gemessen an der Norm – eingeschätzt werden soll. Aus der Verfügbarkeit der Normdaten geht auch die Anwendbarkeit des Tests in Bezug auf das Alter der ProbandInnen hervor: Wird mit einer Normgruppe verglichen, so ist darauf zu achten, dass das Alter des/der ProbandIn innerhalb der Altersbandbreite der Normgruppe liegt. Diese Anwendbarkeitsgrenze macht jedoch keine direkte Aussage darüber, ob der Test altersgerecht einsetzbar ist. Folgende Aussage von Bleidorn und Ostendorf (2009) mit Blick auf das FFM-Instrument verdeutlicht dies:

„Im Hinblick auf die vergleichsweise geringe Anzahl der 11- und 15jährigen Kinder in unserer Stichprobe empfehlen wir außerdem, die Schlussfolgerungen auf den Altersbereich von 12 bis 14 Jahren zu begrenzen. Es wäre somit höchst wünschenswert, das Verfahren künftig in möglichst repräsentativen Stichproben verschiedener Altersgruppen einzusetzen, um die Generalisierbarkeit der hier berichteten Ergebnisse zu überprüfen und Daten für die Normierung des Inventars zu gewinnen“ (Bleidorn und Ostendorf 2009, S. 172).

Vielmehr wird hier die Möglichkeit von den AutorInnen in Betracht gezogen, das Instrument auf jüngere (sowie ältere) Kinder anzuwenden, mit dem Ziel, den Normierungsbereich des Tests zu erweitern.

Die Normgruppe als Nebengütekriterium (und die daraus hervorgehende Altersbegrenzung) spielt für diese Arbeit eine untergeordnete Rolle, da stattdessen eine Vergleichsgruppe herangezogen wird. Insofern ist die Bedeutung dieses Nebengütekriterium als gering einzustu-

fen, wovon die Hauptgütekriterien (Objektivität, Reliabilität und Validität) nicht beeinträchtigt werden. Die Literatur zeigt auf, dass bezüglich der Einhaltung von Nebengütekriterien durchaus Spielraum besteht. Während die Hauptgütekriterien „von zentraler Bedeutung für die Qualität von Tests sind [...], auf die fast ohne Einschränkung ein strenger Maßstab anzulegen ist“ (Amelang und Schmidt-Atzert 2006, 135, 138), „[mögen über] die relative Wichtigkeit der Nebengütekriterien zueinander (und auch zu den Hauptgütekriterien) [...] die Meinungen auseinandergehen“ (ebd. 2006, S. 161).

Für die vorliegende Arbeit bedeutet dies, dass eine Abweichung von der Altersbegrenzung möglich ist, umso mehr, als ohnehin kein Vergleich zur Normgruppe gezogen wird.⁶ Sabeva äußerte sich in der ExpertInnen-Befragung in einem ähnlichen Sinne (Int. Sabeva, 65). Dennoch wird empfohlen, die betreffenden Instrumente mit Vorsicht zu verwenden und insbesondere darauf zu achten, ob die Fragen von den ProbandInnen verstanden werden. Tabelle 7 zeigt im Resultat die Anwendungsmöglichkeiten und Grenzen der Instrumente auf, wobei grün (rot) eingefärbte Bereiche Anwendbarkeit anzeigen (ablehnen). Hellrote Bereiche zeigen eine Anwendung an, die darauf bedacht ist, zu prüfen, ob die Mentees die Fragen gut verstehen. Für das Forschungsdesign geht daraus hervor, dass FFM für Kinder und Jugendliche, MGT Noten sowie die Selbstwertskalen im Selbstbericht anwendbar sind. Auf die Anwendung von MGT *growth mindset* und MGT *school connectedness* muss verzichtet werden.

⁶ Auch in Bezug auf das Alter der Normdaten selbst gibt es Einschränkungen. Dieses sollte beim Vergleich mit der Norm 8 Jahre nicht überschreiten. Da im entwickelten Forschungsdesign nicht mit der Normgruppe verglichen wird, ist diese Einschränkung ebenfalls hinfällig.

Tabelle 7: Übersicht der ausgewählten Instrumente für das Forschungsdesign mit den limitierenden Faktoren.

Instrument	AutorIn, Jahr	Sprache	Normdaten
FFM für Kinder und Jugendliche	Bleidorn und Ostendorf 2009	Deutsch	12 bis 15 Jahre
MGT Noten	o.J.	Unabhängig	8 bis 18 Jahre
MGT Growth mindset	Castella und Byrne 2015	Englisch	12 bis 19 Jahre
MGT School connectedness	Karcher und Sass 2010	Englisch	11 bis 18 Jahre
Selbstwertskalen	Schütz und Selling 2006 (nach Daig et al. 2008)	Deutsch	Ab 14 Jahre

6.8 Akzeptanz

Ob eine langfristige Evaluierung umsetzbar ist, hängt auch mit der Akzeptanz der AkteurInnen zusammen. Die Meinungen und Aussagen der OrganisatorInnen für und gegen eine Wirkungsanalyse ähneln dabei den verschiedenen Positionen innerhalb der wissenschaftlichen Diskussion (vgl. Kapitel 2). Der Projektleiter Günther Leeb äußerte sich gegenüber dem Vorhaben ambivalent. Einerseits bezweifelte er zwar grundsätzlich die Messbarkeit von *Soft-Skills*, andererseits sei er offen genug, diese Meinung auch zu revidieren, wenn entsprechende Ergebnisse dafür sprechen würden. Er hebt jedoch hervor, dass die knappen Ressourcen primär für die eigentliche Arbeit zugunsten der Kinder, anstatt für eine Wirkungsanalyse eingesetzt werden sollen. Für die beiden anderen Organisatorinnen bietet eine Wirkungsanalyse einen großen Mehrwert für ihre jetzige Arbeit und für die Verbesserung des Programmes. Leskowa ist aus einer pädagogischen Perspektive sehr interessiert und argumentiert, dass eine so angelegte Wirkungsanalyse helfen könnte, die Arbeit mit und zugunsten der Kinder zu verbessern. In einem ähnlichen Sinne äußerte sich Sabeva, die aufgrund der Professionalisierungspotentiale für das Programm einem solchen Vorhaben sehr positiv gegenübersteht. Beim Anstreben von Professionalisierung und Steigerung der Effizienz sollten allerdings darauf geachtet werden, dass die Möglichkeit zu intuitiver Arbeit nicht eingeschränkt wird. Insgesamt sind sich die OrganisatorInnen unklar darüber, ob die Dauer des

Programmes ausreicht, um langfristige Wirkungen aufweisen zu können. Wie Rhodes bereits gezeigt hat, hat die Dauer des Mentorings einen Einfluss auf die Wirkung: Programme, die länger als ein Jahr dauerten, hatten positiveren Einfluss auf die schulische Leistung und andere psycho-soziale Bereiche, als Programme, die kürzer als ein Jahr (z.B. zwischen drei und sechs Monaten) andauern (vgl. Rhodes et al. 2005, S. 148). Eine kurze Dauer kann zu niedrigen Outcomes in der Wirkungsanalyse führen, wobei ein solches Resultat als Argument für die Verlängerung des Programmes und entsprechender Steigerung der finanziellen Mittel der Fördergeber verwendet werden könnte.

Grundsätzlich wäre das Forschungsdesign (bei entsprechender Adaptierung der Sprache der Instrumente) auch an anderen Standorten einsetzbar, müsste dann aber wieder auf Akzeptanz und Umsetzbarkeit geprüft werden. Offen bleibt die Frage, ob ein Vergleich der Wirksamkeit zwischen mehreren nationalen Standorten möglich ist. Bei einem Vergleich verschiedener nationaler Zielgruppen wären unterschiedliche Vergleichsgruppen involviert, was die Vergleichbarkeit der Ergebnisse potentiell reduziert könnte. Es müsste daher zunächst plausibel gemacht werden, dass von einer ausreichenden Ähnlichkeit der Vergleichsgruppen ausgegangen werden kann. Weiters bestünde die Gefahr, dass bei solchen Standortvergleichen ein Wettbewerbsdenken Einzug hält, das schädliche Auswirkungen auf das Programm haben könnte. Ob ein solcher Vergleich innerhalb des Netzwerkes sinnvoll oder nützlich ist, kann an dieser Stelle nicht erörtert werden.

Durch die Mitgliedschaft des ‚Nightingale‘-Netzwerkes beim ‚European Center for Evidence-Based Mentoring‘ ist das Interesse an weiterführender Wirkungsforschung durch die Managerin Lönroth begründet. Ihr Interesse an Forschung im Mentoring- und Wirkungsbereich dient der Weiterentwicklung des Programmes und der Erhaltung oder Steigerung der Qualität. Das Interesse an der Wirkung von ‚Nightingale‘ ist jedenfalls groß: Auf der aktuellen Konferenz des Mentoring-Programmes wurden zwei von drei Workshops zu diesem Thema abgehalten. Eine rezente, noch nicht veröffentlichte Studie zu den Outcomes der Kinder und Studierenden wurde dort vorgestellt. Die Kooperation mit dem ECEB-Mentoring unterstreicht damit das Interesse an einer stärkeren Wirkungsorientierung.

7 Conclusio und Ausblick

Soziale Programme und Projekte müssen sich mehr und mehr mit ihrer Wirkung beschäftigen und diese nachweisen. Dabei wird die Langfristigkeit der Wirkungen häufig vernachlässigt, einerseits weil die finanziellen Ressourcen für eine langfristige Beobachtung nicht ausreichen, andererseits weil grundsätzlich angezweifelt wird, ob ‚weiche‘ Wirkungen messbar sind.

In der vorliegenden Arbeit wurde den Fragen nachgegangen, wie die langfristigen Wirkungen des Mentoring-Programmes ‚Nightingale‘ auf die Zielgruppe der teilnehmenden Kinder (Mentees) gemessen werden können. Zudem sollten die Möglichkeiten und Grenzen bei einer Planung und hypothetischen Umsetzung einer solchen Wirkungsanalyse aufgezeigt werden. Um diese Fragen zu beantworten, wurde zunächst anhand von Literatur zu Wirkungsforschung und zu psychologischen Instrumenten ein Forschungsdesign für eine Wirkungsanalyse für die Ziele ‚Stärkung der Bildungsmotivation‘ und der ‚Steigerung des Selbstwertes‘ erarbeitet. Um die Wirkung feststellen zu können, bedarf es einer Vergleichsgruppe und mehrerer Messzeitpunkten, an denen Tests zur Erhebung der Persönlichkeitsveränderung durchgeführt werden. Das Forschungsdesign wurde in einem zweiten Schritt mittels teilstrukturierter ExpertInneninterviews mit diversen AkteurInnen diskutiert. Die OrganisatorInnen, MentorInnen, Mentees, Eltern und Lehrpersonen wurden anhand von Leitfragebögen zu unterschiedlichen Teilaspekten der Wirkungsanalyse befragt, zusätzlich wurden ausgewählte *Items* eines Onlinefragebogens aus der aktuellen Evaluierung ausgewertet. Die Ergebnisse zeigen, dass die Planung und theoretische Umsetzung einer Wirkungsanalyse grundsätzlich durchführbar ist, aber das vorgeschlagene Forschungsdesign einige Limitationen beinhaltet. Die Wirkung der eigenen Arbeit festzustellen ist den AkteurInnen wichtig, auch wenn häufig kritische Aspekte und Unsicherheiten in Bezug auf die Messbarkeit von Wirkungen benannt wurden. Eine Wirkungsanalyse ist mit einem hohen Ressourcenaufwand verbunden, weswegen sich Kooperationen mit Hochschulen – auch langfristig – anbieten würden.

Der hier entwickelte Ansatz ist ein wesentlicher Fortschritt gegenüber gegenwärtiger Evaluierungen, die nur auf einen Messzeitpunkt abstützen. Das Forschungsdesign wurde anhand des aktuellen Forschungsstands aus dem Programm entwickelt und mit Instrumenten geplant, die den wissenschaftlichen Gütekriterien entsprechen. Aufgrund des wissenschaftli-

chen Ansatzes, der sich in die *Evidence-Based-Practice* einordnen lässt, stellt das Forschungsdesign einen Kontrapunkt zur ökonomischen Determinierung dar. Die Zielgruppe wird dabei in den Mittelpunkt gestellt, die Kindern stärker in den Evaluierungsprozess miteinbezogen. Im Selbstbericht sollen das Fünf-Faktoren-Modell (FFM) für Kinder und Jugendliche, Teile des *Measurement Guidance Toolkit* (MGT) und die Selbstwertkala (SWS) bearbeitet werden. Durch standardisierten Fragen, die nicht auf die subjektive Einschätzung oder Erinnerungsfähigkeit der TeilnehmerInnen zielt, kann eine langfristige Betrachtungsweise ermöglicht werden. Dennoch wurden bei den ausgewählten Instrumenten Einschränkungen festgestellt. Zwei Verfahren können aufgrund der englischen Sprache nicht angewandt werden, bei den deutschsprachigen Instrumenten muss die Verständlichkeit der Sprache berücksichtigt werden. Einsetzbar sind das Fünf-Faktoren-Modell, die Selbstwertskalen und der Notenvergleich. Anzudenken ist eine alleinige Erhebung der Wirkungen durch das FFM, da es in der Literatur Hinweise auf Korrelationen zwischen Selbstwert und bestimmten Faktoren des FFM gibt. Diese Möglichkeit wäre allerdings noch genauer zu überprüfen.

Das FFM misst nicht die eigentliche Wirkung, sondern fungiert als ‚Übersetzungstool‘ zwischen der Persönlichkeit der TeilnehmerInnen und möglichen erzielten Wirkungen. Dies bringt den Vorteil mit sich, dass mit einmal erhobenen Daten korrelationstheoretisch eine ganze Vielzahl von (erreichten und nicht erreichten) Wirkungen untersucht werden können. In diesem Sinne hilft die Anwendung des FFM im Kontext einer Wirkungsanalyse nicht nur das Erreichen von intendierten persönlichkeitsorientierten Wirkungszielen zu überprüfen, sondern erlaubt vielleicht sogar die Entdeckung weiterer, bisher unbekannter Wirkungsziele.

Die Erkenntnisse aus dieser Arbeit zeigen in vielen Punkten die notwendige Kooperation zwischen Forschung und Praxis auf. ‚Nightingale‘ ist in diesem Bereich erfahren, da es bereits eine Zusammenarbeit mit der Universität Wien gibt. In einem nächsten Schritt könnten das konkrete Erhebungsinstrument, die Durchführungsplanung, die Auswertungsplanung und evtl. das Hinzufügen weiterer Wirkungsziele von wissenschaftlicher Seite – beispielsweise einer weiteren Masterarbeit zum Thema – geleistet werden. Die tatsächliche Durchführung wäre im Anschluss ohne allzu großen Mehraufwand an Ressourcen möglich. Wesentlich ist dabei, dass Wirkungsorientierung nicht allein für den Fördergeber und dessen Schublade anzusehen ist. Ein Projekt nach Wirkungszielen auszurichten bietet nicht nur in der Planung, sondern auch für die Beteiligten und Kunden oder KlientInnen einen großen

Mehrwert. Die Wirkungsorientierung kann als Projektplanungstool genutzt werden und erlaubt es, sich durch die regelmäßige Überprüfung der Wirkungen immer wieder der eigenen Arbeit bewusst zu werden und diese zu verbessern.

Literaturverzeichnis⁷

- Amelang, Manfred; Schmidt-Atzert, Lothar (2006): Psychologische Diagnostik und Intervention. 4. Aufl. Berlin, Heidelberg: Springer Medizin.
- Amirazodi, Fatemeh; Amirazodi, Maryam (2011): Personality traits and Self-esteem. In: *Procedia - Social and Behavioral Sciences* 29, S. 713–716.
- Asendorpf, Jens B.; Neyer, Franz J. (2012): Psychologie der Persönlichkeit. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Astleithner, Florentina (2012): Diskussion des Nutzens und der Herausforderung von Wirkungsorientierung am Beispiel der Wiener Wohnungslosenhilfe. In: *soziales_kapital* (8), S. 1–20.
- Beywl, Wolfgang (2009): Der Programmbaum von Univation. Univation - Institut für Evaluation. Köln. Online verfügbar unter http://www.univation.org/download/Programmbaum_Demokratieerziehung.pdf, zuletzt geprüft am 25.05.2017.
- Blaßnig, Raphaela (2012): Mentoring Migrant Kids. 'Nightingale - Mentoring für Kinder mit Migrationshintergrund' - ein Beitrag zur Chancengleichheit? Diplomarbeit. Universität Wien, Wien. Pädagogik.
- Bleidorn, Wiebke; Ostendorf, Fritz (2009): Ein Big Five-Inventar für Kinder und Jugendliche. In: *Diagnostica* 55 (3), S. 160–173, zuletzt geprüft am 25.01.2017.
- Bogner, Alexander; Littig, Beate; Menz, Wolfgang (2014): Interviews mit Experten. Wiesbaden: Springer Fachmedien Wiesbaden.
- Bowers, Ed; DuBois, David L.; Elledge, Chris; Hawkins, Stephanie; Herrera, Carla; Neblett, Enrique; Alem, Fasika (2016a): Mentoring Program Reviews. Hg. v. Michael Garringer und Mandy Howard. National Mentoring Resource Center Research Board. Washington D.C. Online verfügbar unter <http://www.nationalmentoring-resourcecenter.org/index.php/what-works-in-mentoring/reviews-of-mentoring-programs.html>, zuletzt geprüft am 13.08.2017.
- Bowers, Ed; DuBois David L.; Elledge, Chris; Hawkins, Stephanie; Herrera, Carla; Neblett, Enrique; Alem, Fasika (2016b): Academic performance. Unter Mitarbeit von

⁷ Zitiert nach Citavi-Basisstil.

- Michael Garringer und Mandy Howard. Hg. v. National Mentoring Resource Center Research Board. Washington D.C. Online verfügbar unter <http://www.nationalmentoringresourcecenter.org/index.php/toolkit/item/267-academic-performance.html>, zuletzt geprüft am 14.01.2017.
- Bowers, Ed; DuBois David L.; Elledge, Chris; Hawkins, Stephanie; Herrera, Carla; Neblett, Enrique; Alem, Fasika (2016c): Growth mindset for intelligence. Unter Mitarbeit von Michael Garringer und Mandy Howard. Hg. v. National Mentoring Resource Center Research Board. Washington D.C. Online verfügbar unter <http://www.nationalmentoringresourcecenter.org/index.php/toolkit/item/268-growth-mindset-for-intelligence.html>, zuletzt geprüft am 14.01.2017.
- Bowers, Ed; DuBois David L.; Elledge, Chris; Hawkins, Stephanie; Herrera, Carla; Neblett, Enrique; Alem, Fasika (2016d): Measurement Guidance Toolkit for Mentoring Programs. Unter Mitarbeit von Michael Garringer und Mandy Howard. Hg. v. National Mentoring Resource Center Research Board. Washington D.C. Online verfügbar unter <http://www.nationalmentoringresourcecenter.org/index.php/toolkit/item/239-toolkit.html>, zuletzt geprüft am 20.12.2016.
- Bowers, Ed; DuBois David L.; Elledge, Chris; Hawkins, Stephanie; Herrera, Carla; Neblett, Enrique; Alem, Fasika (2016e): School connectedness. Unter Mitarbeit von Michael Garringer und Mandy Howard. Hg. v. National Mentoring Resource Center Research Board. Washington D.C. Online verfügbar unter <http://www.nationalmentoringresourcecenter.org/index.php/toolkit/item/269-school-connectedness.html>, zuletzt geprüft am 14.01.2017.
- Caspari, Alexandra; Barbu, Ragnhild (2008): Wirkungsevaluierungen – Zum Stand der internationalen Diskussion und dessen Relevanz für die Evaluierung der deutschen Entwicklungszusammenarbeit. Evaluation Working Papers. Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung.
- Castella, Krista de; Byrne, Donald (2015): My intelligence may be more malleable than yours. The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. In: *Eur J Psychol Educ* 30 (3), S. 245–267.

- Collani, Gernot von; Herzberg, Philipp Yorck (2003): Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. In: *Zeitschrift für Differentielle und Diagnostische Psychologie* 24 (1), S. 3–7.
- Daig, Isolde; Gunzelmann, Thomas; Brähler, Elmar (2008): Die Multidimensionale Selbstwertkala (MSWS) von A. Schütz und I. Selling (2006). [Göttingen: Hogrefe, Testmappe komplett 49,00€]. In: *Diagnostica* 54 (7), S. 166–169.
- Digman, John M. (1990): Personality Structure: Emergence of the Five-Factor Model. In: *Annual Review of Psychology* (41), S. 417–440, zuletzt geprüft am 25.01.2017.
- Dresing, Thorsten; Pehl, Thorsten (2015): Praxisbuch Interview, Transkription & Analyse. Anleitungen und Regelsysteme für qualitativ Forschende. 6. Aufl. Marburg: Dresing. Online verfügbar unter http://www.audiotranskription.de/download/praxisbuch_transkription.pdf?q=Praxisbuch-Transkription.pdf, zuletzt geprüft am 14.01.2017.
- DuBois, David L.; Holloway, Bruce E.; Valentine, Jeffrey C.; Cooper, Harris (2002): Effectiveness of mentoring programs for youth: a meta-analytic review. In: *American journal of community psychology* 30 (2), S. 157–197.
- DuBois, David L.; Portillo, Nelson; Rhodes, Jean E.; Silverthorn, Naida; Valentine, Jeffrey C. (2011): How Effective Are Mentoring Programs for Youth? A Systematic Assessment of the Evidence. In: *Psychological science in the public interest : a journal of the American Psychological Society* 12 (2), S. 57–91. Online verfügbar unter <http://journals.sagepub.com/doi/pdf/10.1177/1529100611414806>, zuletzt geprüft am 12.01.2017.
- ECEB Mentoring: European Center for Evidence-Based Mentoring. Unter Mitarbeit von Marc Coenders und Szilvia Simon. Niederlande. Online verfügbar unter <http://www.ecebmentoring.eu/>, zuletzt geprüft am 28.07.2017.
- ECEB Mentoring (2017): Declaration on Mentoring in Europe. Online verfügbar unter <http://www.ecebmentoring.eu/websites/mentorprogrammefriesland.nl/mediadepot/16892e2daae2.pdf>, zuletzt geprüft am 28.07.2017.
- ETS (2012): Relationship between Big Five and Academic and Workforce Outcomes. Educational Testing Service. Online verfügbar unter https://www.ets.org/s/workforce_readiness/pdf/21334_big_5.pdf, zuletzt geprüft am 13.12.2016.

- Feu Gelis, Jordi (2015): How an intervention project contributes to social inclusion of adolescents and young people of foreign origin. In: *Children and Youth Services Review* 52, S. 144–149.
- Geppl, Monika; Hajek, Wolfgang; Thaller, Andreas (2011): Wirkungsorientierte Steuerung in der österreichischen Bundesverwaltung. In: Helfried Bauer, Peter Biwald und Elisabeth Dearing (Hg.): *Gutes Regieren. Konzepte - Realisierungen - Perspektiven*. Wien, Graz: NWV Neuer Wiss. Verl. (Öffentliches Management und Finanzwirtschaft, 13), S. 421–430.
- Gettings, P. E.; Wilson, S. R. (2014): Examining commitment and relational maintenance in formal youth mentoring relationships. In: *Journal of Social and Personal Relationships* 31 (8), S. 1089–1115.
- Gläser, Jochen; Laudel, Grit (2013): Life With and Without Coding: Two Methods for Early-Stage Data Analysis in Qualitative Research Aiming at Causal Explanations. In: *Forum Qualitative Sozialforschung/ Forum: Qualitative Social Research* 14 (2), Rn. 1–96.
- Goldberg, Lewis R. (1992): The Development of Markers for the Big-Five Factor Structure. In: *Psychological Assessment* 4 (1), S. 26–42.
- Gössler, Martin (2014): Wirkungen und Nebenwirkungen der Wirkungsorientierung. Eine Packungsbeilage für das neue Managementrezept von NPOs. In: *OrganisationsEntwicklung* (3), S. 67–73.
- Haunberger, Sigrid; Baumgartner, Edgar (2017): Wirkungsevaluation in der Sozialen Arbeit mittels Realistic Evaluation: empirische Anwendungen und methodische Herausforderungen. Eine systematische Literaturreview. In: *Zeitschrift für Evaluation* 16 (1), S. 121–145.
- Helfferich, Cornelia (2011): *Die Qualität qualitativer Daten. Manual für die Durchführung qualitativer Interviews*. 4. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden GmbH.
- Herrera, Carla; DuBois David L.; Grossmann, Jean Baldwin (2013): *The Role of Risk. Mentoring Experiences and Outcomes for Youth with varying Risk Profiles*. New York: A Public/Private Ventures project distributed by MDRC.

- Himmelbauer, Magdalena Clarissa (2016): Evaluierung des Nightingale-Projektes anhand der grammatikalischen Fähigkeiten im deutschen Zweitspracherwerb. Masterarbeit. Universität Wien, Wien. Psychologie.
- Kaiser, Robert (2014): Qualitative Experteninterviews. Konzeptionelle Grundlagen und praktische Durchführung. Wiesbaden: Springer VS.
- Karcher, Michael J.; Sass, Daniel (2010): A multicultural assessment of adolescent connectedness: testing measurement invariance across gender and ethnicity. In: *Journal of counseling psychology* 57 (3), S. 274–289.
- Kellogg Foundation, W. K. (2004): Logic Model Development Guide. Using Logic Model to Bring Together Planning, Evaluation, and Action. Michigan. Online verfügbar unter <https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf>, zuletzt geprüft am 11.08.2017.
- Kofler, Verena (2017): Die „Stärkung der Bildungsmotivation“ als SMARTes Wirkungsziel im Mentoring-Programm The Nightingale, Wien. In: IMC FH Krems GmbH (Hg.): 11. FH-Forschungsforum 2017. Research-Innovation-Value. Tagungsband. 19. und 20. April 2017. Krems, S. 428–437. Online verfügbar unter <http://ffho-arep.fh-ooe.at/handle/123456789/943>, zuletzt geprüft am 19.08.2017.
- Krause, Tobias; Proeller, Isabelle (2017): New Public Management (NPM). Hg. v. Gabler Wirtschaftslexikon. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Definition/new-public-management-npm.html>, zuletzt geprüft am 23.06.2017.
- Kurz, Bettina; Kubek, Doreen (2013): Kursbuch Wirkung. Das Praxishandbuch für alle die Gutes noch besser tun wollen : mit Schritt-für-Schritt-Anleitungen & Beispielen. 1. Auflage. Berlin: PHINEO gAG. Online verfügbar unter <http://www.phineo.org/fuerorganisationen/kursbuch-wirkung>, zuletzt geprüft am 02.12.2016.
- Leeb, Günther (2014): "Nightingale": Stärkung von bildungsbenachteiligten Kindern durch Mentorin in der Freizeit. Programm 2014. Hg. v. Österreichische Kinderfreunde. Initiative Interkulturelle Arbeit - gemeinsam. Wien.
- Leeb, Günther (2016): Du kannst die Welt eines Kindes retten! Als Mentor/in eines/einer Schüler/in... Österreichische Kinderfreunde. Wien.
- Leutwyler, Bruno; Aegerter, Miriam; Meierhans, Claudia (2014): "Nightingale" in Teacher Education: Programm Evaluation. Research Report. Zug.

- Lönroth, Carina Sild (2007): *The Nightingale scheme - A Song For The Heart*. Training reports 2/2007. Holmbergs i Malmö AB: Malmö University.
- Lönroth, Carina Sild (2013): *Näktergalen. The Effects of Mentoring*. Online verfügbar unter https://www.mah.se/upload/Student/N%C3%A4ktergalen/N%C3%A4ktergalen%202014/CarinaSL%C3%B6nroth30_T.pdf, zuletzt geprüft am 29.06.2017.
- Mayring, Philipp (2002): *Qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. 5. Aufl. Weinheim und Basel: Beltz Verlag.
- Melling, Atethea; Gurjee, Ridwanah (2013): *Researching the Impact of Student Mentoring in the Community*. In: *International Journal of Social Work and Human Services Practice* (1:1), S. 1–8.
- Mervielde, Ivan; Fruyt, Filip de (1999): *Construction of the Hierarchical Personality Inventory for Children (HIPIC)*. In: Ivan Mervielde, Ian Deary, Filip de Fruyt und Fritz Ostendorf (Hg.): *Personality Psychology in Europe. Proceedings of the Eighth European Conference on Personality Psychology*. Tilburg: Tilburg University Press, S. 107–127.
- Oelerich, Gertrud; Otto, Hans-Uwe (Hg.) (2011): *Empirische Forschung und Soziale Arbeit. Ein Studienbuch*. Wiesbaden: VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden, Wiesbaden.
- Plé, Bernhard (2015): *Wirkungsanalysen und ihre Forschungsdesigns: Kritische Gegenüberstellung ihrer konzeptionellen und methodischen Herangehensweisen im Hinblick auf mögliche Komplementaritäten und Anwendungen für die Beratungswissenschaften*. In: *Zeitschrift für Beratungs- und Managementwissenschaften* 2. Jahrgang, S. 25–38.
- Polutta, Andreas (2010): *Wirkungsorientierung und Profession. Neue Professionalisierung oder Ende professioneller Sozialer Arbeit?* In: *Soz Passagen* 2 (1), S. 47–62.
- Rammstedt, Beatrice; John, Oliver P. (2014): *BFI-10. Big Five Inventory-10*. In: Christoph J. Kemper, A. E. Abele, Elmar Brähler und Markus Zenger (Hg.): *Psychologische und sozialwissenschaftliche Kurzskalen. Standardisierte Erhebungsinstrumente für Wissenschaft und Praxis*. Berlin, Germany: Medizinisch Wissenschaftliche Verlagsgesellschaft, S. 35–45.

- Rauthmann, John (2016a): Lexikalischer Ansatz. Dorsch Lexikon der Psychologie. Online verfügbar unter <https://portal.hogrefe.com/dorsch/lexikalischer-ansatz/>, zuletzt geprüft am 09.07.2017.
- Rauthmann, John (2016b): Normierung. Dorsch Lexikon der Psychologie. Online verfügbar unter <https://portal.hogrefe.com/dorsch/normierung/>, zuletzt geprüft am 11.08.2017.
- Reidinger, Veronika (2015): Wirkungsorientierung in der Sozialen Arbeit. Professionalisierungsprojekt oder managerialistische "cookbook social work"? In: *soziales_kapital* (14), S. 171–180.
- Rhodes, Jean; Reddy, Ranjini; Roffman, Jennifer; Grossman, Jean B. (2005): Promoting successful youth mentoring relationships: a preliminary screening questionnaire. In: *The journal of primary prevention* 26 (2), S. 147–167.
- Rhodes, Jean E. (2004): Stand by Me. The Risks and Rewards of Mentoring Today's Youth. Cambridge, Mass: Harvard University Press.
- Rhodes, Jean E. (2005): A Model of Youth Mentoring. In: David L. DuBois und Michael J. Karcher (Hg.): Handbook of youth mentoring. Thousand Oaks CA u.a.: Sage, S. 30–43.
- Sabeva, Nadejda (2012): Der Unterschied zwischen Vätern und Müttern in der Einschätzung der Persönlichkeit ihrer Kleinkinder im Junior Temperament und Charakter Inventar. Diplomarbeit, Universität Wien. Psychologie.
- Scherer, Hanna Sabine (2016): Evaluation des Projekts 'Nightingale-Mentoring'. Analyse des Mentoring-Prozesses, Universität Wien. Psychologie.
- Schober, Barbara; Finsterwald, Monika; Wagner, Petra; Spiel, Christiane: Lebenslanges Lernen als Herausforderung der Wissensgesellschaft: Die Schule als Ort der Förderung von Bildungsmotivation und selbstreguliertem Lernen. In: Sprecht (Hg.) 2009 – Nationaler Bildungsbericht, S. 124–143.
- Schober, Christian; Rauscher, Olivia (2014): Was ist Impact? Gesellschaftliche Wirkung von (Nonprofit) Organisationen. Von der Identifikation über die Bewertung bis zu unterschiedlichen Analyseformen. Wien.
- Sozialministerium (o. J.): Nationaler Sozialbericht Österreich 2014. Hg. v. Sozialministerium. Wien, zuletzt geprüft am 15.11.2016.

- Spora Sinergies (2017): Evaluation of the mentoring program "Rossinyol". Social and educational evaluation. (Power-Point-Folien). Barcelona, 15.06.2017. Online verfügbar unter <https://drive.google.com/file/d/0B8yP7QdK6KJ-ZGFPZHgwNXZQeXc/view>, zuletzt geprüft am 28.07.2017.
- SRI - Social Reporting Initiative e.V. (2017a): Beispielberichte. Online verfügbar unter <http://www.social-reporting-standard.de/anwender/beispielberichte/>, zuletzt geprüft am 13.08.2017.
- SRI - Social Reporting Initiative e.V. (2017b): Über uns. Online verfügbar unter <http://www.social-reporting-standard.de/sri-ev/ueber-uns/>, zuletzt geprüft am 13.08.2017.
- Täubner, Mischa (2016): Das Leben der Moglis. Soziale Organisationen sind heute gefordert, für Kosten und Nutzen ihrer Arbeit Zeugnis abzulegen. Aber lässt sich das wirklich berechnen? In: *brand eins - Wirtschaftsmagazin* 18 (04), S. 65–67.
- The Nightingale Mentoring Network (2017a): Members. Online verfügbar unter http://nightingalementoring.org/?page_id=197, zuletzt geprüft am 25.06.2017.
- The Nightingale Mentoring Network (2017b): Paid fee and the Nightingale Finances. Online verfügbar unter <http://nightingalementoring.org/?p=4430>, zuletzt geprüft am 25.06.2017.
- The Nightingale Mentoring Network (2017c): The Network. Online verfügbar unter http://nightingalementoring.org/?page_id=81, zuletzt geprüft am 25.06.2017.
- The Nightingale Mentoring Network (2017d): The Nightingale concept. Online verfügbar unter <http://nightingalementoring.org/?cat=35>, zuletzt geprüft am 02.12.2016.
- Trautsmawieser, Andrea (2016): Evaluierung des Projektes Nightingale-Mentoring: Fremdevaluierung der Mentoring-Ergebnisse. Masterarbeit. Universität Wien, Wien. Psychologie.
- Uebelhart, Beat; Zängl, Peter (2013): Wirkungsevaluation im Social-Impact-Modell. In: Beat Uebelhart und Peter Zängl (Hg.): *Praxisbuch zum Social-Impact-Modell*. Baden-Baden: Nomos (36), S. 269–288.
- Wöhler, Veronika; Arztmann, Doris; Wintersteller, Teresa; Harrasser, Doris; Schneider, Karin (2017): Partizipative Aktionsforschung mit Kindern und Jugendlichen. Von

Schulsprachen, Liebesorten und anderen Forschungsdingen. Wiesbaden: Springer VS.

Ziegler, H. (2010): Ist der experimentelle Goldstandard wirklich Gold wert für eine Evidenzbasierung der Praxis Früher Hilfen? In: *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 53 (10), S. 1061–1066, zuletzt geprüft am 25.05.2017.

Zierer, Brigitta (2017): Soziale Arbeit: von der Einzelfallhilfe zur wirkungsorientierten Praxis. In: Armin Wöhrle, Agnès Fritze, Thomas Prinz und Gotthart Schwarz (Hg.): *Sozialmanagement – Eine Zwischenbilanz*. Wiesbaden: Springer Fachmedien Wiesbaden, S. 65–83.

Zilberstein, Karen; Spencer, Renée (2014): Breaking bad. An attachment perspective on youth mentoring relationship closures. In: *Child & Family Social Work*, S. 1–10.

Abbildungsverzeichnis

Abbildung 1: Instrumente zur Messung von Bildungsmotivation und Selbstwert (eigene Darstellung, erweitert nach Kofler 2017, S. 443).	49
Abbildung 2: Berechnung der Outcomes anhand der Double-Difference-Methode (eigene Darstellung, erweitert nach Kofler 2017, S.443)	50
Abbildung 3: Angaben der MentorInnen (n=92) in der Onlinebefragung für ganz Österreich. Frage: Welche Kompetenzen hat dein Mentee durch das Projekt erworben. Antwortoption: Mehr Selbstvertrauen. Eigene Darstellung.	60
Abbildung 4: Angaben der MentorInnen (n=92) in der Onlinebefragung für ganz Österreich. Frage: Welche Kompetenzen hat dein Mentee durch das Projekt erworben? Antwortoption: ‚Mehr Motivation zum Lernen‘. Eigene Darstellung.	61

Tabellenverzeichnis

Tabelle 1: Wirkungstreppe nach Kurz und Kubek 2013, S. 35 (eigene Darstellung).	11
Tabelle 2: Eigene Kategorisierung der Ziele für Mentees nach dem Informationsblatt von ‚Nightingale‘ Wien für den Jahrgang 2016 (Leeb 2016).	26
Tabelle 3: Übersicht über die durchgeführten ExpertInnen-Interviews einschließlich Durchführungsdatum und Dauer.	33
Tabelle 4: Übersichtliche Darstellung der ausgewählten Items aus den Daten der Onlinebefragung der MentorInnen für das Mentoring-Programm ‚Nightingale‘, Jahrgang 2016/17 (eigene Darstellung).	38
Tabelle 5: Übersicht zur Literaturrecherche der Erhebungsinstrumente (eigene Darstellung).	48
Tabelle 6: Bezirke und Anzahl der Schulen, die an ‚Nightingale‘ in Wien teilnehmen.	53
Tabelle 7: Übersicht der ausgewählten Instrumente für das Forschungsdesign mit den limitierenden Faktoren.	86

Anhang

Kategoriensystem vor Kodierung

Code-Kategorie	Abkürzung	Definition/Fragestellung
Programmverlauf/Organisation	ORG	Wie ist ‚Nightingale‘ derzeit organisiert, was wird wie gemacht?
Programm-Evaluation	EV	Wie wird ‚Nightingale‘ aktuell evaluiert?
Wirkungen		
Subjektive Wirkungen	S-WIR	Welche Wirkung wird von wem wie eingeschätzt?
Intendierte Wirkungen	I-WIR	Welche Wirkungen sind vom Programm intendiert?
Nicht intendierte Wirkungen	Nicht I-WIR	Welche Wirkungen sind vom Programm nicht intendiert
Bildungsmotivation	BILD	
Selbstwert	SELBST	
MentorInnen		
<i>Incomes</i> MentorInnen	INCOM MENTOR	Was bringen die MentorInnen in das Programm mit?
Beziehung	BEZ	Wie verläuft die Beziehung zwischen MentorIn und Mentee? Was gilt zu beachten? Wo liegen die Unterschiede?
Umsetzbarkeit		
Umsetzbarkeit des Forschungsdesigns	UMS	Inwieweit kann das Forschungsdesign derzeit umgesetzt werden?
Grenzen der Umsetzbarkeit	UMS – GRENZE	Wo liegen die Grenzen
Messpunkte		
Vor dem Programm	MP – VOR	Gibt es Evaluierung vor dem Programm. Kann diese durchgeführt werden?
Dritter Messpunkt	MP – 3	Wann macht ein dritter Messpunkt Sinn?

Vergleichsgruppe

Vergleichsgruppe	VG	Wie kann eine Vergleichsgruppe gebildet werden? Was muss bedacht werden?
Psychologische Tools		
Anwendung von psychologischen Tools	PSY	Was muss bedacht werden, wenn psychologische Tools angewendet werden?
Wirkungsmessung		Wie stehen die Befragten zur Wirkungsmessung?
Wirkungsmessung Pro	WM +	
Wirkungsmessung Contra	WM -	
Ressourcen	RES	Welche Ressourcen werden gebraucht, Grenzen der Ressourcen

Kategoriensystem nach Kodierung (Output MAXQDA)

Liste der Codes	Memo	#
Codesystem		446
Organisation	Organisation des Programmes. Wie ist ‚Nightingale‘ organisiert, was wird wie gemacht?	2
PsychologiestudentInnen als MentorInnen		4
Abbrüche		3
PH Wien		11
Auswahl der Schulen		1
Dauer des Programmes		6
Ressourcen		6
Fördergeber		13
Auswahl der Mentees		20
Eltern		3
Evaluierung	Evaluierung: Wie wird ‚Nightingale‘ aktuell evaluiert?	10

Evaluierung Netzwerk		2
Evaluierung Fragebögen MentorInnen 2017		5
Sprachtest Uni Wien		12
Lehrpersonen		4
MentorInnen		2
Befragung von Kindern		12
Wirkungsanalyse		2
Fördergeber Wirkungsanalyse		3
Ressourcen Wirkungsanalyse		5
Akzeptanz		1
Akzeptanz Eltern		2
Akzeptanz Lehrpersonen		8
Akzeptanz/Mehraufwand Eltern		3
Akzeptanz/Mehraufwand MentorInnen		5
Vergleichsgruppe		13
SMART Ziele		7
Psychologische Tools		1
Übersetzung der Instrumente aus dem Englischen		2
Anwendung	Was muss beachtet werden?	12
<i>Incomes</i> Mentees		6
<i>Incomes</i> MentorInnen		5
Wirkungsmodell		37
Langfristigkeit		13
Dritter Messpunkt		13
Umsetzbarkeit	Inwieweit kann das Forschungsdesign umgesetzt werden?	17
Umsetzbarkeit Grenzen		16
Wie kann die Wirkung gemessen werden?		12
Reflexion mit Kinder		8

Intendierte Wirkungen	12
Unintendierte Wirkungen	9
Selbstwert	30
Bildungsmotivation	22
Quantitativ	4
Qualitativ	17
Wirkungsmessung Contra	12
Wirkungsmessung pro	9
Professionalität	1
Effizienz	2
ECEB - Kooperation	3

Onlinebefragung der MentorInnen Nightingale 2016/17

Nr. Frage	Beantwortet von:	Übersprungen von:
F1 Du studierst ...	95	0
F2 Wie würdest Du die Rolle(n) beschreiben, die Du für Deinen Mentee hattest?	95	0
F3 Wie würdest Du die Rolle(n) beschreiben, die Du für die Eltern Deines Mentee hattest?	91	4
F4 Hast Du Dich in der Projektlaufzeit ausreichend betreut gefühlt?	93	2
F5 Wo bestünde Verbesserungsbedarf?	45	50
F6 Welche der folgenden Kompetenzen hat Dein Mentee durch das Projekt erworben?	92	3
F7 Du warst/bist als Mentorin tätig in...	92	3
F8 Welche der folgenden Einrichtungen hat Dein Mentee erstmals kennen- gelernt oder erneut besucht (Nur für Teiln. in Wien)?	58	37

F9	Welche praktischen Fertigkeiten hat Dein Mentee mit Deiner Hilfe im Projekt erlangt?	92	3
F10	Von welcher der durchgeführten Aktivitäten hat Dein Mentee am meisten profitiert? Und warum?	90	5
F11	Kann Dein Mentee seine/ihre Muttersprache lesen und schreiben?	92	3
F12	Besucht Dein Mentee den muttersprachlichen Unterricht oder einen Deutsch-Förderkurs (oder beides)?	83	12
F13	Was würdest Du als Deinen eigenen größten Lerngewinn durch das Projekt beschreiben?	91	4
F14	Was hat Dir das Projekt im Hinblick auf Dein Studium und/oder hinsichtlich Deines angestrebten Berufes gebracht?	89	6
F15	Wie gut ist es Dir gelungen, eine Beziehung zu den Eltern herzustellen?	92	3
F16	Warum?	85	10
F17	Was, glaubst Du, können die Eltern aus dem Projekt mitnehmen?	82	13
F18	Welche Unterstützung würden die Eltern am ehesten benötigen, um ihre Kinder im schulischen Lernen besser fördern zu können?	86	9
F19	Wie viele Treffen mit den Mentees hast Du durchgeführt?	85	10
F20	Wie viele Treffen mussten abgesagt werden?	86	9
F21	Warum mussten Treffen abgesagt werden?	84	11
F22	Konntest Du die stattgefundenen Treffen planmäßig gestalten?	89	6
F23	Welche Hindernisse gab es dabei?	70	25
F24	Deine ursprünglichen Erwartungen betrachtend: Warst Du zufrieden mit der Entwicklung der Beziehung zwischen Deinem Mentee und Dir?	89	6
F25	Warum?	76	19
F26	Welche war die schwierigste Phase im Projekt und warum?	87	8
F27	Fandest Du, waren die Gruppentreffen ausreichend?	89	6

F28 Wie hilfreich fandest Du die Gruppentreffen?	88	7
F29 Hast Du Verbesserungsvorschläge (z.B. zur Organisation und zu den Inhalten der Treffen, zu Berichten, zur Vernetzung etc.)?	68	27
F30 Welche Schulnote würdest Du dem Gesamtprojekt geben?	89	6
F31 Hast Du Änderungsvorschläge bezüglich der Organisation (z.B. Training und Informationen vor Beginn der Treffen, Startday etc.)?	60	35
F32 Was Du uns sonst noch mitteilen möchtest:	61	34

Leitfäden

Günther Leeb (Organisator)

Einstieg

Erstmal vielen Dank, dass du dir die Zeit genommen hast – das Interview wird ca. eine Stunde in Anspruch nehmen. Ich werde nun die Aufzeichnung starten, wenn das für dich in Ordnung ist.

Ich möchte mich und mein Vorhaben kurz vorstellen. Wie du weißt, ist mein Name Verena Kofler, ich studiere an der FH Campus Wien im europäischen Masterstudium Sozialwirtschaft und Soziale Arbeit. Nach unserem Gespräch im September habe ich mit Hilfe von Inputs meiner KollegInnen und der Betreuung mein Thema konkretisieren können und habe nun vor, eine Wirkungsanalyse für The ‚Nightingale‘ zu konzipieren, um die langfristige Wirkung des Projektes auf die Mentees feststellen zu können. Gerne erzähle ich dir nach dem Interview mehr über das Vorhaben. Vorher möchte ich dir aber gerne ein paar Fragen stellen. Bitte versuche diese so frei wie möglich zu beantworten, es geht um deine subjektive Sichtweise und Einschätzung. Ist es für dich in Ordnung, dich mit deinem Namen in der Arbeit zu zitieren?

Einstiegsfrage

Du bist nun seit sechs Jahren Projektleiter von The ‚Nightingale‘ Wien. Ich würde dich bitten zu Beginn etwas von dir zu erzählen und wie du zu The ‚Nightingale‘ gekommen bist.

- Warum braucht es The ‚Nightingale‘?

Organisation – Netzwerk

Fragen zum Netzwerk

The ‚Nightingale‘ ist als Mentoring-Netzwerk in ganz Europa organisiert und kommt ursprünglich aus Malmö. Wie kommt The ‚Nightingale‘ nach Wien?

- Aus welchem Grund engagieren sich die Kinderfreunde in diesem Bereich?
- Ist The ‚Nightingale‘ gegenüber den Kinderfreunden erfolgsnachweispflichtig?
- Wie funktioniert das Netzwerk: Abhängigkeiten, Verantwortlichkeiten, Freiheiten?
- Was genau wird vom Netzwerk zur Verfügung gestellt? Woran muss man sich halten im Rahmen der Franchise-Beziehung?
- Welche Lizenzgebühren sind als Partner an das Netzwerk zu bezahlen?

Praxis - Evaluierung

Fragen zur Evaluierung

Ist in Bezug auf die Arbeit von The ‚Nightingale‘ eine verpflichtende Evaluierung vorgesehen und wenn ja von wem (‚Nightingale‘ selber, Kinderfreunde, Netzwerk, Fördergeber)?

- Von wem wird die Evaluation durchgeführt (intern/extern)?
- Mit welcher Begründung wird die Evaluation durchgeführt?
- Wie wird die Evaluierung derzeit durchgeführt (Verfahren)?
- Du hast mir beim letzten Gespräch schon mitgeteilt, dass es schwierig ist, von den Mentees eine Einschätzung/Bewertung zu erhalten. Kannst du auf diesen Punkt noch einmal näher eingehen?
- Wo siehst du Lücken, Verbesserungspotential oder Vernachlässigtes?

Wirkungsanalyse – Risiken und Chancen

Fragen zu Risiken und Chancen der Wirkungsanalyse

Die Soziale Arbeit ist dabei, sich umfassend zu verändern: Wirkungsorientierung wird zum zentralen Element. In einem Text (Polutta 2010), den ich kürzlich gelesen habe, ist mir bewusst geworden, wie wichtig es ist, dass eine Wirkungsanalyse umsichtig gestaltet.

Ich zitiere die deutsche Bundesregierung:

„Dringend muss die Lücke im Bereich der Jugendhilfe-Wirkungsforschung geschlossen werden; Jugendhilfe muss ihre Erfolge auch mit ‚harten Fakten‘ beweiskräftiger machen.“ Bundesregierung DE, 2005

Weiters heißt es vom BMFSFJ:

„Allerorten fehlt es an entsprechenden Studien und vor allem an dem Bewusstsein, dass Evaluation ein unverzichtbares Moment jeglicher professionellen Praxis ist.“
BMFSFJ, 2002

Wie siehst du eine solche Entwicklung – welche Chancen und Risiken erkennst du?

Machst du einen Unterschied zwischen professioneller/persönlicher Meinung?

- Wurden an dich schon Forderungen nach wirkungsorientierter Evaluierung herangetragen (z.B. von Seiten der Fördergeber -> BMBF, Kinderfreunde)?
- Welche konkreten positiven und negativen Effekte können Wirkungsorientierung und darauf aufbauende Steuerung deiner Meinung nach auf The ‚Nightingale‘ haben?

Wirkungsziele

Allgemeine Fragen zu Wirkungszielen

Im Informationsblatt zu The ‚Nightingale‘ für den Jahrgang 2016 sind sieben Ziele in Bezug auf die Mentees formuliert. Wer hat diese Ziele festgelegt?

- Was sind die Wirkungsziele (Mentees) aus deiner Sicht?
- Wie schätzt du die Wirkung von The ‚Nightingale‘ ein?

Fragen zur Operationalisierung der Wirkungsziele

Aus unserem letzten Gespräch weiß ich, dass es keine langfristige Evaluierung von ‚Nightingale‘ gibt. Die Ziele beinhalten aber Änderungen in der Verhaltensweise der TeilnehmerInnen, diese können kurzfristig nicht gemessen werden.

Daher stellt sich die Frage, wie die (langfristige) Wirkung des Mentoring-Programmes The ‚Nightingale‘ Wien, in Bezug auf Mentees, festgestellt werden kann. In meinem Forschungsvorhaben einer Wirkungsanalyse soll ein Tool entwickelt werden, das es in Zukunft möglich macht, den Outcome und Impact von The ‚Nightingale‘ zu erfassen.

Du hast in deiner eigenen Magisterarbeit eine Evaluierung durchgeführt. Im Folgenden würde ich gerne mit dir die formulierten Ziele durchgehen und nach deinen ersten Gedanken bitten, wie man diese Ziele überprüfen könnte – Anhand welcher Kriterien würde es funktionieren, anhand welcher nicht? Gibt es Theorien zu der Entwicklung der jeweiligen Ziele? Orientieren sich die Ziele an bestimmten Disziplinen? Sind kreative Ansätze zur Messung (Beobachtung) möglich?

- Selbstwert stärken
- Orientierung in Wien verbessern
- Kennenlernen der Bildungs- und Freizeitmöglichkeiten in der Stadt
- Kompetenzen erwerben wie Mobilität, Selbsteinschätzung
- Kulturelle Einrichtungen kennenlernen
- Bildungsmotivation stärken durch den Kontakt mit einer Studentin, die den in Österreich höchstmöglichen Bildungsabschluss anstrebt
- Unterstützung der Familie bei Lern- und Bildungsfragen.

Ich danke dir ganz herzlich für deine Zeit und deine Informationen. Wir werden auch weiterhin in Kontakt bleiben, aber gerne biete ich dir jetzt schon mal an, dass du die Masterarbeit nach Veröffentlichung von mir geschickt bekommst.

Elisabeth Leskowa (Organisatorin)

Einstieg

- Darf ich aufzeichnen?
- Darf ich Ihren Namen nennen oder lieber anonymisieren?
- Nach einem Gespräch mit Günther Leeb habe ich bereits erfahren, dass Sie TH als Wahlpflichtfach an der PH anbieten, auch habe ich schon von Praxisvolkschulen gehört. Können Sie mir zum Einstieg ein bisschen davon erzählen, wie The ‚Nightingale‘ an der PH organisiert wird und welche Rolle Sie dabei spielen.
- Wie wählen Sie die Kinder aus?
- Gibt es Unterschiede zu den anderen Programmen in Wien, zb die von Leeb?

Wie würden Sie die Wirkung von The ‚Nightingale‘ auf die Kinder beschreiben?

- Ist Ihrer Meinung nach eine standardisierte Form der Erhebung der Entwicklung der Mentees umsetzbar? Ist die Umsetzung einer langfristigen Wirkungsmessung mit drei Messpunkten durchführbar?
- Müssen die Ziele konkretisierter werden - SMART Ziele, Herausforderungen. Sollwerte, Creaming-Effekt, ...

SWOT

Welche Stärken, Schwächen, Chancen und Risiken können in der bis jetzt entwickelten Form des Programmes/der Evaluierung gesehen werden.

Welche Stärken, Schwächen, Chancen und Risiken werden in der Wirkungsmessung grundsätzlich gesehen?

Szenario

Phase 1: Szenariofeld: Wirkungsanalyse wird in das Programm eingeführt.

Schlüsselfaktor 1 – Was wäre, wenn eine verpflichtende Wirkungsmessung von FördergeberInnen gefordert würde?

Schlüsselfaktor 2 – von TN selbst entwickelt

Schlüsselfaktor 3 - Langfristigkeit

Phase 2: Schlüsselfaktoren filtern und danach genauer betrachten

Phase 3: einzelne Schlüsselfaktoren werden daraufhin analysiert, welche möglichen zukünftigen Ausprägungen jeweils vorstellbar sind.

Phase 4: aus den verschiedenen Ausprägungen der diskutierten Entwicklungen werden Szenarien entwickelt – in sich stimmige Bilder möglicher Zukünfte.

Phase 5: Szenarien werden weiterverwendet.

Praxis – Evaluierung

Fragen zur Evaluierung

Ist in Bezug auf die Arbeit von The ‚Nightingale‘ eine verpflichtende Evaluierung von Seiten der PH vorgesehen?

- Von wem wird die Evaluation durchgeführt (intern/extern)?

- Mit welcher Begründung wird die Evaluation durchgeführt?
- Wie wird die Evaluierung derzeit durchgeführt (Verfahren)?
- Du hast mir beim letzten Gespräch schon mitgeteilt, dass es schwierig ist, von den Mentees eine Einschätzung/Bewertung zu erhalten. Kannst du auf diesen Punkt noch einmal näher eingehen?
- Wo siehst du Lücken, Verbesserungspotential oder Vernachlässigtes?

Wirkungsanalyse – Risiken und Chancen

Fragen zu Risiken und Chancen der Wirkungsanalyse

Die Soziale Arbeit ist dabei, sich umfassend zu verändern: Wirkungsorientierung wird zum zentralen Element. In einem Text (Polutta 2010), den ich kürzlich gelesen habe, ist mir bewusst geworden, wie wichtig es ist, dass eine Wirkungsanalyse umsichtig gestaltet.

Ich zitiere die deutsche Bundesregierung:

„Dringend muss die Lücke im Bereich der Jugendhilfe-Wirkungsforschung geschlossen werden; Jugendhilfe muss ihre Erfolge auch mit ‚harten Fakten‘ beweiskräftiger machen.“ Bundesregierung DE, 2005

Weiters heißt es vom BMFSFJ:

„Allerorten fehlt es an entsprechenden Studien und vor allem an dem Bewusstsein, dass Evaluation ein unverzichtbares Moment jeglicher professionellen Praxis ist.“
BMFSFJ, 2002

Wie siehst du eine solche Entwicklung – welche Chancen und Risiken erkennen Sie?

Wurden an dich schon Forderungen nach wirkungsorientierter Evaluierung herangetragen (z.B. von Seiten der Fördergeber -> BMBF, Kinderfreunde)?

- Welche konkreten positiven und negativen Effekte können Wirkungsorientierung und darauf aufbauende Steuerung deiner Meinung nach auf The ‚Nightingale‘ haben?
- Wie viel mehr würde eine Wirkungsmessung an Ressourcen (€) brauchen?

Carina Sild Lönroth (Organisatorin)

1. What are the most important effects of The ‚Nightingale‘ on the mentees in your opinion?
2. Are there standardized methods of evaluating the members of the ‚Nightingale‘ Network?
3. And if so, how do you consider the results of the evaluations of individual countries/members?
4. Do you know if there are any countries/members which carry out methods of longterm-evaluation?
5. Are the strategies/results from different countries being compared? What in your opinion, could be the advantages and disadvantages of doing that?
6. When did you start cooperating with ECEB? What were the reasons/motivations for it?
7. What is your expectation on the cooperation with ECEB?
8. Which meaning has the cooperation with ECEB for the ‚Nightingale‘ network?
9. Which opportunities and risks do you see in applying evidence based research for mentoring programs?
10. A long term research on effects needs at least three chronological points of data collection – one before the beginning of the program, one at the end of the program and a last, not yet defined one. At which point in the program would you suggest setting the third measurement point?

Nadejda Sabeva (Organisatorin)**Einstieg**

- Darf ich aufzeichnen?
- Darf ich Ihren Namen nennen?
- Vorstellung

Eigene Rolle

- Können Sie mir zum Einstieg etwas über Sie und **Ihre Rolle bei ‚Nightingale‘** erzählen?

Wirkung

-
- **Wie würden Sie die Wirkung von The ‚Nightingale‘ auf die Kinder beschreiben?**
-
- Wie könnte man Ihrer Meinung nach die Wirkung des Programms auf die Kinder am ehesten messbar machen?

Psychologische Tools

- Ich habe ansatzweise ein theoretisches Forschungsdesign entwickelt, das auf das Big Five Inventar für Kinder und Jugendliche nach Bleidorn und Ostendorf aufbaut. Offenheit und Gewissenhaftigkeit korrelieren mit Bildungsmotivation. Sehen Sie dies auch so, bzw. wie schätzen Sie die Umsetzung diese Tools – in Kombination mit anderen, zumeist psychologischen Fragebögen, ein (Beispiel Selbstwertskalen)?
- Soweit ich informiert bin, braucht es zur Anwendung bzw. zur Auswertung von psychologischen Fragebögen ausgebildete PsychologInnen. Können Sie dies bestätigen? Wie arbeitsintensiv wäre eine solche Auswertung? Wie viele Ressourcen würde dies zusätzlich benötigen?
- MGT -> Fragen Englisch, können diese – auch meist psychologischen Instrumente – „einfach“ auf Deutsch übersetzt werden?

Umsetzung und dritter Messpunkt

- Ist Ihrer Meinung nach eine standardisierte Form der Erhebung der Entwicklung der Mentees umsetzbar? Ist die Umsetzung einer langfristigen Wirkungsmessung mit drei Messpunkten durchführbar?
- Wann schätzen Sie einen dritten Befragungszeitpunkt am sinnvollsten ein?

Reflexion/Befragung Kinder

- Findet eine regelmäßige Reflexion mit den Mentees statt?
- Kinder befragen – wie schätzen Sie dies ein?

Evaluierung/Wirkungsmessung

- Wo sehen Sie Lücken in der derzeitigen Evaluierung?

- Sehen Sie einen verpflichtenden Nachweis der Wirkung des Programmes kommen?
- Wie stehen Sie zur Wirkungsmessung von sozialen Programmen, konkret von TN?

SWOT

Welche Stärken, Schwächen, Chancen und Risiken können in der bis jetzt entwickelten Form des Programmes/der Evaluierung gesehen werden.

Welche Stärken, Schwächen, Chancen und Risiken werden in der Wirkungsmessung grundsätzlich gesehen?

Szenario

Phase 1: Szenariofeld: Wirkungsanalyse wird in das Programm eingeführt.

Schlüsselfaktor 1 – Was wäre, wenn eine verpflichtende Wirkungsmessung von FördergeberInnen gefordert würde?

Schlüsselfaktor 2 – von TN selbst entwickelt

Schlüsselfaktor 3 - Langfristigkeit

Phase 2: Schlüsselfaktoren filtern und danach genauer betrachten

Phase 3: einzelne Schlüsselfaktoren werden daraufhin analysiert, welche möglichen zukünftigen Ausprägungen jeweils vorstellbar sind.

Phase 4: aus den verschiedenen Ausprägungen der diskutierten Entwicklungen werden Szenarien entwickelt – in sich stimmige Bilder möglicher Zukünfte.

Phase 5: Szenarien werden weiterverwendet.

Lehrpersonen

1. In welcher Klasse unterrichten Sie, bzw. welche Klasse ist Teil von ‚Nightingale‘?
2. Ist es das erste Jahr mit Teilnahm an ‚Nightingale‘ für Sie?
3. Wenn nicht das erste Jahr: Welche Wirkung hat die Teilnahme am Programm, Ihrer Meinung nach, auf die Kinder?
4. Wie viele Kinder aus Ihrer Klasse/Ihren Klassen nehmen an ‚Nightingale‘ teil?
5. Wie viele Jungs, wie viele Mädchen?
6. Wie wählen Sie die Kinder aus?

7. Wie viele Kinder gibt es dieses Jahr, die in das Programm mit aufgenommen worden wären, wenn es mehr Plätze gäbe?
8. Was protokollieren Sie?
9. Welche Veränderungen merken Sie bei den Kindern?
10. Wie hoch ist der Aufwand für Sie?
11. Finden Reflexionen mit den Kindern statt?
12. Dritter Messpunkt
13. Umsetzbarkeit eine langfristigen Wirkungsanalyse
14. Umsetzbarkeit einer Vergleichsgruppe – Kinder aus derselben Klasse oder andere Klasse?

MentorInnen

Vorstellung – Aufzeichnung

1. Ist dein Mentee ein Junge oder ein Mädchen?
 - 1.1. Wie alt ist dein Mentee?
2. Wirst du dich mit deinem Mentee weiterhin treffen? Warum?
3. Hast du bemerkt, dass sich dein Mentee seit dem Kennenlernen verändert hat?
 - 3.1. Wie hat sich dein Mentee verändert?
 - 3.2. Wie hast du das bemerkt?
4. Was hat dein Mentee deiner Meinung nach vom Programm mitgenommen?
 - 4.1. Wie würdest du die Bildungsmotivation deines Mentees einstufen, als du ihn/sie kennengelernt hast? Wie schätzt du ihre/seine Bildungsmotivation heute ein? 1–5 (1 sehr gut, 5 sehr schlecht)
 - 4.2. Wie würdest du den Selbstwert deines Mentees einstufen, als du ihn/sie kennengelernt hast? Wie schätzt du ihren/seinen Selbstwert heute ein? 1–5 (1 sehr gut, 5 sehr schlecht)
5. Wärest du bereit, neben den bisherigen monatlichen Aufzeichnungen auch ganz konkrete Fragen zur Veränderung deines Mentees zu beantworten?
 - 5.1. Zum Beispiel am Anfang und am Ende des Programmes
 - 5.2. Oder regelmäßiger, z.B. monatlich?
6. Wie könnte man deiner Meinung und deiner Erfahrung nach am besten feststellen, wie sich die Mentees durch das Programm verändert haben, bzw. was sie vom Programm mitnehmen?

7. Würdest du an einem weiteren Treffen wie heute teilnehmen?
 - 7.1. Wann sollte ein solches Treffen stattfinden?

Eltern

1. Ist Ihr Kind ein Junge oder ein Mädchen?
 - 1.1. Wie alt ist Ihr Kind?
2. Hatte Ihr Kind eine Studentin oder einen Studenten?
3. Ist Ihr Kind das erste Kind, das bei diesem Programm mitgemacht hat?
4. Wie hat sich Ihr Kind durch den Studenten/die Studentin verändert?
 - 4.1. Was ist Ihnen aufgefallen?
 - 4.2. Haben Sie an Ihrem Kind etwas beobachtet, dass es erst macht oder sagt, seit es den/die StudentIn kennt?
5. Nächstes Jahr machen andere Kinder bei diesem Programm mit – was würden Sie den Eltern sagen, warum ihr Kind mitmachen soll (oder nicht mitmachen soll)?
 - 5.1. Was würden Sie den Eltern sagen, was das Kind davon hat?
6. Möchten Sie, dass sich Ihr Kind weiter mit dem/der StudentIn trifft?
7. Wären Sie bereit, konkrete Fragen zur Veränderung Ihres Kindes zu beantworten?
 - 7.1. Zum Beispiel am Anfang und am Ende des Programmes
 - 7.2. Oder regelmäßiger, z.B. monatlich?
8. Würden Sie an einem weiteren Treffen wie heute teilnehmen?
 - 8.1. Wann sollte ein solches Treffen stattfinden?

Mentees

Vorstellung: Ich möchte mit dir ein echtes Interview machen, wie für die Zeitung, darum hab ich ein paar Fragen über deine Zeit mit den Studenten. Ich werde das Interview aufnehmen, weil ich mir nicht alles merken kann =)

1. Wie **alt** bist du?
 - 1.1. Hast du Geschwister?
2. Was ist dein **Berufswunsch**, was möchtest du werden, wenn du groß bist?
3. Hattest du einen **Studenten** oder eine **Studentin**?
 - 3.1. Hast du einen Mann oder eine Frau?
 - 3.2. Hättest du lieber einen Mann oder eine Frau gehabt?

4. Kannst du dich erinnern, ganz am Anfang als du deine/n Studenten/Studentin gesehen hast, wie hast du dich dabei gefühlt?
 - 4.1. War das komisch für dich, oder hast du dich gleich wohlfühlt bei ihm/ihr?
5. Hast du das Gefühl, dass du jetzt mehr Freude am Lernen hast oder ist es gleich?
 - 5.1. Hast du es genossen, dass er/sie sich Zeit nur für dich alleine genommen hat/nur für dich da war?
6. **Stell dir mal vor, du bist jetzt Student/Studentin**, was würdest du mit deinem Mentee/Kind machen? Was würdest du ihm zeigen/erzählen?
7. Nächstes Jahr bekommen andere Kinder eine/n StudentIn: Was möchtest du denen sagen, **warum sie mitmachen sollen?**
8. Was möchtest du ihnen sagen, **was man davon hat**, wenn man eine/n Studentin hat?
9. Was wirst du mit deinem **Tagebuch** machen?
 - 9.1. Wirst du es aufheben, es wegschmeißen, dein Zimmer damit tapezieren oder es als Klopapier benutzen?
10. Hast du mit deiner Familie/deinen Freunden/Geschwistern über die **Treffen** mit der Studentin/dem Studenten geredet
 - 10.1. Was genau hast du ihnen erzählt?
11. Hast du Lust darauf, dass so ein **Tag wie heute nochmal** gemacht wird, wo man alle wiedersehen kann?
 - 11.1. Möchtest du deine/n StudentIn nochmal sehen?
 - 11.2. Was möchtest du mit ihm/ihr machen?
 - 11.3. Wann möchtest du ihn wiedersehen?

Lebenslauf

Personalien	<i>Name</i> Verena Kofler <i>Geburtsdatum und -ort</i> 12.07.1987, Meran, Südtirol (I) <i>Staatsbürgerschaft</i> Italien <i>Kontakt</i> koflerverena@hotmail.com
Ausbildung	Seit WS 2015 Masterstudium Sozialwirtschaft und Soziale Arbeit, FH Campus Wien. Voraussichtlicher Abschluss: Master of Arts in Social Sciences SS 2017 SS 2014–WS 2015 Masterstudium Sozial- und Humanökologie, IFF Wien WS 2010–01/2014 Bachelorstudium Kultur- und Sozialanthropologie, Universität Wien. Abschluss: Bachelor of Arts in Social Sciences 2008–2010 Abendschule der Handelsoberschule „Franz Kafka“ Meran (I). Abschluss: Matura
Berufserfahrung	02/2015–heute Teamleitung und Teamkoordination Persönliche Assistenz, Fachbereich der Assistenz24 gemeinnützige GmbH, Wien 05/2012–06/2015 Persönliche Assistentin für Menschen mit Behinderung, Assistenz24 gemeinnützige GmbH, Wien
Freiwilligenarbeit	11/2014–06/2015 Nightingale Wien: Mentorin für SchülerInnen mit Migrationshintergrund
Konferenzteilnahmen	04/2017 11. FH-Forschungsforum 2017. Research-Innovation-Value: <i>Die „Stärkung der Bildungsmotivation“ als SMARTes Wirkungsziel im Mentoring-Programm The Nightingale, Wien.</i>
