

Weiterbildung an der Universität –

Evaluation des Projekts PROFiL-TT am
Institut für Psycholinguistik der
Ludwig-Maximilians-Universität München

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Philosophie an der Ludwig-Maximilians-Universität
München

Fakultät für Sprach- und Literaturwissenschaften

Dekan: Prof. Dr. Andreas Höfele

Vorgelegt von
Manfred Paul Kostolnik

aus
München

Gutachter: 1. Prof. Dr. Gerd Kegel (LMU München)

2. Prof. Dr. Dieter Frey (LMU München)

Tag der mündlichen Prüfung: 04.Juli 2007

Inhaltsverzeichnis

<i>Inhaltsverzeichnis</i>	5
<i>Danksagung</i>	10
<i>Verzeichnis der Abbildungen, Diagramme und Tabellen</i>	11
<i>Vorwort</i>	15
1 Weiterbildung	16
1.1 Aufgabenfelder der Erwachsenen- und Weiterbildung nach Kraft	17
1.2 Lebenslanges Lernen	18
1.3 Berufliche Weiterbildung	18
1.4 Allgemeine Weiterbildung	19
1.5 Politische Weiterbildung	20
1.6 Wissenschaftliche Weiterbildung	21
1.6.1 Expansion wissenschaftlicher Weiterbildung	22
1.6.2 Akkreditierungskriterien und -prozesse	22
1.7 Professionalisierung und Qualitätssicherung in der beruflichen Weiterbildung	23
1.7.1 Professionalisierung in der Weiterbildung	23
1.7.2 Qualität in der Weiterbildung	25
1.8 Beurteilung guter Lehre	26
1.8.1 Instrumente der Datenerhebung	26
1.8.2 Befragungsformen	27
1.8.3 Fragebogen – Eigenschaften und Vergleich mit anderen Verfahren	27
1.8.4 Fragebogenkonstruktion	28
1.8.5 Alternative Datenerhebungsmethoden in Anlehnung an Rindermann	29
1.8.6 Itemgenerierungsverfahren	30
1.8.7 Frageformulierung	31
1.8.8 Frage mit Einschätzungsskalen	31
1.9 Train-The-Trainer	32
1.9.1 Erwartungen an Weiterbildner	32
1.9.2 Hintergrund	32
1.9.3 Lernziele	33
1.9.4 Modularität	33
1.9.5 Zielgruppe	33
1.9.6 Aktuelle Tendenzen im Bereich der professionellen Weiterbildung	34

2	Evaluation	36
2.1	Sozialwissenschaftliche Bedeutung von Evaluation	36
2.2	Zielsetzungen und drei klassische Unterscheidungen von Evaluation	37
2.3	Merkmale und Kennzeichen von Evaluation	38
2.4	Nutzenerwartung der Auftraggeber einer Evaluation nach Wottawa	38
2.5	Typische Untersuchungsdesigns / Evaluationsdesigns	41
2.5.1	Experimente / Feldexperimente	41
2.5.2	Quasi-Experimente	43
2.5.3	Ex-Post-Facto-Designs	43
2.5.4	Gutachtenmodell	45
2.5.5	Untersuchungsdesigns im Überblick	46
2.6	Störfaktoren	47
2.7	Deskriptive Statistik	49
2.7.1	Population und Grundgesamtheit	49
2.7.2	Quantile und Median	50
2.7.3	Quartilabstand	51
2.7.4	Interpretation ordinal- und intervallskalierter Daten	51
2.7.5	Statistische Mortalität	52
2.8	Grundsätzliche Evaluationsvorgehensweisen	52
2.8.1	Formative Evaluation nach Scriven	52
2.8.2	Summative Evaluation nach Scriven	53
2.8.3	Vergleich wesentlicher Merkmale von formativer und summativer Evaluation	54
2.8.4	Interne Evaluation – Selbstevaluation	56
2.8.5	Externe Evaluation – Fremdevaluation	57
2.8.6	Evaluationsvorgehensweise bei PROFIL-TT	57
2.9	Evaluationsmodelle	58
2.9.1	Übersicht	58
2.9.2	Effizienzkriterienorientiert Evaluation (Rossi, Freeman, Lipsey)	59
2.9.3	Versuchsplanorientierte Evaluation (Hager, Patry, Brezing)	59
2.9.4	Programmziel- / lernzielgesteuerte Evaluation (objective-based study)	59
2.9.5	Theoriegeleitete Evaluation (Chen)	60
2.9.6	Gegnerschaftsgesteuerte Evaluation (adversary Evaluation)	60
2.9.7	Entscheidungsgesteuerte Evaluation (Stufflebeam)	60
2.9.8	Beteiligteninteressengesteuerte Evaluation	60
2.9.9	Qualitätsentwickelnde Evaluation	61
2.9.10	Nutzenorientierte Evaluation (Utilization-Focused Evaluation; Patton)	62
2.9.11	CIPP-Modell (Stufflebeam)	62

2.9.12	Zielorientierte Evaluation	63
2.9.13	Goal-free evaluation (Scriven)	63
2.9.14	Partizipative Evaluation	64
2.9.15	Cluster-Evaluation (Haubrich)	64
2.9.16	Deskriptive Evaluation – Explorative Evaluation – Hypothesenprüfende Evaluation	65
2.9.17	Askriptive Evaluation	65
2.9.18	Black-box Evaluation	66
2.9.19	Experten-orientierte Evaluation	66
2.9.20	Analytic / analytical Evaluation (Scriven)	66
2.9.21	Global Evaluation (Scriven)	67
2.9.22	Alternative Evaluationsmodelle	67
2.9.23	Modelle in der Evaluation PROFiL-TT	68
2.10	Evaluationsstandards	68
2.10.1	Übersicht	68
2.10.2	Hintergrund der Evaluationsstandards	69
2.10.3	Nützlichkeitsstandards	70
2.10.4	Durchführbarkeitstandards	74
2.10.5	Korrektheitsstandards	75
2.10.6	Genauigkeitsstandards	78
2.10.7	Zusammenfassung der Standards	84
2.10.8	Durchführungsentscheidung einer Evaluation	84
2.10.9	Anwendung der Evaluationsstandards im Projekt PROFiL-TT	85
2.11	Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE) nach Rindermann (1994)	85
2.11.1	Übersicht	85
2.11.2	Dimensionen und Grundstruktur	86
2.11.2.1	Dozent	87
2.11.2.2	Studierende	89
2.11.2.3	Rahmenbedingungen	91
2.11.2.4	Lehrerfolg	92
2.11.3	HILVE-II: Dimensionen und Items	93
2.11.4	Anwendung der HILVE-II Dimensionen und Items auf die Evaluation von PROFiL-TT	98
3	Projekt PROFiL-TT	99
3.1	Projektbeschreibung	99
3.1.1	Projekthintergrund	99
3.1.2	Modularer Aufbau	100
3.1.3	Projektziele	101
3.1.4	Modulziele – inhaltliche Schwerpunkte	101
3.1.5	Zielgruppe / Teilnahmevoraussetzungen	105

3.1.6	Mitarbeiter	105
3.1.7	Träger / Förderer	106
3.2	Evaluation PROFiL-TT	107
3.2.1	Inhaltliche Vorgehensweise bei der Evaluation	108
3.2.2	Technische Vorgehensweise – Datenerhebung und -auswertung	109
3.2.3	Zeitmodell – Untersuchungsplan	114
3.2.4	Ergebnisse der Evaluation von PROFiL-TT	116
3.2.4.1	Eigenschaften der Population	117
3.2.4.2	Publizitätsmaßnahmen	118
3.2.4.3	Struktur	118
3.2.4.4	Auseinandersetzung	120
3.2.4.5	Redundanz	121
3.2.4.6	Anforderungen	126
3.2.4.6.1	Stoffschwere	127
3.2.4.6.2	Stoffumfang	130
3.2.4.6.3	Kurstempo	133
3.2.4.6.4	Anforderungen	137
3.2.4.7	Lernen – quantitativ	141
3.2.4.7.1	Lern- Quantität	141
3.2.4.7.2	Subjektiver Wissenszuwachs	145
3.2.4.8	Verarbeitung	149
3.2.4.9	Betreuung	151
3.2.4.9.1	Feedback	152
3.2.4.9.2	Betreuung außerhalb der Veranstaltung	152
3.2.4.10	Beteiligung	153
3.2.4.11	Kommunikative Unterrichtsformen	154
3.2.4.11.1	Diskussionen – quantitativ	154
3.2.4.11.2	Einsatz kommunikativer Unterrichtsformen	155
3.2.4.12	Dozentenbewertung	155
3.2.4.13	Allgemeineinschätzung	160
3.2.5	Qualitative Auswertung der Verbatims	161
3.2.5.1	Anmerkungen aus der Vorab- Untersuchung	161
3.2.5.2	Zusammenfassung positiver Anmerkungen	162
3.2.5.3	Zusammenfassung konstruktiver Anmerkungen	163
3.2.5.4	Kommentare und Anmerkungen zur Erhebung	164
3.2.6	Resümee	164
3.2.6.1	Organisation – Gesamtveranstaltung	165
3.2.6.2	Modulübergreifende Themen	166
3.2.6.3	Modulbezogene Themen	167
3.2.7	Fazit	173

<i>Literatur- und Quellenverzeichnis</i>	174
<i>Anhang</i>	181
Anhang I: Program Evaluation Standards des Joint Committee on Standards for Educational Evaluation	181
Anhang II: Heidelberger Inventar zur Lehrveranstaltungsevaluation HILVE-II Dimensionen und Items	185
Anhang III: Datenerhebungsinstrumente – Screenshots - Beispiele	188
Anhang IV: Verbatims	190
<i>Lebenslauf</i>	193
<i>Eidesstattliche Erklärung</i>	<i>Fehler! Textmarke nicht definiert.</i>

Danksagung

Herzlichen Dank an all diejenigen, die mich bei der Erstellung dieser Arbeit unterstützt haben.

Besonderer Dank gilt Prof. Dr. Gerd Kegel für die Inspiration genauso, wie für die Unterstützung und die konstruktiven Ratschläge während des gesamten Prozesses.

Herzlichen Dank an Prof. Dr. Dieter Frey vom Lehrstuhl für Sozialpsychologie der LMU München für seine spontane Bereitschaft, für diese Arbeit ein Gutachten zu erstellen.

Vielen Dank an Dr. Andreas Hendrich, der mir sowohl in der Rolle des Projektadministrators und Dozent von PROFiL-TT, wie auch als Angestellter des Instituts für Psycholinguistik stets ein offenes Ohr für Fragen und Unklarheiten bot.

Danke auch an die Mitarbeiter des Projekts PROFiL-TT, namentlich genannt seien hier Prof. Dr. Jochen Gerstenmaier vom Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie der LMU München, Dr. Marija Kulis und Dr. Susanne Günther, Cornelia Gericke, Imme Schönfeld und Stefanie Vits.

An dieser Stelle dürfen auch nicht die Teilnehmer von PROFiL-TT und der Evaluation vergessen werden, auf deren Aussagen, Mitarbeit und Engagement die Ergebnisse dieser Arbeit letztendlich beruhen.

Vielen Dank an Rupert Heigl, der mir vor allem in der Anfangsphase bei der Konzeption und Umsetzung des Online- Erhebungstools beratend und unterstützend zur Seite stand.

Das Engagement der Mitarbeiter des Leibniz- Rechenzentrums trug darüber hinaus zu einem reibungslosen Ablauf der Datenerhebung wesentlich mit bei.

Danke auch an meinen Arbeitgeber und die Kollegen von der Solution 24 GmbH & Co. KG für die menschliche Unterstützung, den Rückhalt und die Freiräume, die diese Arbeit neben einer Vollzeitbeschäftigung mit ermöglicht haben.

Besonders herzlicher Dank gilt meiner Familie und meiner Freundin Jenny.

Herzlichen Dank!

München im März 2007

Verzeichnis der Abbildungen, Diagramme und Tabellen

Abbildungen	Seite
<i>Abbildung 1: Multifaktorielles Modell der Lehrveranstaltungsqualität</i>	87
<i>Abbildung 2: Auszug aus Datenstrukturmodell</i>	112
<i>Abbildung 3: Auszug aus der Rohdatenübersicht in Excel mit Median / Quartilabweichung</i>	112
<i>Abbildung 4: Auszug aus den kumulierten Daten / Relativ- Werten in Excel</i>	113
<i>Abbildung 5: Auszug Instrument Prospektive Erhebung</i>	188
<i>Abbildung 6: Auszug Instrument Post- Erhebung</i>	189
<i>Abbildung 7: Auszug Instrument Retrospektive Erhebung</i>	189
Diagramme	Seite
<i>Diagramm 1: zeitliche Ablaufplanung der Evaluation PROFiL-TT</i>	116
<i>Diagramm 2: Aktuelle Trainingshäufigkeit im Vergleich zur zukünftigen</i>	117
<i>Diagramm 3: Publizitätsmaßnahmen von PROFiL-TT</i>	118
<i>Diagramm 4: Untersuchungsergebnisse Evaluation PROFiL-TT Aufbau</i>	119
<i>Diagramm 5: Untersuchungsergebnisse Evaluation PROFiL-TT Organisation</i>	119
<i>Diagramm 6: Untersuchungsergebnisse Evaluation PROFiL-TT Veranschaulichung</i>	120
<i>Diagramm 7: Untersuchungsergebnisse Evaluation PROFiL-TT Theorie-Praxis-Bezug</i>	121
<i>Diagramm 8: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Rhetorik</i>	122
<i>Diagramm 9: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Stimmbildung / Sprecherziehung</i>	123
<i>Diagramm 10: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Einsatz von Präsentationstechnik / Kursunterlagen</i>	123
<i>Diagramm 11: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul kohärenter Sprachgebrauch</i>	124
<i>Diagramm 12: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Lehrkompetenz</i>	125
<i>Diagramm 13a: Untersuchungsergebnisse Evaluation PROFiL-TT inhaltliche Überschneidungen vorher</i>	125

<i>Diagramm 13b: Untersuchungsergebnisse Evaluation PROFiL-TT inhaltliche Überschneidungen nachher / retrospektiv</i>	126
<i>Diagramm 14: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Rhetorik</i>	127
<i>Diagramm 15: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Stimmbildung und Sprecherziehung</i>	127
<i>Diagramm 16: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	128
<i>Diagramm 17: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul kohärenter Sprachgebrauch</i>	129
<i>Diagramm 18: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Lehrkompetenz</i>	129
<i>Diagramm 19: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Rhetorik</i>	130
<i>Diagramm 20: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Stimmbildung und Sprecherziehung</i>	131
<i>Diagramm 21: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Einsatz von Präsentationstechnik / Gestaltung von Kursunterlagen</i>	131
<i>Diagramm 22: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul kohärenter Sprachgebrauch</i>	132
<i>Diagramm 23: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Lehrkompetenz</i>	133
<i>Diagramm 24: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Rhetorik</i>	133
<i>Diagramm 25: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Stimmbildung und Sprecherziehung</i>	134
<i>Diagramm 26: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Einsatz von Präsentationstechnik / Gestaltung von Kursunterlagen</i>	135
<i>Diagramm 27: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul kohärenter Sprachgebrauch</i>	136
<i>Diagramm 28: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Lehrkompetenz</i>	136
<i>Diagramm 29: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul Rhetorik</i>	137
<i>Diagramm 30: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul Stimmbildung und Sprecherziehung</i>	138
<i>Diagramm 31: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul Einsatz</i>	139

<i>von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	
<i>Diagramm 32: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul kohärenter Sprachgebrauch</i>	140
<i>Diagramm 33: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul Lehrkompetenz</i>	140
<i>Diagramm 34: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Rhetorik</i>	141
<i>Diagramm 35: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Stimmbildung und Sprecherziehung</i>	142
<i>Diagramm 36: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	143
<i>Diagramm 37: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul kohärenter Sprachgebrauch</i>	144
<i>Diagramm 38: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Lehrkompetenz</i>	144
<i>Diagramm 39: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Rhetorik</i>	145
<i>Diagramm 40: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Stimmbildung und Sprecherziehung</i>	146
<i>Diagramm 41: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	147
<i>Diagramm 42: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul kohärenter Sprachgebrauch</i>	148
<i>Diagramm 43: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Lehrkompetenz</i>	149
<i>Diagramm 44: Untersuchungsergebnisse Evaluation PROFiL-TT aktives Mitdenken</i>	150
<i>Diagramm 45: Untersuchungsergebnisse Evaluation PROFiL-TT kritische Auseinandersetzung</i>	151
<i>Diagramm 46: Untersuchungsergebnisse Evaluation PROFiL-TT Feedback</i>	152
<i>Diagramm 47: Untersuchungsergebnisse Evaluation PROFiL-TT Betreuung außerhalb der Veranstaltung</i>	152
<i>Diagramm 48: Untersuchungsergebnisse Evaluation PROFiL-TT Beteiligung</i>	153
<i>Diagramm 49: Untersuchungsergebnisse Evaluation PROFiL-TT Diskussionen - quantitativ</i>	154
<i>Diagramm 50: Untersuchungsergebnisse Evaluation PROFiL-TT kommunikative Unterrichtsformen</i>	155

<i>Diagramm 51: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /- beurteilung Modul Rhetorik</i>	156
<i>Diagramm 52: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /- beurteilung Modul Stimmbildung und Sprecherziehung</i>	157
<i>Diagramm 53: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /- beurteilung Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	158
<i>Diagramm 54: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /- beurteilung Modul kohärenter Sprachgebrauch</i>	159
<i>Diagramm 55: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /- beurteilung Modul Lehrkompetenz</i>	160
<i>Diagramm 56: Untersuchungsergebnisse Evaluation PROFiL-TT Allgemeineinschätzung</i>	161
Tabellen	Seite
Tabelle 1: Übersicht typischer Untersuchungsdesigns (vgl. Stockmann 2006, 229)	47
<i>Tabelle 2: Überblick Median und Quartilabstand Aufbau, Organisation und Betreuung außerhalb der Veranstaltung</i>	165
<i>Tabelle 3a: Überblick Median und Quartilabstand modulübergreifende Themen</i>	166
<i>Tabelle 3b: Überblick Median und Quartilabstand modulübergreifende Themen</i>	167
<i>Tabelle 4a: Überblick Median und Quartilabstand Modul Rhetorik</i>	167
<i>Tabelle 4b: Überblick Median und Quartilabstand Modul Rhetorik</i>	168
<i>Tabelle 5a: Überblick Median und Quartilabstand Modul Stimmbildung und Sprecherziehung</i>	168
<i>Tabelle 5b: Überblick Median und Quartilabstand Modul Stimmbildung und Sprecherziehung</i>	169
<i>Tabelle 6a: Überblick Median und Quartilabstand Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	169
<i>Tabelle 6b: Überblick Median und Quartilabstand Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen</i>	170
<i>Tabelle 7a: Überblick Median und Quartilabstand Modul kohärenter Sprachgebrauch</i>	170
<i>Tabelle 7b: Überblick Median und Quartilabstand Modul kohärenter Sprachgebrauch</i>	171
<i>Tabelle 8a: Überblick Median und Quartilabstand Modul Lehrkompetenz</i>	171
<i>Tabelle 8b: Überblick Median und Quartilabstand Modul Lehrkompetenz</i>	172

Vorwort

Universitäre Weiterbildung – Evaluation – Train-the-Trainer. Eine in sich geschlossene Abfolge von Prozessen und Teilprozessen, die beliebig weitergeführt und zueinander in rekursive Abhängigkeit gesetzt werden kann. Die einzelnen Elemente dieser Kette für sich betrachtet sind dabei hoch komplex und vielschichtig, darüber hinaus äußerst interessant in der Praxis näher zu untersuchen.

Bezogen auf meine persönliche Situation vor dem Hintergrund des Themas dieser Arbeit, befinde ich mich im fortgeschrittenen Stadium dieser Entelechie. Diese Arbeit behandelt thematisch den Prozess, den ich gerade selbst durchlaufe – universitäre Weiterbildung, Evaluation.

Nach den Prämissen vieler Weiterbildungsveranstaltungen zeigt der Aufbau dieser Arbeit eine generelle Zweiteilung auf. Einerseits beschäftigt sie sich mit theoretischen Modellen zur Weiterbildung und Evaluation, Standards und Richtlinien, die bei Evaluationen berücksichtigt werden sollten, Lehrevaluationsinventaren und Möglichkeiten, die Qualität von Lehre zu beurteilen, auf der anderen Seite steht die praktische Implementierung und Bezugnahme auf eine reale Weiterbildungsveranstaltung im Mittelpunkt. Ein Praxisbezug ist dabei – meiner Meinung nach – wesentlich und äußerst wichtig, um die komplexe und abstrakte Thematik darstellen und vermitteln zu können. Die theoretischen Inputs unter Kapitel 1 und 2 sollen dabei zur praktischen Anwendung (Kapitel 3) hinführen und die Vorgehensweise bei der Evaluation, sowie die Umgebungsvariablen und Randbedingungen erklären.

Generell befindet man sich im Rahmen einer Evaluation fortwährend auf einem schmalen Grat zwischen empirisch- methodischer Wissenschaftlichkeit und pragmatischer Anwendungsorientierung – diese beiden Ansätze können unter Umständen stark differieren. Hier gilt es einen „goldenen Weg“ zu finden, um sowohl der empirischen Wissenschaftlichkeit, als auch der Implementierung und Anwendung dieser Methodik gerecht werden zu können.

Ein ausschlaggebender Punkt bei der Durchführung des empirischen Teils dieser Arbeit, lag in der Zielsetzung, die theoretischen Methoden und Ansätze auch in der Realität, somit im Untersuchungsfeld, anwenden zu können. Meiner Meinung nach sollte grundsätzlich eine klare Nutzenorientierung sowie ein Realitätsbezug des Untersuchungsgegenstands – kurz-, mittel- oder langfristig – absehbar sein, wobei ein klar erkennbarer, konsistent durchgeführter, empirischer Hintergrund keineswegs fehlen darf. *Evaluation* sollte dabei dennoch nicht auf die „Anwendung empirischer Forschungsmethoden mit primär aus der Praxis heraus motivierten Fragestellungen, zur Lösung praktischer Problemstellungen bzw. mit nicht-wissenschaftlichen Zielen“ reduziert, sondern aufgrund der „sozialen Funktionen [...], ihrer spezifischen Akteurskonstellationen und des Auftragskontexts“¹ als eigene Disziplin gesehen und auch als solche verstanden werden.

¹ Hense J (2007). *Online-Wörterbuch Evaluation*. In: evolution.de - Evaluation und Qualitätssicherung im Bildungswesen. Abrufbar unter <http://www.evaluation.de/glossary> (Abgerufen: 16. Januar 2007)

1 Weiterbildung

Um die komplexe Thematik „Weiterbildung“² konkretisieren zu können, wurden von verschiedenen Seiten Definitions- und Begriffsklärungsversuche unternommen.

Das Bundesministerium für Bildung und Forschung (BMBF) bezeichnet Weiterbildung generell als „Fortsetzung jeder Art des Lernens nach Abschluss der Bildungsphase in der Jugend“.³

Der Deutsche Bildungsrat definiert 1970 Weiterbildung folgendermaßen:

„Weiterbildung wird hier als Fortsetzung oder Wiederaufnahme organisierten Lernens nach Abschluß einer unterschiedlich ausgedehnten ersten Bildungsphase bestimmt. Das Ende der ersten Bildungsphase und damit der Beginn möglicher Weiterbildung ist in der Regel durch den Eintritt in die volle Erwerbstätigkeit gekennzeichnet; [...] Die Grenzen zwischen erster Bildungsphase und Weiterbildung bleiben fließend, ein Mindestzeitraum der Erwerbstätigkeit als Voraussetzung für Weiterbildung läßt sich nicht angeben. Das kurzfristige Anlernen oder Einarbeitung am Arbeitsplatz gehört nicht in den Rahmen der Weiterbildung.“⁴

Weiterbildung ist demnach eine weitere Stufe im Bildungsprozess eines Individuums, die einerseits auf privaten Initiativen, andererseits häufig auch auf beruflichen Hintergründen basiert. Generell kann eine starke Tendenz zu Handlungsorientierung und Praxisbezug festgestellt werden, die häufig auf berufsberufsbezogene Erfahrungen zurückgreift, allerdings auch auf alltäglich- pragmatische Erwartungen Bezug nimmt.

Eine genaue Festlegung des Zeitraums, in dem Weiterbildung stattfindet, kann nicht getroffen werden. Aus der Definition geht allerdings hervor, dass Weiterbildung eine bereits durchlaufene aktive Teilnahme am Berufsleben voraussetzt. Trotz einer gesetzlichen Verankerung des Rechts auf Weiterbildung in der EU-Verfassung, nach dem jeder „Person [...] das Recht auf Bildung sowie auf Zugang zur beruflichen Ausbildung und Weiterbildung“⁵ zusteht, ist es verwunderlich, dass Weiterbildung in Deutschland – statistisch betrachtet – einem „Flickenteppich“ gleicht.⁶

Um die generell sehr vielschichtige Thematik zu konkretisieren, soll im folgenden Abschnitt Weiterbildung nach unterschiedlichen Aspekten gegliedert und diese Aspekte näher erläutert werden.

1. Lebenslanges Lernen
2. Berufliche Weiterbildung

² Erwachsenen- und Weiterbildungen werden vielfach synonym verwendet. (Erwachsenen- und Weiterbildungen werden vielfach synonym verwendet. (vgl. Tippelt 1999, 11)

³ Vgl. BMBF. Artikel *Weiterbildung*

⁴ Deutscher Bildungsrat: Empfehlungen der Bildungskommission: Strukturplan für das Bildungswesen, 13. Februar 1970, 197 ff.

⁵ EU-Verfassung, Artikel II-74, Ziffer 1

⁶ Vgl. Pehl K (2001)

3. Allgemeine Weiterbildung
4. Politische Weiterbildung
5. Wissenschaftliche Weiterbildung

Neben diesen generellen Formen der Weiterbildung, werden häufig auch Formen arbeitsintegrierten Lernens, Fernunterricht, Computer Based Training, Training on the job, Training near the job und Training off the job, sowie selbstorganisiertes Lernen zur Weiterbildung gezählt.⁷

1.1 Aufgabenfelder der Erwachsenen- und Weiterbildung nach Kraft⁸

Die Aufgabenfelder der Erwachsenen- und Weiterbildung unterteilt Kraft (2006) historisch bedingt in vier Gruppen.

1. *Persönlichkeitsbildung*

Im Rahmen der Persönlichkeitsbildung besteht die Aufgabe der Weiterbildung darin, die „Entfaltung individueller Potentiale zur Gestaltung eines erfüllten Lebens und zum Ausbau sozialer Kompetenzen für soziales und gemeinschaftliches Zusammenleben“⁹ zu schaffen. Der Hintergrund für die Persönlichkeitsbildung ist eher allgemeiner Natur und kann sowohl auf berufliche, wie auch auf private Interessen zurückgeführt werden.

2. *Qualifikation*

Das Aufgabenfeld *Qualifikation* orientiert sich am „Erwerb oder [der] Ergänzung berufsorientierter Fähigkeiten und Kompetenzen zur individuellen Beschäftigungsfähigkeit sowie zur gesellschaftlichen Wettbewerbsfähigkeit“¹⁰ und zielt damit vorwiegend im ökonomischen Sinne auf die Aufrechterhaltung bzw. Steigerung des individuellen „Marktwerts“. Weiterbildungen in diesem Rahmen fokussieren in erster Linie eine Höherqualifizierung von Individuen, meist verbunden mit einem zunehmenden Anforderungsfeld.

3. *Kompensation*

Der weiterbildnerische Hintergrund in der *Kompensation* besteht im Nachholen von, in früheren Bildungsphasen – z.B. der Erst- oder Schulausbildung – nicht erworbenem Wissen, wodurch ein höheres Wissensniveau erreicht werden kann. Kraft fordert, dass „Erwachsene [...] fehlendes Wissen und mangelnde Schulausbildung durch die Teilnahme an Weiterbildung ausgleichen können“.¹¹

4. *Demokratisierung*

⁷ Vgl. BMBF. Artikel *Neue Formen der betrieblichen Weiterbildung*

⁸ Vgl. Kraft S (2006). Aufgaben und Tätigkeiten von Weiterbildungler/inne/n – Herausforderungen und Perspektiven einer weiteren Professionalisierung in der Weiterbildung

⁹ Ibidem

¹⁰ Ibidem

¹¹ Ibidem

Das letzte, hier erwähnte Aufgabenfeld von Erwachsenen- / Weiterbildung soll Individuen die Möglichkeit „zur demokratischen Teilhabe an gesellschaftlichen Aufgaben“¹² bieten; eine Integration in die Gesellschaft ist häufig das Resultat dieser Demokratisierung.

1.2 Lebenslanges Lernen

Wissen und Kenntnisse aus der Berufsausbildung sowie die gesammelte Berufserfahrung reichen häufig nicht aus, um während eines meist mehrere Jahrzehnte langen Arbeitslebens zurechtzukommen zu können. Aus diesem Grund ist es für Arbeitnehmer notwendig, sich während der Dauer ihres Berufslebens – teilweise auch darüber hinaus – durch die Aneignung von neuem, innovativem Wissen weiterzubilden und somit ihre Qualifikation zu erhöhen. Dieser Lernprozess zeichnet sich nicht nur durch den Besuch von Weiterbildungsveranstaltungen, sondern auch durch im Berufsleben gesammelte Erfahrungen und die Orientierung an sich verändernden Arbeitsumgebungen und daraus resultierenden höheren Ansprüchen an den Arbeitnehmer als Individuum aus.¹³

Auf politischer Ebene bildet lebenslanges Lernen auch einen wesentlichen Faktor, der zur Aufrechterhaltung, Verstärkung und Fundierung des individuellen Soziallebens beiträgt:

„Lebenslanges Lernen hilft, den Zusammenhalt in der Gesellschaft zu stärken und Ausgrenzung soweit wie möglich zu vermeiden. Im Rahmen einer Gesamtstrategie soll das Ziel verfolgt werden, die Bildungsteilhabe zu erhöhen, allen Menschen mehr Chancen zur persönlichen, ihren Begabungen entsprechenden gesellschaftlichen und beruflichen Entwicklung zu ermöglichen und den Standort Europa mitzugestalten.“¹⁴

Kritik am Ansatz des lebenslangen Lernens richtet sich in erster Linie auf den starken ökonomischen Hintergrund, nach dem die Wissens- und Bildungssteuerung von wirtschaftlichen und arbeitsmarktbezogenen Interessen abhängig ist. Wer sich diesem Trend nicht unterwirft – so die Kritik weiter – ist quasi selbst für die eigene, gesellschaftliche und soziale Ausgrenzung verantwortlich.¹⁵

Eine Definition von lebenslangem Lernen ist aufgrund dieser Vielschichtigkeit nur schwer möglich.

1.3 Berufliche Weiterbildung

Berufliche Weiterbildung – häufig synonym verwendet mit Fortbildung – ist ein wesentlicher Bestandteil von *Weiterbildung* und wird als eine Form der Erwachsenenbildung definiert, die jegliche Art des beruflichen Fortschritts unterstützen, sichern oder vorantreiben soll.

¹² Ibidem

¹³ Vgl. BMBF. Artikel *Lebenslanges Lernen*

¹⁴ Ibidem

¹⁵ Vgl. Wikipedia. Artikel *Lebenslanges Lernen*

Im Berufsbildungsgesetz wird *berufliche Bildung* als eine Form der Erwachsenenbildung bezeichnet, die dazu dient, „die berufliche Handlungsfähigkeit zu erhalten und anzupassen oder zu erweitern und beruflich aufzusteigen.“¹⁶

Generell begründet sich berufliche Weiterbildung aus Anforderungsänderungen im beruflichen Alltag, die eine Weiterqualifizierung fordern. Solche Anforderungsänderungen können durch hohe Fluktuation oder schnellen technischen Fortschritt ausgelöst oder forciert werden. Im Rahmen des lebenslangen Lernens werden die Erhaltung, Anpassung und Erweiterung vorhandenen Wissens im Berufsalltag zur Notwendigkeit. Dabei ist für manche Berufsgruppen eine „Aktualisierung“ ihres Wissens erforderlich, um einerseits den Anschluss an technische Innovationen nicht zu verlieren und somit ihre Konkurrenzfähigkeit auf dem Arbeitsmarkt zu erhalten. Einflussfaktoren, wie z.B. Elternurlaub oder Arbeitslosigkeit können ebenso eine Weiterqualifizierung unumgänglich machen, um sich leichter wieder ins Arbeitsleben integrieren zu können. Von äußeren Einflussfaktoren abgesehen, kann berufliche Weiterbildung auch aus eigenem Antrieb – beispielsweise dem Streben einer höheren Position innerhalb Firmen- Hierarchien – im Rahmen von Zusatzqualifikationen initiiert werden.

Da berufliche Weiterbildung meist auf freiwilliger Basis geschieht, wird die Finanzierung in der Regel vom Betroffenen selbst übernommen.¹⁷ Ausnahmen bilden dabei Maßnahmen, die von der Bundesagentur für Arbeit oder anderen Träger finanziert werden.

1.4 Allgemeine Weiterbildung

Die Definition von Allgemeinbildung oder allgemeiner Weiterbildung – nach dem Bundesministerium für Forschung und Wissenschaft ein wesentlicher Bestandteil der Weiterbildung¹⁸ – wird in verschiedener Hinsicht diskutiert.

Eine grundlegende Definition von Allgemeinbildung bietet Olechowski (1997):

„Allgemeinbildung ist die Gewinnung von Grundkompetenzen in möglichst allen Bereichen des Lebens für die kritische Auseinandersetzung mit der gesamten physischen und geistigen Wirklichkeit des Lebens.“¹⁹

Die Festlegung, welche Aspekte *alle Bereiche des Lebens* einschließen ist wiederum stark diskutiert. In diesem Umfeld wird häufig auf den Begriff *Bildungskanon* Bezug genommen, der die Essenz der Bildung für eine Gesellschaft oder Kultur festzulegen versucht und somit einem steten Wandel und gesellschaftlichen Veränderungen unterworfen ist. Abgesehen von dieser fortschreitenden Entwicklung besteht eine generelle Uneinigkeit über die Inhalte, die der Bildungskanon umfassen sollte.

¹⁶ BGBl. I 931. §1 Berufsbildungsgesetz Abs. 4). Geänderte Fassung vom 23. März 2005

¹⁷ Unterschieden werden sollte hierbei zwischen *betrieblicher* und *beruflicher* Weiterbildung. Im Rahmen von betrieblichen Weiterbildungsmaßnahmen werden die Kosten häufig komplett oder zumindest teilweise von Unternehmen getragen (vgl. Wikipedia. Artikel *Berufliche Weiterbildung*).

¹⁸ Vgl. BMBF. Artikel *Weiterbildung*

¹⁹ Olechowski R (1997)

Unter den Aspekten der kritisch- konstruktivistischen Didaktik teilt Klafki Allgemeinbildung in drei wesentliche Aspekte ein.²⁰ Allgemeinbildung sollte demnach zunächst „für alle“²¹ bestimmt sein und auf einer Chancengleichheit beruhen, die nicht durch den gesellschaftlichen oder sozialen Status des Individuums beeinflusst wird. Ferner sollte Allgemeinbildung „allseitig“²² sein, indem sie methodisch unterschiedliche Lernaspekte und -ebenen abdeckt und für die Lernenden interessenbezogen gestaltet wird. Darüber hinaus sollte Allgemeinbildung „durch das Allgemeine“²³ bestimmt werden.

Vor dem Hintergrund der Informationsgesellschaft, in der der Zugang zu Informationen über Medien und Informationsquellen – das Internet spielt aktuell eine wesentliche Rolle – für eine breite Personengruppe zum täglichen Leben zählt, wird das Allgemeinwissen häufig der Fertigkeit gegenübergestellt, Informationen zu beschaffen, zu bewerten und unterschiedliche Informationen zu vergleichen.²⁴

1.5 Politische Weiterbildung

Wie die allgemeine Weiterbildung, ist auch die politische Weiterbildung häufig nicht berufsbezogen und zielt vorwiegend auf die Vermittlung von Toleranz und Kritikfähigkeit den demokratischen Grundsätzen zufolge.

Im Rahmen der Kultusministerkonferenz (1998) werden die Ziele und Perspektiven politischer Weiterbildung folgendermaßen definiert:

„Aus Sicht der Kultusministerkonferenz soll die politische Weiterbildung einen konstitutiven Beitrag zu einer demokratischen Staats- und Gesellschaftsordnung und zur Weiterentwicklung der demokratischen politischen Kultur leisten. Sie sei dem Ziel der mündigen Bürgerinnen und Bürger verpflichtet und eine am Grundgesetz und den Länderverfassungen sowie an internationalen Konventionen orientierte Aufgabe.“²⁵

Die generelle Zielsetzung von politischer Bildung in demokratischen Gesellschaften besteht demnach darin, Bürger zur sozialen Mündigkeit und Autonomie zu erziehen eine Variation dieser Zielsetzungen kann allerdings zwischen verschiedenen politischen Hintergründen – demokratisch, konservativ, liberal, sozial oder sozialistisch

²⁰ Diese Einteilung basiert auf dem Postulat von Johann Amos Comenius, der bereits im 17. Jahrhundert lerntheoretische Prinzipien (u.a. die allgemeine Schulpflicht, eine Verbindung von Handlung, Bild und Wort im Lernprozess, sowie eine Aufteilung der Schüler in unterschiedliche wissens- und altersbezogene Gruppen) formulierte, die aktuell noch gültig sind. Die Essenz von Comenius „Omnes, omnia, omnino“ („Man kann alle Menschen alle Dinge der Welt in grundlegender Weise lehren“; Didaktik Uni Jena. Artikel *Jan Amos Komensky*.) setzt einen Grundstein für eine moderne Sichtweise auf die Allgemeinbildung. (vgl. *ibidem*)

²¹ Didaktik Uni Jena. Artikel *Allgemeinbildung*.

²² *Ibidem*

²³ *Ibidem*; das „Allgemeine“ charakterisiert Klafki als „epochaltypische Schlüsselprobleme unserer kulturellen, gesellschaftlichen, politischen, individuellen Existenz“ (*ibidem*).

²⁴ Vgl. Wikipedia. Artikel *Allgemeinbildung*

²⁵ Kultusministerkonferenz (18. September 1998). Überlegungen zur politischen Weiterbildung.

– beobachtet werden. Die Zielsetzung tendiert dabei in die jeweilige ideologische Ausprägung der politischen Denkweise.²⁶

Die Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung (GPJE) formulierte 2004 drei Schwerpunkte für die politische Bildung. Demnach sollen politische Urteilsfähigkeit, politische Handlungsfähigkeit und methodische Fähigkeiten die Grundelemente für politische Bildung darstellen.²⁷

1.6 Wissenschaftliche Weiterbildung

Das Spektrum der wissenschaftlichen Weiterbildung – der letzten, hier erwähnten Kategorie der Weiterbildung – ist im universitären Rahmen relativ breit gefächert und umfasst Kurse bis hin zu kompletten Studiengängen. Als wesentlicher Bestandteil universitärer Aufgaben – neben klassischer Lehre und Forschung – sollen nachfrageorientierte Weiterbildungsangebote eine Verbindung zwischen Wissenschaftsorientierung und Praxisanforderungen knüpfen und somit eine Vernetzung des Wissens von Hochschullehrkräften mit dem Wissen von Fachkräften herstellen.²⁸

Im Rahmen der Professionalisierung und der Qualitätssicherung der Weiterbildung besteht jedoch noch Handlungsbedarf bei der „Schaffung von Transparenz des Angebots, Modularisierung, Zertifizierung und Akkreditierung“²⁹ sowie bei der „Zusammenarbeit der Hochschulen untereinander“.³⁰

Die Zugangsvoraussetzungen zur wissenschaftlichen Weiterbildung sind ebenso unterschiedlich, wie der Besuchsgrund oder die thematischen Inhalte. Bei manchen Veranstaltungen ist ein Hochschulabschluss nicht unbedingt Voraussetzung für den Besuch, bei anderen ist eine derartige Qualifikation obligatorisch.³¹

Die Angebotsform von weiterbildenden Qualifizierungen wird meist über lokale Prüfungsordnungen geregelt, wobei das Spektrum von Einzelveranstaltungen in überschaubarem zeitlichen Rahmen bis hin zu berufsbegleitenden Teilzeitstudiengängen reichen kann. Die Problematik – bereits im vorigen Absatz angesprochen – liegt generell in der Heterogenität der Abschlüsse, die selten modular konzipiert sind oder aufeinander aufbauen.³²

Bei einigen Varianten besteht die Möglichkeit einen Abschluss, ein Zertifikat oder eine Teilnahmebescheinigung zu erlangen, derartige Zeugnisse sind jedoch nicht grundsätzlich Bestandteil von wissenschaftlichen Weiterbildungsmaßnahmen.³³

²⁶ Vgl. Wikipedia. Artikel *Politische Bildung*

²⁷ Vgl. Weißeno G (2005)

²⁸ Vgl. BMBF. Artikel *Wissenschaftliche Weiterbildung*

²⁹ Ibidem

³⁰ Ibidem; siehe auch Kapitel 1.7

³¹ Vgl. ibidem

³² Vgl. Faulstich P (2004)

³³ Zur Anerkennung, Standardisierung und Vereinheitlichung von Abschlüssen und Zertifikaten, sowie den damit verbundenen Schwierigkeiten im Rahmen der inter- universitären Zusammenarbeit mehr im Kapitel 1.7;

Eine Orientierung an den Bedürfnissen von Berufstätigen kann sowohl an den zeitlichen Rahmenbedingungen, in denen wissenschaftliche Weiterbildung organisiert wird, als auch an den thematischen Inhalten, die meist auf beruflicher Erfahrung fußen und diese teilweise auch voraussetzen, erkannt werden.³⁴

Die Universitäten stehen bei der wissenschaftlichen Weiterbildung im Zwiespalt zwischen einem steigenden Bedarf ihrer Weiterbildung einerseits und immer stärker eingeschränkten finanziellen Mitteln andererseits. Die wissenschaftliche Weiterbildung wird vor diesem Hintergrund zunehmend als Mittel zur Ressourcengewinnung für den universitären Betrieb eingesetzt.³⁵

1.6.1 Expansion wissenschaftlicher Weiterbildung

Basierend auf dem wachsenden Bedarf an wissenschaftlicher Weiterbildung – abhängig von der immer geringer werdenden Halbwertszeit von Wissen und der steigenden Wissensfluktuation – nimmt die Anzahl der Anbieter wissenschaftlicher Weiterbildung stark zu, wobei die Rolle von Hochschulen hierbei untergeordnet ist. Ein interessanter Aspekt dabei ist allerdings, dass die universitären Dozenten³⁶ und Hochschullehrer generell sehr stark in wissenschaftliche Weiterbildung involviert sind – allerdings häufig nicht im Auftrag der Hochschulen. Die Universitäten versuchen diesen Umstand zu ändern, und die Ressource „Hochschuldozent“ im Rahmen wissenschaftlicher Weiterbildung an den Hochschulen einzugliedern.³⁷

1.6.2 Akkreditierungskriterien und -prozesse³⁸

Faulstich schlägt, bezogen auf Akkreditierungskriterien und -prozesse im Rahmen wissenschaftlicher Weiterbildung, ein dreistufiges Modell vor, das als wesentliche Kriterien Qualitätsmanagement, Zertifizierung und Akkreditierung beinhaltet.

der Europäische Sozialfonds zum Beispiel unterscheidet mehrere Zertifikatsformen: Trägerzertifikat (im Sinne einer Teilnahmebescheinigung), anerkannte Zertifikate (gesetzlich reguliert), sowie Kammerzertifikate (z.B. IHK; vgl. BMAS. Artikel *Der ESF in Deutschland*). Im Rahmen des Projekts PROFiL-TT hatten die Teilnehmer die Möglichkeit, ein solches Trägerzertifikat zu erhalten.

³⁴ Vgl. BMBF. Artikel *Wissenschaftliche Weiterbildung*

³⁵ Vgl. Faulstich P (2004); bezogen auf das vorliegende Projekt PROFiL-TT trifft dieser Aspekt durchaus zu, da mittels der Weiterqualifizierung universitärer Weiterbildner vor dem Hintergrund der Multiplikation universitären Wissens in der Wirtschaft, finanzielle Vorteile für die Universitäten und somit Ressourcengewinnung verbunden werden (siehe auch Kapitel 3).

³⁶ Die Begriffe *Dozent, Trainer, Hochschullehrer, Weiterbildner, Teilnehmer, Evaluator* etc. beziehen sich in dieser Arbeit sowohl auf Frauen, als auch auf Männer. Aus Gründen der einfacheren Leseart wird nur die maskuline Form verwendet.

³⁷ Vgl. Faulstich P (2004). 2

Zur Sicherung der Qualität sollten dabei eine grundlegende Vergleichbarkeit von Maßnahmen und Anbietern, sowie transparente Strukturen gewährleistet sein, die mittels einer Kombination von Selbst- und Fremdevaluation beurteilt werden – wobei eine Berücksichtigung und Umsetzung dieser Ergebnisse empfehlenswert wäre.

1.7 Professionalisierung und Qualitätssicherung in der beruflichen Weiterbildung

Im Rahmen der Erwachsenenbildung und Weiterbildung generell ist Professionalisierung und die Diskussion um Qualitätsentwicklung ein zentrales Thema. Der Verbraucherschutz am Bildungsmarkt fordert dabei Transparenz einerseits für die Nutzer von Weiterbildungsveranstaltungen, andererseits auch für die Anbieter von Qualifizierungsmaßnahmen zur Qualitätssicherung.³⁹

1.7.1 Professionalisierung in der Weiterbildung

Der Begriff Profession ist im Rahmen dieser Diskussion grundlegend.

Aus berufssoziologischer Sicht werden unter Profession Berufe zusammengefasst, die folgende Kriterien erfüllen müssen:⁴⁰

- hauptamtliche Erwerbstätigkeit
- Mandat
- Lizenz
- Existenz eines eindeutigen Berufsbildes
- eindeutige Regelung des Zugangs und der Qualifikationsanforderungen
- soziales Prestige als Experte
- Berufsverband als Interessenvertretung
- planmäßige wissenschaftliche Berufsausbildung
- Berufsethos

Einige dieser Aspekte – z.B. Mandat oder Lizenz – können auf die Weiterbildung nicht angewendet werden, weswegen Weiterbildung häufig als „unvollendete Profession“⁴¹ bezeichnet wird.

Der Begriff Professionalisierung soll im weiteren Verlauf der Arbeit als „Entwicklung erwachsenenpädagogischer Professionalität“ verwendet werden und somit das Qualitätsbewusstsein und die Optimierung von pädagogischem Handeln fokussieren.

³⁸ Vgl. *ibidem*; 4

³⁹ Vgl. BMBF. Artikel *Qualität in der Weiterbildung – Tests und mehr*

⁴⁰ Vgl. Arnold R (1984). 60 – 65

⁴¹ Vgl. Kraft S (2006). 5

„Professionalität heißt, [...] die Fähigkeit nutzen zu können, breit gelagerte, wissenschaftlich vertiefte und damit vielfältig abstrahierte Kenntnisse in konkreten Situationen angemessen anwenden zu können. Oder umgekehrt betrachtet: in eben diesen Situationen zu erkennen, welche Bestandteile aus dem Wissensfundus relevant sein können. Es geht also darum, im einzelnen Fall das allgemeine Problem zu entdecken. Es wollen immer wieder Relationen hergestellt sein zwischen gelernten Generalisierungen und eintretenden Situationen, zwischen einem umfangreichen Interpretationsrepertoire und dem unmittelbar Erfahrenen“⁴²

Diese Professionalität stützt sich nach Peters (1999) auf drei wesentliche Aspekte, die erwachsenenpädagogisch und didaktisch Handelnden – Voraussetzung für diese Gruppe ist eine angemessene Kompetenz und eine Klarheit bezogen auf den Tätigkeitsbereich und das Umfeld –, die Inhalte des Lernens – ausgeprägtes und fundiertes Wissen über bezogen auf den Lehr- und Lernstoff – und die lernenden Erwachsenen – bezogen auf die Empathie der Lehrenden den Lernenden gegenüber.⁴³

Es stellt sich die generelle Frage, wozu eine Professionalisierung in der Weiterbildung nützlich sein sollte. Im Rahmen dieser Fragestellung tritt wiederum der Begriff des lebenslangen Lernens auf, dessen Notwendigkeit und Sinnhaftigkeit bereits in zahlreichen Studien untersucht und nachgewiesen werden konnte.⁴⁴

Zentrale Problemfelder der Professionalisierung des Weiterbildungspersonals

Ein wesentliches Problemfeld im Rahmen der Weiterbildung besteht in der Heterogenität der Beschäftigungsgruppen, der Berufsbezeichnungen sowie der Qualifizierungswege – generell in der Standardisierung.

Das Bild der Beschäftigungsgruppen setzt sich aus haupt-, neben- und freiberuflichen Mitarbeitern zusammen, die sowohl an öffentlichen, wie auch an privaten Einrichtungen tätig sein können. Dennoch wird diese Vielzahl unterschiedlicher Gruppen häufig zu „die Lehrenden“ oder „die Weiterbildner“ generalisiert.⁴⁵ Der Grund für diese Verallgemeinerung liegt vorwiegend an den uneinheitlichen Berufsbezeichnungen, die für Weiterbildner aller Art existieren. Dadurch, dass für „Weiterbildner“ keine exakten Aufgaben- und Tätigkeitsbeschreibungen existieren (Weiterbildung ist keine Profession, s.o.) bleibt dieses Tätigkeitsfeld vorerst noch unscharf und kann nur schwerlich eingegrenzt werden. Diese Unschärfe setzt sich in der Vielfalt der Qualifizierungswege und un-

⁴² Tietgens H (1988). 28 – 75

⁴³ Eine ähnliche Aufteilung findet sich bei Rindermann (2001), der im Heidelberger Inventar zur Lehrevaluationsforschung zur Beurteilung guter Lehre unter anderem auf die Hauptdimensionen „Dozent“, „Studierende“ und „Rahmenbedingungen“ Bezug nimmt (vgl. Kapitel 2.10.2); Vgl. Peters R (1999). 92 – 102

⁴⁴ Vgl. u.a. Deutsche UNESCO- Kommission e.V. Artikel *Lebenslanges Lernen in Europa*; Europäische Union. Artikel *Schlüsselkompetenzen für lebenslanges Lernen*; BMBF. Artikel *Lebenslanges Lernen*; siehe auch Kapitel 1.2

⁴⁵ Vgl. Kraft S (2006). 25

terschiedlichen, voneinander unabhängigen Abschlüssen im Rahmen der Weiterbildung fort und sorgt dafür, dass die Transparenz für Nutzer und Betroffene im Weiterbildungsbereich stark eingeschränkt ist.⁴⁶

Vorschläge zur Professionalisierung der Weiterbildung nach Kraft⁴⁷

Basierend auf den Problemfeldern formuliert Kraft (2006) Vorschläge zur Professionalisierung der Weiterbildung, die Transparenz und Standardisierung in diesem Bereich schaffen sollen. Der Konsens dieser Vorschläge beruht auf einer Vereinheitlichung und Systematisierung der Weiterbildung, sowie einer Zusammenarbeit von Weiterbildungsinstitutionen, die, unter anderem durch eine „wechselseitige Anerkennung unterschiedlicher Teilmodule [einen] gemeinsamen Qualifizierungsrahmen“⁴⁸ schaffen sollen. Die Formulierung von Aufgaben- und Tätigkeitsbeschreibung von Weiterbildungnern, beziehungsweise festgelegte Kompetenzprofile sollen diesen Prozess unterstützen.⁴⁹

1.7.2 Qualität in der Weiterbildung

Eigenschaften „guter Lehre“⁵⁰

Hinter der Diskussion über Professionalität der Weiterbildung steht die Grundsatzfrage, welche Eigenschaften eine gute Lehrveranstaltung bzw. Weiterbildungsnahe haben sollte. Der Weg über die Qualität und damit die „Güte“ der Lehre liegt dabei nahe. Es stellt sich nun zunächst die Frage, was „gute Lehre“ eigentlich ist und welche Aspekte dieses Qualitätsmerkmal aufweisen sollte.

Rindermann beschreibt hierzu eine Vielzahl von Kriterien und Verfahren, mit deren Hilfe es möglich ist, den komplexen und vielschichtigen Begriff der „guten Lehre“ einzugrenzen und transparent zu machen.

Für die Definition werden verschiedene Möglichkeiten vorgeschlagen.

Der erste Ansatz geht von der Zielerreichung der Teilnehmer innerhalb eines Kurses aus. Es liegt dann „gute Lehre“ vor, wenn die Kursteilnehmer durch den Kurs ihren persönlichen Zielen – und auch den Zielen der Veranstaltung, z.B. Wissens- und Kompetenzerwerb – näher gebracht werden.

Eine deskriptive Möglichkeit, „gute Lehre“ zu bestimmen, besteht darin, dass Dozenten in der Rolle von „Experten“ die Güte von Veranstaltungen überprüfen. Dabei ist vorausgesetzt, dass diejenigen Dozenten aus eigener Erfahrung wissen, welche Kriterien für die gute Qualität einer Lehrveranstaltung ausschlaggebend sind.

⁴⁶ Vgl. *ibidem*; 26 f.

⁴⁷ Vgl. *ibidem*; 33

⁴⁸ *Ibidem*; 27

⁴⁹ An dieser Stelle kann ein Vergleich zu den in Österreich ausgearbeiteten „Funktionsprofile“ gezogen werden, in denen die unterschiedlichen Aufgabenbereiche der Weiterbildung thematisiert und eingegrenzt werden (vgl. BMUKK / BIFEB. Artikel *Funktionsprofile*)

⁵⁰ Vgl. Rindermann H (2001). 35 – 52

Eine dritte Möglichkeit, die Merkmale „guter Lehre“ zu bestimmen ist eine offene Befragung von Dozenten und den Besuchern von Veranstaltungen.

Merkmale und Kriterien „guter Lehre“⁵¹

Grundsätzlich sind die Merkmale und Kriterien, nach denen Lehrende und Studierende die Qualität und Güte von Lehre beurteilen sehr ähnlich. Sie basiert auf mehreren didaktischen, inhaltlichen und methodischen Aspekten.

So besticht „gute Lehre“ durch eine klar nachvollziehbare und stimmige Struktur im Aufbau und der Organisation der Lehrinhalte. Diese Struktur wird durch einen angemessenen Einsatz von Materialien – Handouts, Medien, u.a. – ergänzt und sollte sich stets auf einem Niveau bewegen, dass für alle Beteiligten angemessen – weder zu hoch, noch zu niedrig – ist. Die Rahmenbedingungen, unter denen der Wissensaustausch, bzw. die Informationsvermittlung stattfindet, sollten darüber hinaus für alle Involvierten akzeptabel und annehmbar sein. Neben einer didaktischen Methodenvielfalt und der sicheren Anwendung dieser Methoden unter verschiedenen Unterrichtsbedingungen, tragen Persönlichkeitsfaktoren (z.B. Engagement, Enthusiasmus, Freundlichkeit, Empathie und Offenheit den Interessen und Bedürfnissen der Veranstaltungsteilnehmer gegenüber) sowie der Einsatz von rhetorischen Hilfsmitteln wesentlich zu einer gelungenen Wissensvermittlung bei.

Diese Faktoren – sofern richtig eingesetzt und angewendet – führen zu Lernerfolg. Dabei ist „gute Lehre“ zwar ein komplexes, allerdings kein diffuses Konstrukt, und als solches auch erlernbar.

Nach Rindermann fließen Produktfaktoren wie der eigentliche Lernerfolg, Wissens- und Fähigkeitszuwachs sowie gute Studienabschlüsse nur bedingt in die Bestimmung von Lehrkompetenz mit ein.

1.8 Beurteilung guter Lehre

1.8.1 Instrumente der Datenerhebung⁵²

Die Entscheidung der Wahl eines standardisierten oder nicht- standardisierten Verfahrens gibt selten Aufschluss über die Präzision und Güte der Ergebnisse.

Wichtig bei der Wahl des Erhebungsinstrumentariums ist der Hintergrund des Einsatzes und die Begründung, aus welchem Grund eine bestimmte Methode eingesetzt wurde, was der Einsatz im Einzelfall bezwecken soll und warum keine andere Methode diese gewünschten Resultate erzielen kann.

⁵¹ Vgl. ibidem; 52 – 54

⁵² Vgl. Kromrey H (2006). 317 ff.

1.8.2 Befragungsformen⁵³

Bei den verschiedenen Befragungsformen kann eine Unterscheidung in zwei Dimensionen getroffen werden.

Eine Befragung generell kann mündlichen – Interview, Gruppeninterview oder Gruppendiskussion – oder schriftlichen Charakter haben. Bei schriftlichen Befragungen handelt es sich häufig um postalische oder via Email – auch als Link auf einen Fragebogen – zugeschickte Befragungsbögen. Um bei schriftlichen Befragungen die Anzahl der Rückläufer zu erhöhen, werden die Fragebögen oftmals anstelle einer postalisch Aussendung, persönlich abgegeben und auch wieder abholt. Bei einer schriftlichen Online- Befragung ist die Sicherung einer hohen Rücklaufquote durch Nachfassen – schicken von Erinnerungsemails – oder durch eine persönliche Ankündigung der Befragung möglich⁵⁴.

Die zweite Dimension, nach der verschiedene Befragungsformen differenziert werden können, ist der Grad der Standardisierung. Befragungen können nicht-, teil- oder voll- standardisiert sein. Der Grad der Standardisierung bezieht sich auf den Frage- und Antwortfreiraum, der dem Interviewer / Fragensteller sowie dem Interviewten / Befragten offen steht.

Bei voll- standardisierten Befragungen – in der Regel alle schriftlichen Befragungen – hat der Interviewte lediglich einen geringen Spielraum für seine Antworten, da hier die Reihenfolge der Fragen sowie die Fragen selbst fix vorgegeben sind.

Die teil- standardisierte Befragung bietet einen Mittelweg. Hier wird dem Befragten ein Fragengerüst vorgesetzt, das die Antwortmöglichkeiten zwar eingrenzt, allerdings auch noch Platz für freie Aussagen bietet. In der Regel verwenden teil- standardisierte Befragungen offene Fragen.

Die Vorgaben bei nicht- standardisierten Befragungen sind sehr frei. Dem Befragten werden lediglich Themen oder Stichpunkte vorgestellt, zu denen er freie Aussagen treffen kann.

Wird eine Befragung einer Zielgruppe zu mehreren Zeitpunkten durchgeführt, spricht man von einem Panel – im Gegensatz zu einer einmaligen Befragung (One shot case studies⁵⁵).

1.8.3 Fragebogen – Eigenschaften und Vergleich mit anderen Verfahren⁵⁶

Die Verwendung von Fragebögen bei Evaluationen bietet eine Reihe von Vorteilen. Sie ist ökonomisch, da zu einem Zeitpunkt eine Vielzahl von Merkmalen abgefragt werden kann. Darüber hinaus bieten Fragebögen überprüfbare Gütekriterien und eine einfache Anwendung.

⁵³ Vgl. ibidem; 387 ff.

⁵⁴ Diese Methode wird aus pragmatischen Gründen vor allem bei einer überschaubaren Stichprobe / Population praktiziert.

⁵⁵ Vgl. Rindermann H (2000). 437 – 440

⁵⁶ Vgl. Rindermann H (2001). 57 – 62

Die Entscheidung über eine Online- / internetbasierte Variante oder Printversion des Fragebogens hängt an infrastrukturellen Voraussetzungen oder dem zeitlichen Rahmen, in dem die Datenerhebung stattfinden soll.⁵⁷

Durch die vorgegebene Struktur innerhalb eines Fragebogens besteht allerdings häufig die Gefahr einer zu starken Lenkung der Befragten und die Tendenz zur Akquieszenz und sozialen Erwünschtheit.

Im Vergleich zu weiteren Datenerhebungsverfahren ist der Fragebogen allerdings das pragmatischste Mittel zur Datenerhebung.

1.8.4 Fragebogenkonstruktion

Ansatz von Rindermann⁵⁸

Bei der Konstruktion und Überprüfung des Fragebogens sollten eine Reihe von Kriterien berücksichtigt werden.

Der erste Schritt sollte dabei die Inhalte des Fragebogens betreffen. Der Fokus dabei liegt auf der Frage, welche Inhalte in welcher Form gemessen werden sollen. Sind Inhalt und Form klar, sollte festgelegt werden, mittels welcher Items die Erhebung stattfinden sollte (vgl. Kapitel 1.8.6 Itemgenerierungsverfahren) und in welcher Reihenfolge diese Items am Besten arrangiert werden. Nach diesen drei Schritten steht das Grundgerüst des Fragebogens. Es besteht Klarheit darüber, was in welcher Form mit welchen Items gefragt werden soll.

Grundsätzlich sollten Fragebogenitems die Kriterien der Verständlichkeit, Kürze, klaren Formulierung und Ein-dimensionalität erfüllen. Doppelte Verneinungen sowie *extreme* Formulierungen sollten vermieden werden. Diese Kriterien werden im folgenden Schritt überprüft.

Als nächstes sollte die Überlegung folgen, mittels welcher Skalen die Items abgefragt werden. Hierfür gibt es unterschiedliche Möglichkeiten, differenziert durch die Anzahl der vorgegebenen Antwortmöglichkeiten, verschiedene Skalen- Stufen, aufsteigende, absteigende oder zentral- fokussierte Skalen.

Eine Prämisse bei der Erstellung eines Fragebogens, bzw. Befragungsinstruments ist die unmissverständliche Führung der Teilnehmer durch eindeutige und klare Instruktionen.

Der fertige Fragebogen sollte vor dem eigentlichen produktiven Einsatz in einer Integrativumgebung empirisch bei einer kleinen Gruppe auf Inkonsistenzen, Verständnisschwierigkeiten und Verwendbarkeit überprüft und getestet werden.

Ansatz von Kromrey⁵⁹

Bei der Fragebogenkonstruktion und Zusammenstellung der einzelnen Fragen sollten nach Kromrey (2006) einige grundlegende Kriterien beachtet werden.

⁵⁷ Vgl. Kapitel 3.2.2

⁵⁸ Vgl. Rindermann (2001). 55 – 61

Zunächst sollte der Befragte an die Befragung herangeführt werden. Das geschieht mittels „leichter“ und neutraler Fragen, die den Teilnehmer bei der Beantwortung lediglich unwesentlich in Anspruch nehmen. Generell sollte auch in Kromreys Ansatz der Fragebogen klar strukturiert und für den Teilnehmer einfach nachvollziehbar sein. Hier bietet sich an, ähnliche Themenblöcke zusammenzufassen und nicht wahllos zwischen unterschiedlichen Themen zu variieren.

Zur Überprüfung der Antwortvalidität bieten Kontrollfragen die Möglichkeit, einzelne Aussagen der Teilnehmer zu hinterfragen und somit die Verlässlichkeit der Antworten sicherzustellen.

1.8.5 Alternative Datenerhebungsmethoden in Anlehnung an Rindermann⁶⁰

Dem klassischen Fragebogen stehen mehrere Erhebungsmethoden gegenüber. Projektive Tests (z.B. Deutung von Bildern und Figuren) bieten den Teilnehmern die Möglichkeit, frei zu assoziieren, wodurch ein Zugang zum Unterbewusstsein des Befragten geschaffen werden soll. Dieses Verfahren ist allerdings sehr aufwändig in der Auswertung und zudem nur schwer zu standardisieren. Für die Evaluation von Lehr- und Weiterbildungsveranstaltungen eignet sich diese Erhebungsmethode nur bedingt.

Durch Verhaltensbeobachtung ist es möglich, interessiertes Verhalten entkoppelt von der Selbsteinschätzung der Teilnehmer zu untersuchen. Ein solches Setting ist allerdings aufgrund von Hawthorne-Effekt⁶¹ sowie Testweisheit⁶² nur sehr schwer umzusetzen.

Befragungen und Interviews bieten die Möglichkeit, offen, beziehungsweise mit Hilfe eines Interviewleitfadens Daten zu Verhaltenweisen zu erheben. Diese Form der Datenerhebung besitzt eine hohe Akzeptanz, ist allerdings sehr aufwändig in der Umsetzung, kann nur schwer normiert werden und besitzt eine geringe Standardisierung. Die Einhaltung der Richtlinien zur Validität der Daten birgt bei diesem Verfahren ebenso Schwierigkeiten. Interviews und Befragungen werden deswegen häufig mit standardisierten Fragebögen kombiniert.

Eine Methode zur Leistungsmessung und Messung von physiologischen Merkmalen sind objektive Tests. Die Schwierigkeit hierbei ist allerdings die stark eingeschränkte Möglichkeit, psychische Merkmale zu untersuchen bzw. die gemessenen Daten psychologischen Merkmalen zuzuordnen.

⁵⁹ Vgl. Kromrey H (2006). 383 ff.

⁶⁰ Vgl. Rindermann H (2001). 58 f.

⁶¹ „allgemein bei reaktiven Erhebungsverfahren gefürchtete Tendenz von Versuchspersonen, wegen des Wissens um ihre Situation als Vp ihr Verhalten vorübergehend zu ändern“ (Sozialpsychologie Universität Hannover. Artikel *Hawthorne Effekt*.)

⁶² „Testweisheit“; Vgl. Rindermann H (2000). 437 – 440

1.8.6 Itemgenerierungsverfahren⁶³

Die Generierung von Fragebogenitems sollte sich grundsätzlich an der Praxis orientieren.

Bei der Itemgenerierung gibt es mehrere Verfahren, die angewendet werden können. Jedes dieser Verfahren weist Vor- und Nachteile auf.

Befragung

Bei dieser Verfahrensweise werden die Zielgruppen – Dozenten und Teilnehmer – nach den Kriterien für „gute Lehre“ befragt. Aus diesen Aussagen kann dann ein Fragebogen erstellt werden. Der Vorteil dieser Methode besteht darin, dass die Generierungskriterien explizit sind, wodurch sie genau festgelegt werden können. Ein schwerwiegender Nachteil der Befragungs- Methode ist die Auswahl der „passenden“ Inhalte und die Umarbeitung der Antwortergebnisse zu den eigentlichen Items. Der Output aus solchen Befragungen ist meistens so umfangreich, dass ein Filtern in der Regel recht kompliziert und aufwändig ist.

Synkretistische Vorgehensweise

Bei diesem Verfahren wird auf die bereits bestehende Struktur von Ergebnissen und Items zurückgegriffen. Durch eine gezielte Auswahl passender Vorgehensweisen können geeignete Items generiert werden. Der maßgebende Vorteil dieses Verfahrens besteht im Einsatz von bereits empirisch überprüften und validen Instrumenten sowie im relativ geringen Aufwand bei der Generierung des Untersuchungsapparats. Bei der synkretistischen Vorgehensweise sollte auf eine mögliche Übertragung der bereits existierenden Items und die dementsprechende Implementierung in den jeweiligen Kontext geachtet werden.

Gremienbeschluss

Der Fragebogen samt Items wird in Gremien abgestimmt und beschlossen. Hierbei fließen die Interessen aller Beteiligten in die Konstruktion des Fragebogens mit ein.

Bei dieser Vorgehensweise ist aus Akzeptanzgründen die Abdeckung aller Interessensgruppen, sowie der einzelnen Zielgruppen der Befragung ausschlaggebend. Die Gefahr bei diesem Verfahren besteht darin, dass durch die inhaltliche Abdeckung aller Interessenvertreter eventuell kritische Item bereits von vorneherein durch die einzelnen Stakeholder ausgeschaltet werden und somit lediglich ein wenig – im schlimmsten Fall überhaupt nicht – aussagekräftiges Instrument generiert wird.

Theoretisches Verfahren

Dieses Verfahren beruht auf der Verwendung bereits vorhandener Theorien und Konzeptionen, die in sich eine geschlossene Struktur bilden. Der wissenschaftliche Verteidigungswert bei diesem Verfahren ist in der Regel relativ hoch. Basis für ein theoretisches Verfahren ist die empirische Absicherung der hinzugezogenen Theorien.

⁶³ Vgl. Rindermann H (2001). 55 – 57

Autor

Der Autor generiert einen neuen Fragebogen, der auf seinen impliziten und expliziten Annahmen beruht. Die Gestaltungsfreiheit, die dem Autor bei diesem Verfahren zur Verfügung steht, kann auf verschiedene Weise bewertet werden. Einerseits besteht darin die Möglichkeit, ein Instrument einfach zu modellieren, andererseits ist diese Methode – bezogen auf den wissenschaftlichen Verteidigungswert und die oftmals fehlende Transparenz – etwas schwierig zu rechtfertigen.

Generell werden in der praktischen Umsetzung sehr häufig Kombinationen der oben genannten Verfahren eingesetzt, um ein möglichst vielseitiges und aussagekräftiges Lehrinventar zu erhalten. Dadurch wird versucht, die Schwächen der einzelnen Verfahren durch Stärken anderer auszugleichen.

1.8.7 Frageformulierung⁶⁴

Um ein einheitliches Verständnis der gefragten Inhalte bei der Zielgruppe gewährleisten zu können – die Fragen sollen bei den Teilnehmern auch so ankommen, wie sie ursprünglich im Untersuchungsplan vorgesehen sind – empfiehlt es sich, bei der Fragenformulierung auf einige Grundsätze zu achten.

„...“

- 1) *Fragen sollen so einfach formuliert sein, wie es mit dem sachlichen Zweck der Fragestellung noch vereinbart werden kann: keine komplizierten Sätze, nicht zu lange Fragen einfache Sachverhalte ansprechen.*
- 2) *Fragen sollen so eindeutig sein, dass mit der Frage ein für alle Befragten einheitlicher Bezugsrahmen geschaffen wird. Auch der gewünschte Genauigkeitsgrad der Antwort ist zu spezifizieren. (...)*
- 3) *Der Befragte darf nicht überfordert werden; d.h. sein Wissensstand darf nicht überstrapaziert, er darf nicht überfragt werden. (...)*
- 4) *Fragen sollen nicht suggestiv, sondern so neutral wie eben möglich gestellt werden.*

...“⁶⁵

1.8.8 Frage mit Einschätzungsskalen⁶⁶

Fragen mit Einschätzungsskalen dienen in der Regel dem Zweck, Daten nach dem Grad der Ausprägung oder der Akzeptanz, bzw. der Ablehnung eines Merkmals zu erheben.

⁶⁴ Vgl. Kromrey H (2006). 373 ff.

⁶⁵ Ibidem

⁶⁶ Vgl. ibidem; 377 f.

Diese Fragen sollten dabei bestimmten Kriterien entsprechen. Der verbalen Formulierung der Einschätzungsskala sollte eine numerische Abstufung zugeordnet sein (z.B. „1“ entspricht „stimme voll und ganz zu“, „4“ entspricht „Zustimmung teils / teils“ und „7“ entspricht „stimme überhaupt nicht zu“).

Diese numerische und verbale Darstellung sollte noch durch ein visuell konsistentes Layout unterstützt werden – beispielsweise durch eine gleiche Größe der einzelnen Antwortfelder.

1.9 Train-The-Trainer

1.9.1 Erwartungen an Weiterbildner⁶⁷

Die Erwartungen an Weiterbildner sind – ebenso wie der Hintergrund von Train-the-Trainer Weiterbildungsanstaltungen – sehr vielschichtig und komplex.

„Der Weiterbildner in der heutigen Berufspraxis ist Lehrender, Tutor, Dozent und Trainer, Projektleiter und Maßnahmenleiter, ehrenamtlicher Stellvertreter des Leiters einer Bildungsstätte, Akteur in kooperativen Netzwerken in der Region, Bearbeiter von Projektanträgen, Dokumenten und Verwendungsnachweisen, Evaluator, sozialpädagogischer Betreuer und Lernbegleiter, Mitglied in Prüfungskommissionen der Kammern und daneben auch als unbezahlte Fachkraft für Arbeitssicherheit, als Datenschutzbeauftragter, Systemadministrator, Qualitätsmanagementsbeauftragter oder als Leiter von Arbeitsgruppen tätig ... Im System der Rollen und im ständigen Wechsel seiner Aufgaben ist der Weiterbildner selbst Lernender und benötigt neben der öffentlichen Anerkennung auch Unterstützung.“⁶⁸

Die Aussage von Hatwieger macht die Zielsetzung von Train-the-Trainer- Seminaren deutlich, Weiterbildner auf dem Weg zu dieser umfassenden Qualifikation zu unterstützen und zu begleiten.

1.9.2 Hintergrund

Der Hintergrund einer Train-the-Trainer- Seminarkonzeption besteht vor allem darin, Personen, die sowohl im Rahmen wissenschaftlicher, als auch berufsorientierter Weiterbildung tätig sind – Dozenten, Trainer, Hochschul-lehrer, Berater oder allgemeines Fachpersonal, das sich mit den Themen Weiterbildung und Training beschäftigt – bezüglich der von ihnen angewendeten Lehre weiterzuqualifizieren und zu entwickeln, um die Qualität der Lehre zu sichern⁶⁹. Eine Fachbindung besteht dabei nicht, Train-the-Trainer ist auf jedes Fachgebiet anwendbar.

⁶⁷ Vgl. Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung e.V. (Hrsg.). QUEM-BULLETIN Jg. 1999 / Heft 5. Berlin

⁶⁸ Hatwieger K (1999) 1

⁶⁹ „Die Qualität der Weiterbildung hängt ganz entscheidend von den Kompetenzen der Lehrenden ab. Ihre fachlichen, erwachsenenpädagogischen, mediendidaktischen und planerischen Fähigkeiten bestimmen die angestrebte Entwicklung einer neuen Lernkultur.“ (BMBF. Artikel *Professionalisierung in der Weiterbildung*)

Die Bandbreite der Inhalte in Train-the-Trainer Seminaren ist in der Regel relativ weit gefächert, wobei wesentliche didaktische und pädagogische Themen sehr häufig in unterschiedlichen Seminarkonzeptionen auftauchen. In deren Mittelpunkt steht der Kompetenzerwerb für die Konzeption und Umsetzung von qualitativ hochwertigen Weiterbildungsmaßnahmen – die Gestaltung und Organisation von Lern- und Entwicklungsprozessen. Die Inhalte betreffen einerseits theoretische Aspekte, allerdings auch die praxisbezogene Anwendung dieser Theorieinputs.

1.9.3 Lernziele

Die Lernziele eines Train-the-Trainer- Seminars variieren zwischen unterschiedlichen Seminartypen. Grundsätzlich werden der Zusammenhang von der Person des Trainers, der Gruppe – in diesem Rahmen auch moderative Techniken und der Umgang mit schwierigen gruppendynamischen Prozessen – sowie inhaltlichen Themen beschrieben und Umsetzungsvarianten diskutiert und erarbeitet. Der Trainer sollen explizit auf die Ermittlung der Schulungsbedarfe – welche Inhalte sind für das Training, das der Trainer später durchführen will, relevant, welche Zielgruppe sollte angesprochen werden –, auf die Planung, Durchführung und Nachbereitung der Schulung, sowie auf die Transfersicherung nach der Schulung – Vertiefung und Nachhaltigkeit der Inhalte – vorbereitet werden.

1.9.4 Modularität

Um eine hochwertige Qualifizierung der Trainer zu erreichen, werden Train-the-Trainer- Seminarkonzepte zu meist modular konzipiert, wodurch eine Vielzahl von theoretischen und praxisbezogenen Inhalten abgedeckt werden kann. Diese Inhalte betreffen einerseits den Dozenten selbst – z.B. den Einsatz von rhetorischen oder stimmlichen Techniken zur Gestaltung einer informativen und belebten Veranstaltung – andererseits betreffen diese Inhalte die Art und Weise, wie Veranstaltungsteilnehmer am besten und effizientesten Lehrinhalte aufnehmen können.

Die Zusammenstellung der Module differiert zwischen den unterschiedlichen Seminarkonzeptionen. Betrachtet man zum Beispiel die Konzeption von PROFiL-TT, findet man die Module Rhetorik, Stimmbildung und Sprecherziehung, Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen, kohärenter Sprachgebrauch und Lehrkompetenz. Im Rahmen dieser Seminarkonzeption wird eine Vielzahl von Aspekten, die den Trainer, die Teilnehmer sowie die Konzeption und Aufbereitung der Kursinhalte betreffen, vermittelt.

1.9.5 Zielgruppe

Die Zielgruppe eines Train-the-Trainer Seminars besteht aus Dozenten, die in Trainings- oder Weiterbildungsreichen im universitären oder nicht- universitären Rahmen tätig sind. Diese Zielgruppe kann sehr variabel und vielschichtig sein, da die Trainingserfahrung der einzelnen Dozenten sehr unterschiedlich ausgeprägt sein kann.

Generell zielt das Train-the-Trainer Konzept sowohl auf Personen ohne Praxiserfahrung, die – aus beruflichen oder privaten Gründen – eine Trainings- oder Weiterbildungstätigkeit anstreben, als auch auf bereits erfahrene, im Weiterbildungsbereich tätige Personen.⁷⁰

1.9.6 Aktuelle Tendenzen im Bereich der professionellen Weiterbildung

Die aktuelle Tendenz im Bereich der professionellen Weiterbildung bezieht sich nach Nussl von Rein (2000) hauptsächlich auf vier Schwerpunkte.⁷¹

Management

Managementkompetenzen im aktuellen Rahmen der professionellen Weiterbildung nehmen vor allem hinsichtlich Bedarfsanalysen, Planung und Positionierung von Weiterbildungsmaßnahmen für die Durchführenden wesentlich an Wichtigkeit und Relevanz zu.⁷²

Methodenkompetenz

Der zunehmende Anspruch an die Qualität der Vermittlung von Lehrinhalten fordert von Weiterbildnern in jedem Bereich innovative, zielgerichtete und effiziente Lehrkonzepte. Die Berücksichtigung der projektierten Zielgruppe geht mit dieser Entwicklung einher. Wesentlich ist dabei nicht nur das Wissen um die pädagogische und didaktische Methodenvielfalt, sondern vor allem auch die aktive Anwendung und praktische Umsetzung dieser Methoden.

Medienkompetenz

Die Art und Weise, mittels der die Informationsvermittlung von Dozenten gehandhabt wird, ist für den Lernerfolg der Teilnehmer ausschlaggebend. Nussl spricht hierbei nicht nur vom Einsatz „pädagogischer Medien“⁷³, sondern generell die Verwendung von interaktiven Medien – Software, Internet, etc.

Sprachkompetenz

Nussl fokussiert mit diesem Punkt vor allem die interkulturelle Sprachkompetenz (Fremdsprachen), die vielerorts bei Weiterbildungsveranstaltungen vorausgesetzt wird. Sprachkompetenz kann allerdings auch auf z.B. den

⁷⁰ Die Zusammensetzung und Konsistenz einer Train-the-Trainer Veranstaltung kann unter verschiedenen Prämissen stattfinden. Zum Einen kann seitens der Veranstaltungskoordination eine heterogene Gruppe gewünscht sein, in der erfahrene Trainer ihr Wissen an weniger erfahrene weitergeben können – der Nutzen dabei meist recht einseitig ausgeprägt – zum Anderen besteht die Möglichkeit, verschiedene Stufen von Train-the-Trainer Seminaren zu konzipieren, in denen Dozenten mit ähnlicher Lehrerfahrung zusammengeführt werden.

⁷¹ Vgl. Nussl E (2000)

⁷² Diese Fähigkeit ist vor allem für die Projekt- und Programmkonzeption, für Mitarbeiter im administrativen Umfeld, allerdings auch für aktive Dozenten von zunehmender Bedeutung.

⁷³ Vgl. Nussl E (2000)

rhetorisch und stimmlich ausgefeilten und durchdachten Einsatz der Sprache bezogen werden, der für eine effiziente und effektive Durchführung von Lehrveranstaltungen wesentlich ist.⁷⁴

Als grundsätzliche Annahme schlägt Nuissl eine Untergliederung der in der Erwachsenenbildung beschäftigten Personen vor. Diese Aufteilung ist nicht von der Intensität, mit der die Weiterbildungstätigkeit betrieben wird, abhängig, sondern wird von der ausgeübten Funktionalität bestimmt. Nuissl unterscheidet – analog zu den aktuellen Tendenzen – vier Typen:⁷⁵

- 1) Manager / Organisator
- 2) Methodenexperten
- 3) Medienexperten / Medienpädagogen
- 4) Berater / Begleiter / Moderator

Im folgenden Verlauf der Arbeit wird der Schwerpunkt vor allem in der Ausbildung des ersten Typs – Methodenexperte, also des didaktisch geschulten Dozenten – liegen, da das Projekt PROFiL-TT diese Zielgruppe als Hauptbezugspunkt aufgreift.

⁷⁴ Vgl. PROFiL-TT. <http://www.profil-tt.lmu.de> (Abgerufen: 03. Januar 2007)

⁷⁵ Vgl. Nuissl E (2000)

2 Evaluation

Aus sozialwissenschaftlicher Sicht besteht eine Evaluation aus der Bewertung eines Programms mit Hilfe empirischer Daten. Diese Betrachtung bewegt sich in der Regel im Rahmen anwendungsbezogener Forschung.⁷⁶

Allgemein definiert bedeutet Evaluation „die systematische Untersuchung der Verwendbarkeit oder Güte eines Gegenstandes“⁷⁷, wobei der Begriff *Gegenstand* in diesem Zusammenhang auf Programme, Projekte, Materialien, Personal, etc. bezogen werden kann.

Im pädagogischen Sinn wird der Begriff „Programm“ als „Pädagogische Tätigkeiten, die kontinuierlich stattfinden“⁷⁸ bezeichnet. Der wesentliche Unterschied zwischen Projekt und Programm besteht in der Kontinuität und Fortführung, bzw. in zeitlichen Intervallen. Programme finden in regelmäßigen Abständen ab und sind auf Dauer angelegt. Projekte unterliegen in der Regel einer zeitlichen Begrenzung und Einschränkung. Dauerhaft eingerichtete Projekte werden zu Programmen.⁷⁹

*„Programme sind beschriebene und durchgeführte, intentional aufeinander bezogene Bündel von Interventionen, Maßnahmen, Projekten oder Teilprogrammen. Sie bestehen aus einer Folge von auf ausgewiesene Ziele hin ausgerichtete Aktivitäten. Sie werden auf der Basis von verfügbaren Ressourcen durchgeführt und sind darauf gerichtet, mittels bereitgestellter Leistungen (outputs) bestimmte (outcomes) bei bezeichneten Zielgruppen oder im sozialen System zu erreichende impacts auszulösen.“*⁸⁰

Im Fokus steht bei dieser Definition zunächst die Regelmäßigkeit und Ordnung der im Rahmen eines Programms durchzuführenden Tätigkeiten, die von personellen, finanziellen oder ähnlichen Rahmenbedingungen abhängen und durch deren Einsatz etwas bewirkt, ein Ziel verfolgt wird.

Der Begriff „Programm“ soll in dieser Arbeit alle Arten von Evaluationsgegenständen (Programme, Projekte, Materialien, Maßnahmen und Interventionen) aus den Bereichen Erziehung, Aus- und Weiterbildung umfassen.

Merkmale, Zielsetzungen und Leitfunktionen von Evaluationen werden im weiteren Verlauf der Arbeit näher spezifiziert.

2.1 Sozialwissenschaftliche Bedeutung von Evaluation⁸¹

Aus sozialwissenschaftlicher Sicht kann im Rahmen einer Evaluation einerseits eine spezifische Fragestellung, andererseits ein spezifisches Design verfolgt und untersucht werden.

⁷⁶ Vgl. Kromrey H (2006). 74

⁷⁷ Beywl W / Widmer T (2006). 28

⁷⁸ Ibidem

⁷⁹ Vgl. ibidem

⁸⁰ Beywl W (2004). Artikel *Programm*

⁸¹ Vgl. Kromrey H (2006). 102 ff.

Spezifische Fragestellung

Vor diesem Hintergrund hat eine Evaluation den Zweck der Bewertung eines bestimmten Sachverhaltes oder Handelns anhand von vorgegebenen Kriterien mit Hilfe empirischer Informationen.

Spezifisches Design

Im Rahmen eines spezifischen Designs nimmt die Evaluation in methodisch kontrollierter Weise Bezug auf die Bewertung des zu evaluierenden Handelns, bzw. des untersuchten Programms und der daraus resultierenden Effekte.

Voraussetzungen für diese Bewertung ist eine fundierte Untersuchung und Beschreibung der existierenden Randbedingungen, des Ist- Zustands zum Zeitpunkt t_0 , der vorgenommen Eingriffe in die Randbedingungen – intendiert und nicht intendiert – während der laufenden Maßnahme, sowie des Zustands nach der Programmdurchführung.

Ein allgemeine Problematik besteht hierbei im Abgleich der Effekte zwischen Anfangszeitpunkt t_0 und Endzeitpunkt. Generell stellt sich hier die Frage, ob die gemessenen Effekte auf die Intervention zurückzuführen – und somit intendiert – sind, oder ob die Effekte auf programmexterne Einflüsse bezogen werden müssen. Eine eigentliche Programmeinschätzung, bzw. Bewertung kann in erster Linie mittels der beabsichtigten Ziele eines Programms durchgeführt werden. Als Fazit kann hieraus gezogen werden, dass eine Intervention dann als „erfolgreich“ bezeichnet werden kann, wenn eine Beeinflussung in Richtung der beabsichtigten Programmziele erkennbar ist. Wünschenswert ist dabei, dass die Zielvariablen in gewünschter Richtung und Stärke beeinflusst werden. Der Einfluss von nicht- intendierten Faktoren sollte dabei nicht außer Acht gelassen werden, da hierdurch oftmals starke Effekte beobachtet werden können.

Eine derart umfassende Untersuchung kann als Ergebnis sowohl Aussagen über den Erfolg oder Misserfolg einer Maßnahme treffen, sowie wichtige Informationen zur Verbesserung liefern.

In der Praxis ist diese Vorgehensweise – Evaluierung einer Maßnahme und aller Rahmenbedingungen, sowie der intendierten und nicht- intendierten Effekte – meist nicht realisierbar. Bei der praktischen Umsetzung von Evaluationsprojekten ist eine Fokussierung auf Schwerpunkte notwendig, da die streng empirische Vorgehensweise nur bedingt der pragmatischen Kriterien entsprechen kann.

2.2 Zielsetzungen und drei klassische Unterscheidungen von Evaluation

Eine Evaluation kann sich auf verschiedene Zielsetzungen fokussieren. Kromrey führt drei übliche Unterscheidungen im Bezug auf den Gegenstand „Evaluation“ an.⁸²

⁸² Vgl. ibidem; 105 f.

Wirkungsanalyse

In einer Wirkungsanalyse werden als Schwerpunkt die Effekte untersucht, die durch ein Programm initiiert werden.

Implementationsforschung

Der Fokus bei der Implementationsforschung liegt auf der systematische Planung, Durchsetzung und Umsetzung eines Programms.

Akzeptanzforschung

Wird eine Evaluation unter dem Gesichtspunkt der Akzeptanzforschung durchgeführt, stehen im Vordergrund, ob und inwiefern eine Maßnahme bei der Zielgruppe angenommen wird, beziehungsweise das Angebot auf Zustimmung oder Ablehnung stößt.

2.3 Merkmale und Kennzeichen von Evaluation⁸³

Führt man den Begriff „Evaluation“ zurück auf seine lateinischen Wurzeln, stößt man in der Übersetzung des Wortstamms „valere“ auf Begriffe wie „Einfluss haben“, „gültig sein“ und „wert sein“. Eine Evaluation dient somit unter semantischen und etymologischen Aspekten neben dem Werkzeug als Planungs- und Entscheidungshilfe auch als Bewertungsinstrumentarium von Handlungs- und Programmalternativen.

Ein wesentliches Merkmal von Evaluation ist nach Wottawa ein generell ziel- und zweckorientierter Charakter. Evaluation ist ein Hilfsmittel zur Überprüfung praktischer Maßnahmen und als solches vorwiegend handlungs- und praxisorientiert.

Nach Wottawa besteht ebenso Einigkeit darüber, dass Evaluationsmaßnahmen stets an den aktuellen Forschungs- und Entwicklungsstand der Wissenschaft adaptiert werden.

Evaluation bedient sich der Ergebnisse der aktuellen empirisch- wissenschaftlichen Forschung, um diese in der Praxis zielorientiert einzusetzen. Sie wandert somit auf dem schmalen Grad zwischen theoretischer Empirie und praktischer Anwendungsorientierung.

2.4 Nutzenerwartung der Auftraggeber einer Evaluation nach Wottawa⁸⁴

Evaluation ist in der Regel Auftragsforschung und hat als solche ein vielschichtiges Spektrum an Nutzenerwartungen seitens der Auftraggeber, allerdings auch der Anwender.

⁸³ Vgl. Wottawa H, Thierau H (2003). 14

⁸⁴ Vgl. ibidem; 29 ff.

Evaluationen können generell mit folgenden Zielsetzungen, beziehungsweise Hintergründen in Auftrag gegeben werden:

1. Bewertung ohne detaillierte Zielsetzung
2. Verantwortungsdelegation
3. Durchsetzungshilfe
4. Entscheidungshilfe
5. Optimierungsgrundlagen
6. Kontrolle der Zielerreichung

Bewertung ohne detaillierte Zielsetzung

Wird eine Evaluation ohne detaillierte Zielsetzung in Auftrag gegeben und durchgeführt, hat dies meist ein Potential an Misstrauen der Auftraggeber gegenüber der fokussierten Maßnahme als Hintergrund. Im Untersuchungsfokus steht bei solchen Evaluationen in der Regel die Frage, welchen Nutzen eine Intervention hat. Wotawa (2003) spricht in diesem Umfeld auch von „Überprüfung“ und „Kontrolle“; der Evaluator sieht sich im Rahmen einer solchen Evaluation häufig mit geringer Akzeptanz und Widerständen innerhalb der Maßnahme – zum Beispiel Fehlinformationen, information-hiding oder Ähnliches – konfrontiert.

Verantwortungsdelegation

Der Hintergrund einer Evaluation mit der Nutzenerwartung *Verantwortungsdelegation* besteht in der Regel aus der Ungewissheit einem neuen Projekt oder einer neuen Maßnahme gegenüber. Der Auftraggeber versucht die Verantwortung für eine – meist selbst initiierte – Innovation abzugeben und auf den Evaluator zu übertragen.

Häufig werden Interventionen mit einem knappen Mehrheitsbeschluss durchgesetzt und haben innerhalb der durchführenden Organisation lediglich einen schwachen Rückhalt.

Ziel der Auftraggeber ist es in der Regel die Verantwortung, die sie für die Umsetzung einer Innovation tragen müssen, durch das Heranziehen von Evaluatoren zu verringern oder ganz abzugeben.

Diese Vorgehensweise impliziert vorsichtiges – in gewisser Weise auch skeptisches – Verhalten der Auftraggeber gegenüber der eingeführten Intervention.

Die Erwartungshaltung gegenüber der Evaluation ist in der Regel eine Verbesserung des aktuellen Zustands, bzw. zumindest eine Veränderung des Status quo durch den evaluativen Eingriff.

Durchsetzungshilfe

Die Auftraggeber erwarten sich im Rahmen dieser Zielsetzung von den Evaluationsergebnissen Rückhalt, die Akzeptanz der evaluierten Maßnahme zu stärken, beziehungsweise im Gegenteil, diese zu schwächen. Das gewünschte Resultat ist dabei eine erleichterte Durchführung der Maßnahme oder ein Verhindern.

Dieser Hintergrund wird häufig bei größeren Projekten, umfangreichen Systemänderungen und weit reichenden Umorganisationen verwendet. Untersucht wird in der Regel im Rahmen eines Test- Umfelds, um einerseits ein gewisses Maß an Akzeptanz für die Änderung zu schaffen und andererseits die Innovation zu etablieren.

Befürworter und Gegner einer innovativen Änderung ziehen gleichermaßen Vorteile aus dieser Vorgehensweise, da sie den konträren Seiten Zeit verschafft, sich mit der Innovation – wenn auch nur in einem abgegrenzten Rahmen umgesetzt – zu beschäftigen und die Vor- und Nachteile in der Praxis zu erfahren.

Die Deckung der Evaluationsergebnisse mit den Zielen der Auftraggeber ist dabei gewünscht.

Entscheidungshilfe

Voraussetzung für diese evaluative Zielsetzung sind Alternativen zum untersuchten Programm, dessen Inhalte, Auswirkungen und Effekte genau beschrieben werden sollen.

Der Evaluator untersucht die Maßnahme anhand der Entscheidungskriterien und vergleicht sie im Anschluss mit den Alternativen. Der Vergleich sollte dabei umfassend die Aspekte der unterschiedlichen Alternativen beleuchten.

Die finale Entscheidung für die beste Alternative dieser Auswahl liegt bei den Auftraggebern der Studie.

Optimierungsgrundlage

Der Evaluator nimmt in diesem Setting eine insofern angenehme Rolle wahr, dass er nicht konkurrierende Alternativen gegenüberstellen und selektieren muss. Die Vor- und Nachteile einer Maßnahme werden herausgearbeitet und explizit gemacht. Voraussetzung für den Evaluator ist in dieser Form der Evaluation allerdings neben der Methodenkenntnis in der Evaluation auch ein fundiertes Fachwissen, bezogen auf die Maßnahme (vgl. Kapitel 2.8.4 Interne Evaluation / Selbstevaluation).

Kontrolle der Zielerreichung

Ziel der Auftraggeber bei dieser Vorgehensweise ist die Untersuchung der Effizienz eines Programms / einer Intervention. Der Evaluator untersucht die Qualität – wurde die geplante Kosten- Nutzen- Rechnung erreicht oder nicht. Der Abgleich der projektierten Programmziele mit den tatsächlich erreichten Zielen ist dabei wesentlich.

2.5 Typische Untersuchungsdesigns / Evaluationsdesigns⁸⁵

Der Anspruch der Evaluatoren, wissenschaftlich fundierte und objektive Untersuchungsverfahren einzusetzen ist generell nachvollziehbar, muss allerdings auch stets unter der Prämisse der Umsetzbarkeit und der praktischen Anwendung betrachtet werden. Unter diesen Voraussetzungen ist es „sinnvoller (...) mit einer ‚weichen‘, weniger exakten und elaborierten Methode ein interessantes Kriterium zu messen, als mit hoher Präzision etwas, was inhaltlich nicht zu den eigentlichen Evaluationszielen paßt“⁸⁶. Der „Verteidigungswert“ bei wissenschaftlich fundierten Methoden ist allerdings ungleich höher, als bei solchen „ad-hoc“- Instrumenten.

Grundsätzlich stellt sich somit die Frage nach dem Grad der Wissenschaftlichkeit der angestrebten Untersuchung. Es besteht dabei die Möglichkeit zwischen „echten“ wissenschaftlichen Experimenten, Quasiexperimenten bis hin zu stark pragmatischen Designs wie Pretest- Posttest- Untersuchungen oder Gutachtenmodellen zu wählen. Wesentlich bei dieser Auswahl ist die Häufigkeit und Art der Datenerhebungszeitpunkte und die Existenz oder Beschaffenheit einer Kontrollgruppe.

Die Designs zur Untersuchung von Maßnahmen beruhen in der Regel auf Vergleichen. Es werden zwei Gruppen gebildet:

- eine Experimentalgruppe, an der die Intervention vorgenommen wird und
- eine Kontrollgruppe, die von der Intervention nicht betroffen ist

Diese Vorgehensweise ist für den Vergleich der Resultate (hat die Intervention einen Effekt, wenn ja einen positiven oder einen negativen) notwendig.

2.5.1 Experimente / Feldexperimente⁸⁷

Echte Experimente entsprechen den Anforderungen zur Überprüfung einer kausalen Anordnung. Sie tragen der zeitlichen Abfolge von Maßnahme und Wirkung, dem Zusammenhang zwischen Maßnahme und Wirkung sowie der Kontrolle von Drittvariablen (mittels Randomisierung und Matching) Rechnung.

In realen sozialen Situationen ist allerdings die Verwendung von echten experimentellen Designs nie vollständig möglich.⁸⁸

Echte experimentelle Designs sind trotz der – gegenüber anderen Designtypen – Überlegenheit in der Wirkungsmessung nicht unproblematisch umzusetzen.

Theoretische und empirische Restriktionen verhindern die Vollständigkeit der einzubeziehenden, erklärenden Faktoren. Eine Intervention durch einen Forscher löst durch die Untersuchung automatisch eine Veränderung

⁸⁵ Vgl. Stockmann R (2006). 229 ff.

⁸⁶ Wottawa H, Thierau H (2003). 132

⁸⁷ Vgl. Stockmann R (2006). 230 ff.

⁸⁸ Vgl. Kromrey H (2002). 96 ff.

aus, deren Effekte von den Wirkungen der Intervention getrennt werden müssen (Hawthorne- Effekt).⁸⁹ Dieser Effekt kann lediglich mittels eines schwer zu realisierenden Solomon-Vier-Gruppen-Designs vermieden werden, also eine Untersuchung mit mehreren Kontrollgruppen, die zu mindestens zwei Messzeitpunkten mit den gleichen Instrumenten untersucht werden.

Da die Teilnahme an einer Untersuchung in der Regel aus freiem Willen erfolgt, besteht von Beginn an in diesen Designs die Gefahr einer Selbstselektion der Teilnehmer.

Vor allem bei formativen Untersuchungen ist eine Wiederherstellung der Versuchsbedingungen schwierig, da bereits durch die Maßnahme interveniert wurde, was eine Rückkehr zum Ausgangspunkt nahezu unmöglich macht. Somit muss in diesem Fall bei Messwiederholungen stets mit einer Verzerrung der Untersuchungsbedingungen gerechnet werden.⁹⁰

Über die Ergebnisse, die durch den Einsatz von echten experimentellen Designs bei Untersuchungen erzielt werden, besteht generell Uneinigkeit.

So konnten z.B. in vielen Untersuchungen des „Great Society Programs“ unter der Verwendung von experimentellen Designs keine signifikanten Programmeffekte nachgewiesen werden, was allerdings auch so interpretiert werden kann, dass es tatsächlich keine Programmeffekte gab.⁹¹ Das Resultat aus dieser Ungewissheit war der Trend, Designs immer aufwändiger und kostenintensiver zu gestalten, weswegen sich die Meinung konstatierte, dass dieser *optimale* Evaluationsansatz zu anspruchsvoll, kostenintensiv und letztendlich für die Durchführung von Forschungsprojekten ungeeignet sei.⁹²

Eine weitere Schwierigkeit, die aus der Anwendung von experimentellen Designs resultiert, besteht in der Komplexität der Untersuchungsgegenstände. So können echte Experimente bei Untersuchungen benutzt werden, die eine relativ geringe Anzahl von Parameter umfassen. Auf diese Weise können die Effekte der einzelnen Parameter schrittweise untersucht werden. Bei der Untersuchung von komplexen Phänomenen – beispielsweise Organisationen oder vielschichtigen Weiterbildungsinstrumenten – ist die Verwendung von experimentellen Untersuchungsdesigns weniger zielführend.⁹³

Stockmann beschreibt weitere Probleme experimenteller Designs. So können „auch im Sinne der experimentellen Logik ‚schwache‘ Designs [...] durchaus eine hohe interne Validität besitzen, wenn die Beobachtungsbefunde nur auf eine Weise plausibel erklärt werden können, und umgekehrt erzielt ein ‚echtes‘ Experiment nur eine schwache interne Validität, wenn es für das Zustandekommen der Befunde gleichermaßen plausible konkurrierende Erklärungen gibt.“⁹⁴

⁸⁹ Vgl. Thompson RJ (1990). 379ff

⁹⁰ Vgl. Stockmann R (2006). 231

⁹¹ Vgl. ibidem

⁹² Vgl. Binnendijk AL (1989). 209

⁹³ Vgl. Stockmann R (2006). 232

⁹⁴ Ibidem

Laborergebnisse müssen ebenso nicht unbedingt intern und extern hoch valide sein, da in diesem Umfeld häufig Versuchleitereffekte oder Messeffekte, bedingt durch die Laborsituation auftreten können.⁹⁵

Andererseits besteht auch bei Feldexperimenten die Möglichkeit, dass Ergebnisse – z.B. aufgrund des Hawthorne-Effekt – verzerrt dargestellt werden. Die Situation, die mittels des Untersuchungseingriffs geschaffen wird stellt fast immer einen Eingriff in den Alltag und das gewohnte Umfeld der Zielpersonen dar.

Im Rahmen von sozialwissenschaftlichen Untersuchungen verhindern oftmals ethische und rechtliche Aspekte den Einsatz von empirischen Untersuchungsdesigns, da es beispielsweise bei der Implementierung eines Sozialprogramms in der Regel nicht zu rechtfertigen ist, einer sozialen Gruppe den Zugang zu einer Intervention zu verweigern – um eine Kontrollgruppe bilden zu können (siehe Korrektheitsstandards Kapitel 2.10.5).⁹⁶

2.5.2 Quasi-Experimente

Vor dem Hintergrund solcher praktischen Einschränkungen wird häufig auf Quasi-Experimente zurückgegriffen, die den Einsatz der Untersuchungsinstrumente und die Rahmenbedingungen nicht auf derartig strenge Weise reglementieren, allerdings auch nicht einen so hohen wissenschaftlichen Verteidigungswert besitzen wie „echte“ Experimente.

Bei Quasi-Experimenten wird die Zuteilung der Versuchspersonen zu den einzelnen Versuchsgruppen nicht randomisiert oder parallelisiert. Im Rahmen eines Quasiexperiments können Abhängigkeiten von Variablen untersucht und bestimmt werden, es kann jedoch kein Rückschluss auf die Richtung der Abhängigkeit zwischen den Variablen gezogen werden.⁹⁷

2.5.3 Ex-Post-Facto-Designs

Ex-Post-Facto-Designs werden für Untersuchungen verwendet, wenn die Rahmenbedingungen weder experimentelle noch quasiexperimentelle Designs erlauben. Ähnlich wie bei quasiexperimentellen Untersuchungsplänen werden auch in diesem Typ abhängige und unabhängige Variablen untersucht, wobei auch in diesem Fall Störvariablen nicht überprüft werden können. Der wesentliche Vorteil von Ex-Post-Facto-Untersuchungen besteht in einer meist hohen Effizienz des Untersuchungsapparats, der mit relativ geringem Aufwand an Ressourcen Rückschlüsse und Verallgemeinerungen erlaubt.

Grundsätzlich können Ex-Post-Facto-Designs in Quer- und Längsschnittuntersuchungen eingeteilt werden.⁹⁸

⁹⁵ Vgl. *ibidem*; 233

⁹⁶ Vgl. *ibidem*

⁹⁷ Vgl. Campbell DT, Stanley JC (1963). 171–246

Querschnittstudien⁹⁹

Im Rahmen dieses „Ex-Post-Facto-Designs“ wird aus der Zielpopulation eine Stichprobe gezogen, wobei Personen erfasst werden, die an einer Intervention oder Maßnahme teilgenommen haben, sowie Personen, die diese Intervention nicht betroffen hat. Die Ergebnisse beider Gruppen werden miteinander verglichen, um die Wirkung der Zielvariablen zu bestimmen. Eine kausale Reihenfolge kann bei diesem Design nicht berücksichtigt werden, da es keinen zeitlichen Verlauf der Untersuchungsergebnisse gibt und die Erhebung zu einem bestimmten Zeitpunkt stattfindet.

Längsschnittstudien¹⁰⁰

Während die Untersuchungsprämisse bei Querschnittuntersuchungen, darauf liegt, dass Daten verschiedener Personengruppen an einem bestimmten Zeitpunkt erhoben werden, liegt der Schwerpunkt bei Längsschnittuntersuchungen auf der wiederholten Durchführung einer Untersuchung. Ein Vergleich innerhalb eines zeitlichen Verlaufs ist somit bei diesem Design möglich. Die Auswahl der Untersuchungszielgruppe wird dabei nach unterschiedlichen Aspekten getroffen. Bei Panel- Untersuchungen bleiben die an der Untersuchung teilnehmenden Personen die gleichen, wohingegen bei Trendstudien sich die Untersuchungszielgruppe – personell – verändert; die Auswahlkriterien sind allerdings auch bei Trendstudien konstant.

Pretest- Posttest- Untersuchungen¹⁰¹

Im klassischen Pretest- Posttest- Designs findet eine Datenerhebung vor der geplanten und eine nach der durchgeführten Intervention statt. Im Hintergrund dieses Untersuchungsdesigns steht die Annahme, dass beide Messungen ohne eine dazwischenliegende Intervention identisch sein sollten. Der Nettoeffekt einer Maßnahme sollte über die Differenz der Ergebnisse aus Pretest und Posttest bestimmt werden. Die Besonderheit bei dieser Vorgehensweise – wie bei vielen Längsschnittstudien – besteht darin, dass Kontrollgruppe und Experimentalgruppe identisch sind. Eine Problematik – vor allem bei Wirkungsanalysen – wird häufig in der schwierigen Differenzierungs- und Abgrenzungsmöglichkeit zwischen Wirkung und Störfaktoren gesehen. Diese Problematik wird bei einer speziellen Variante dieses Designs – anstelle der Vorher- Messung wird bereits nach Beginn der Maßnahme eine retrospektive Befragung der Situation vor Interventionsbeginn durchgeführt – weiter verschärft, da in dieser Form der Datenerhebung ein zusätzlicher Faktor (bedingt durch Vergessenseffekte), einfließt.

⁹⁸ Vgl. Schnell R, Hill PB, Esser E (2005). 228 – 231

⁹⁹ Vgl. Stockmann R (2006). 234

¹⁰⁰ Vgl. ibidem; 235 f.

Panelanalysen¹⁰²

Der wesentliche Vorteil bei Panelanalyse besteht in der wiederholten Durchführung einer Untersuchung mit derselben Befragungsgruppe zu unterschiedlichen Zeitpunkten. Durch diese Wiederholung können abhängige und unabhängige Variablen aufeinander bezogen werden. Aufgrund dieses Vergleichs der verschiedenen Erhebungswellen können differenzierte Aussagen über die Entwicklung der untersuchten Variablen getroffen werden. Der wesentliche Nachteil von nicht- experimentellen Designs – das Problem der kausalen Reihenfolge –, kann durch die wiederholte Durchführung der Untersuchung abgeschwächt werden, da dieselben Variablen mit derselben Operationalisierung an derselben Personengruppe mehrfach erhoben werden.

Zeitreihenanalysen¹⁰³

Bei Zeitreihenanalysen handelt es sich um diskrete, in endlichen zeitlichen Abständen anfallende Datenerhebungen.

Mittels Zeitreihenanalysen kann eine Prognose über zukünftige Zeitreihenwerte gestellt werden.¹⁰⁴

Aus diesen regelmäßigen, größtenteils amtlichen Statistiken wird – als Ersatz für eine Kontrollgruppe – die Entwicklung von Personen geschätzt, die nicht an einer Intervention teilgenommen haben.

Die Vorgehensweise bei der Zeitreihenanalyse ist dabei wie folgt: Vor der Intervention werden in zeitlichen Abständen Messungen unternommen, woraus sich ein Trend errechnen lässt, der darüber Aufschluss gibt, wie die Entwicklung ohne Intervention abgelaufen wäre.

Das größte Problem bei der Verwendung der Zeitreihenanalyse ist nach Stockmann (2006) die umfangreiche Anzahl von Messungen, die vor der eigentlichen Intervention notwendig ist (Stockmann empfiehlt mindestens 30 Messungen), um eine verlässliche Aussage über den Trend machen zu können. Aus diesem Grund wird die Zeitreihenanalyse in der Regel nur bei Projekte angewendet, bei denen dieser Datengrundstock ohnehin bereits existiert (z.B. bei amtlich erhobenen Daten).

2.5.4 Gutachtenmodell

Das Gutachtenmodell stellt die kostengünstigste, allerdings auch die – nach Stockmann (2006) – unzuverlässigste und ungenaueste Variante eines Untersuchungsdesigns dar.

¹⁰¹ Vgl. ibidem

¹⁰² Vgl. ibidem

¹⁰³ Vgl. ibidem; 236

¹⁰⁴ Vgl. Schlittgen R, Streitberg BHJ (2001). 1 – 8

Kernstück sind dabei Bewertungen von Experten, administrativen Kräften und Interventionsteilnehmern, die zur Schätzung der Nettowirkung einer Maßnahme herangezogen werden. Rossi, Freeman und Lipsey (1999) charakterisieren dieses Design als eine eher fragwürdige Möglichkeit zur Durchführung einer Evaluation.¹⁰⁵ Ferner stehen Rossi, Freeman und Lipsey dieser Untersuchungsmethode kritisch gegenüber, da sie der Ansicht sind, dass es bei der Befragung der Nutzer einer Dienstleistung als „Experten“ äußerst schwierig ist, ein Urteil über den Nettoeffekt einer Intervention abzugeben, da dieser Zielgruppe häufig das nötige Hintergrundwissen für eine solche Einschätzung fehlt.¹⁰⁶

Kromrey (2006) schlägt für diesen Untersuchungstypus bei dem „individuelle parteiische Werturteile von Personen, die in einer besonderen Beziehung – eben als Nutzer, als Betroffene – zum Untersuchungsgegenstand stehen“¹⁰⁷ die Bezeichnung „Akzeptanzerhebung“ vor.

Stockmann (2006) ist der Ansicht, dass das Gutachtenmodell zwar wichtige Merkmale zur Bewertung von Programmen liefert, allerdings keinen Ersatz für eine Wirkungsuntersuchung darstellt.¹⁰⁸

Ein wesentlicher Vorteil dieses Verfahrens ist – neben der unkomplizierten, kostengünstigen und ergebnisorientierten Umsetzung – ein Entgegenwirken der allgemeinen Akzeptanzproblematik im Rahmen von Evaluationsstudien, in dem eine besonders skeptische Betrachtung und Beurteilung der Evaluationsergebnisse immer mehr zunimmt.¹⁰⁹

2.5.5 Untersuchungsdesigns im Überblick¹¹⁰

Bei der Wahl des Evaluationsdesigns ist es generell wichtig, dass die aus der Erhebung hervorgegangenen Daten valide und reliabel sind, damit das Programmmanagement und die Stakeholder aus der Erhebung einen möglichst großen Nutzen ziehen können.

Wichtig ist ebenso die Transparenz für alle Beteiligten – dies gilt im Besonderen bei der Wahl des Untersuchungsdesigns. Schwächen und Stärken sollten dabei offen kommuniziert und die Auswahl begründet werden.

¹⁰⁵ Vgl. Rossi, Freeman und Lipsey (1999). 269

¹⁰⁶ Vgl. ibidem; diese Aussage trifft in der vorliegenden Untersuchung allerdings nicht zu, da Evaluation als Bestandteil der Weiterbildungsmaßnahme PROFIL-TT im Modul Lerntheorie ist. Es kann somit nicht davon ausgegangen werden, dass die Teilnehmer mangelndes Wissen über Evaluation im Allgemeinen und in der Anwendung im Besonderen haben.

¹⁰⁷ Kromrey H (2006). 105 f.

¹⁰⁸ Vgl. Stockmann R (2006). 238

¹⁰⁹ Vgl. ibidem

¹¹⁰ Vgl. ibidem; 229

Design	Art der Kontrollgruppe	Datenerhebungszeitpunkte
'Echte' Experimente / Feldexperimente	Randomisierte Kontrollen, oft zusätzlich statistische Kontrollen	Minimum: nur nach der Intervention. Meist vorher und nachher; oft mehrere Messungen während der Intervention
Quasi-Experimente	Konstruierte und / oder statistische Kontrollen	Minimum: nur nach der Intervention. Meist vorher und nachher; oft mehrere Messungen während der Intervention
Querschnittanalysen	Statistische Kontrollen	Nur Nachher- Messungen
Pretest- Posttest-Untersuchungen	Reflexive Kontrollen	Minimum: Vorher- und Nachher-Messung
Retrospektive- Vorher-/Nachher-Untersuchungen	Retrospektive reflexive Kontrollen	Nachher- Messungen mit retrospektiven Messungen der Ausgangssituation
Panel-Untersuchungen	Reflexive Kontrollen	Mehr als zwei Messungen während der Intervention
Zeitreihen Analysen	Reflexive Kontrollen	Viele Messungen vor und nach der Intervention
Gutachtenmodell	Generische und / oder Schattenkontrollen	Nur Nachher- Messungen

Tabelle 1: Übersicht typischer Untersuchungsdesigns (vgl. Stockmann 2006, 229)

2.6 Störfaktoren

Im Rahmen von Evaluationsuntersuchungen können Störfaktoren – zum Beispiel verzerrte Auswahlkriterien bei der Zusammensetzung der Experimental- und Kontrollgruppen¹¹¹ – auftreten, die die Güte der Datenerhebung teilweise beträchtlich mindern.

Die Messergebnisse können darüber hinaus durch Verwendung unterschiedlicher Messinstrumente oder ein divergierendes Verhalten des Evaluators bei den Untersuchungsgruppen beeinträchtigt werden.¹¹²

Es gibt verschiedene Möglichkeiten Störfaktoren auszuschalten, beziehungsweise ihr Auftreten zu kontrollieren.

¹¹¹ Der Unterschied zwischen Experimental- und Kontrollgruppe betrifft unter Umständen nicht nur die Intervention, sondern bezieht sich auch auf weitere Merkmale.

¹¹² Vgl. Stockmann R (2006). 226

Möglichkeiten, Störfaktoren zu verhindern¹¹³

Randomisierte Kontrollen

Die untersuchten Probanden werden über zufällige Auswahl der Experimental- bzw. Kontrollgruppe zugewiesen. Jeder Teilnehmer hat dabei die gleiche Chance, in die Experimental- sowie in die Kontrollgruppe zu kommen.

Konstruierte Kontrollen

Die Experimental- und Kontrollgruppe wird in diesem Verfahren durch eine Art „coupling“ gebildet. Es werden immer zwei Personen ausgewählt, die sich in bestimmten Merkmalen ähnlich sind. Diese „Partner“ werden dann jeweils einer auf die Kontroll- und einer auf die Experimentalgruppe verteilt. Nach Rossi, Lipsey und Freeman (2004) entsteht durch die „Parallelisierung“ eine konstruierte Kontrollgruppe.¹¹⁴

Statistische Kontrollen

Die Kontrollgruppe bei der statistischen Kontrolle wird erst nach der Datenerhebung gebildet. Die Vergleichsgruppe entsteht somit erst bei der Datenanalyse. Es wird ein statistisches Abbild mit allen ausschlaggebenden Kriterien und Merkmalen der Versuchsgruppe gebildet.

Reflexive Kontrollen

Bei dieser Form der Kontrolle von Störvariablen bildet die Experimentalgruppe auch die Versuchsgruppe. Die Messung der Experimentalgruppe wird vor und nach der Intervention beobachtet. Dabei wird davon ausgegangen, dass keine Änderungen in den Zielgrößen auftauchen, die nicht direkt auf die durchgeführte Intervention zurückzuführen sind.¹¹⁵ Der Unterschied zwischen der ersten und der zweiten Messung wird somit als Wirkung bzw. interventionsbezogener Effekt interpretiert.

Diese Art der Kontrolle birgt allerdings – neben dem wesentlichen Vorteil einer komplett homogenen Experimental – und Vergleichsgruppe – auch Nachteile:

- Die beobachtete Veränderung kann nicht klar auf die Intervention zurückgeführt werden (diese Problematik lässt sich mit einer größeren Anzahl von Messwerten abschwächen, da in diesem Fall ein Trend sichtbar wird)
- Ein weiteres Problem bei dieser Form der Kontrolle ist, dass die Versuchspersonen – je weiter die Messzeitpunkte in die Vergangenheit zurückreichen und je länger das Programm dauert – zunehmend altersbedingten Effekten sowie historischen Ereignissen ausgesetzt sind.

Generische Kontrollen

Generische Kontrollen sollten nach Rossi, Freeman und Lipsey lediglich dann eingesetzt werden, wenn sie die einzige Kontrollmöglichkeit bieten und andere Kontrollen nicht verwendet werden können,¹¹⁶ da hierbei auf

¹¹³ Vgl. ibidem; 226 ff.

¹¹⁴ Vgl. Rossi, Freeman und Lipsey (2004). 275

¹¹⁵ Vgl. Ibidem; 290

¹¹⁶ Vgl. Rossi, Freeman und Lipsey (1999). 332

allgemein anerkannte empirische Werte zurückgegriffen wird, die allerdings nur in wenigen sozialen Bereichen vorliegen (z.B. Sterbe- und Fruchtbarkeitsraten). Sie dienen als Ersatz für eine Kontrollgruppe.

Schattenkontrollen

Bei der Verwendung von Schattenkontrollen („Shadow Controls“, Rossi, Freeman und Lipsey 1999, 356) wird die interventionsbedingte Wirkung mit der Wirkung verglichen, die ohne die Intervention zu erwarten gewesen wäre – also „normalerweise“. Hierbei werden Expertenmeinungen, bzw. die Meinungen von Programmleitern und Teilnehmer herangezogen. Schattenkontrollen werden vorwiegend aus Kostengründen eingesetzt.

2.7 Deskriptive Statistik

„Die beschreibende (deskriptive) Statistik zielt darauf ab, die in einem Datensatz enthaltenen Informationen möglichst übersichtlich darzustellen, so dass ‚das Wesentliche‘ schnell erkennbar wird.“¹¹⁷

Eine Deskription ist unter Verwendung von graphischen und / oder numerischen Darstellungsverfahren möglich, wobei darauf geachtet werden sollte, sich lediglich auf das Spektrum der tatsächlich erhobenen Daten zu beziehen. Mit Hilfe der dadurch erfolgenden Informationsverdichtung bleibt einerseits die Datenbasis erhalten, andererseits werden Redundanzen ausgefiltert. Daten, die für die Fragestellung nicht relevant sind, können aus der ursprünglichen Datenbasis – im Rahmen einer Informationsreduktion – entfernt werden.

2.7.1 Population und Grundgesamtheit

Kromrey definiert den Begriff „Grundgesamtheit“ folgendermaßen:

„Unter Grundgesamtheit ist diejenige Menge von Individuen, Fällen, Ereignissen zu verstehen, auf die sich die Aussagen der Untersuchung beziehen sollen und die im Hinblick auf die Fragestellung und die Operationalisierung vorher eindeutig abgegrenzt werden muss. Genau genommen handelt es sich hierbei um die angestrebte Grundgesamtheit“¹¹⁸

Zwischen einer angestrebten Grundgesamtheit und einer Auswahl- Grundgesamtheit sollte dabei differenziert werden.

Bezieht sich die Datenerhebung auf die gesamte Population oder Grundgesamtheit, wird von einer Voll- oder Totalerhebung gesprochen. Erstreckt sich die Datenerhebung lediglich auf einen Ausschnitt aus der Grundgesamtheit, handelt es sich um eine Teilerhebung, bei der Stichproben untersucht werden.

Im Fall der Evaluation von ProfIL TT werden alle Teilnehmer von PROFiL-TT in die Untersuchung miteinbezogen, die an der Vorher-, Nachher- und Retrospektionsbefragung teilgenommen haben. Wir sprechen in diesem

¹¹⁷ Kromrey H (2006). 420

¹¹⁸ Ibidem; 269

Fall nicht von einer Stichprobe, sondern von der Untersuchung einer Population oder Grundgesamtheit. Die Evaluation von PROFIL-TT ist somit eine Vollerhebung.

2.7.2 Quantile und Median¹¹⁹

Quantile sind lagetypische Maße zur Bestimmung von Verteilungen in Ordinal- oder Rangdaten.

Die Messwerte werden dabei aufsteigend sortiert und in gleich große Abschnitte unterteilt.

Besonderst häufig wird die Unterteilung in vier gleiche Abschnitte (Quartile) gewählt. Das 1. Quartil ist dabei der gemessene Wert, „der die Reihe der geordneten Merkmalswerte im Verhältnis 1:3 teilt“ (Kromrey 2006, 437). Das 2. Quartil – auch Zentralwert oder Median – unterteilt die gemessenen Werte im Verhältnis 2:2 und liefert somit den Wert, der in der Mitte der aufsteigend geordneten Reihe von Beobachtungswerten liegt. Das 3. Quartil unterteilt diese Reihe im Verhältnis 3:1

Quartile werden folgendermaßen bestimmt:

$$Q_1 = x_k \text{ mit } \frac{n}{4} < k < \frac{n}{4} + 1$$

$$Q_2 = x_k \text{ mit } \frac{n}{2} < k < \frac{n}{2} + 1$$

$$Q_3 = x_k \text{ mit } \frac{3n}{4} < k < \frac{3n}{4} + 1$$

„Das 1. Quartil (Q_1) ist derjenige Wert, in einer geordneten Reihe von Beobachtungswerten, der von einem Viertel der gemessenen Werte nicht über- und von dreiviertel der gemessenen Werte nicht unterschritten wird. Analoges gilt für das 2. und 3. Quartil.“ (Kromrey 2006, 438)

K ist dabei eine natürliche Zahl. Sollten $\frac{n}{4}$ und $\frac{n}{4} + 1$ ebenfalls natürliche Zahlen ergeben, wird für Q_1 ein Wert zwischen $x_{n/4}$ und $x_{(n/4+1)}$ sowie für Q_3 ein Wert zwischen $x_{3n/4}$ und $x_{(3n/4+1)}$ gewählt.

Als Mittelwert hat der Median folgende Eigenschaften:

- 1) Die extremen Werte einer Verteilung werden nicht berücksichtigt.
- 2) Bei einer ungeraden Anzahl von Beobachtungswerten stellt er immer einen real in der Verteilung vorkommenden Wert dar, bei einer geraden Anzahl von Beobachtungswerten nimmt der Median einen in der Verteilung nicht vorkommenden, hypothetischen Wert an, der die Verteilung exakt mittig halbiert.
- 3) Ist eine Verteilung nicht in der Mitte konzentriert, kann der Median ein verzerrtes Bild der Werte liefern.
- 4) Der Median ist der „typische“ Wert innerhalb einer Reihe von Rangdaten.

¹¹⁹ Vgl. ibidem; 437 ff.

2.7.3 Quartilabstand¹²⁰

Der Quartilabstand bezeichnet ein lagetypisches Streuungsmaß und wird durch die Differenz zwischen 1. und 3. Quartil definiert.

$$Q = Q_3 - Q_1$$

Der Quartilabstand spiegelt das Werte- Intervall wieder, in dem 50 % aller beobachteten Fälle liegen. Dabei werden die ersten (niedrigsten) und die letzten (höchsten) 25% der Werte „abgeschnitten“.

„Der Quartilabstand ist also (...) die Länge des Intervalls, das die 50% ‚mittleren Fälle‘ einer Beobachtungsreihe umfasst.“¹²¹

In der Regel setzt der Quartilabstand intervallskalierte Daten voraus. Er kann allerdings auch bei ordinalskalierten Daten herangezogen werden. Dabei wird die Differenz der Positionswerte Q_1 und Q_3 herangezogen.

In Kombination mit dem Median kann der Quartilabstand wichtige Informationen über eine Verteilung geben. So kann mittels Median und Quartilabstand unter anderem darauf geschlossen werden, welche Ausrichtung – links-schief, rechtsschief, symmetrisch – eine Verteilung aufweist.

2.7.4 Interpretation ordinal- und intervallskalierter Daten

Die Abgrenzung von ordinal- und intervallskalierter¹²² Daten kann nach Kromrey auf unterschiedliche Weise interpretiert werden.

„Ein (...) häufig auftretender Streitfall (...), ist die Behandlung ordinalskalierter Werte als ‚annähernd intervallskaliert‘. Einerseits fordern nämlich leistungsfähige statistische Analyseverfahren häufig metrische (d.h. mindestens intervallskalierte) Daten, andererseits erreichen sozialwissenschaftliche Variablen meist nur ordinales Messniveau.“¹²³

¹²⁰ Vgl. ibidem; 447 f.

¹²¹ Ibidem; 447

¹²² Ein ordinales Messniveau zeichnet sich generell durch die Ausprägung der Merkmalsdimensionen hinsichtlich ihrer Stärke und Intensität aus, wobei die Relation zwischen den Ziffern der Messkala gleich oder ungleich sein kann. Ordinalskalierte Daten werden grundsätzlich auf Basis einer Rangordnung (größer oder kleiner) empirisch interpretiert. Bei intervallskalierten Daten ist darüber hinaus eine Differenzierung der Merkmalsdimensionen hinsichtlich ihres genauen Grads an Stärke oder Intensität möglich (z.B. ist y ein um einen bestimmten Wert intensiver ausgeprägt als z). Wesentlich in der Unterscheidung von ordinal- und intervallskalierter Daten ist die Tatsache, dass bei intervallskalierten Daten die Beziehung, in der diese Daten zueinander stehen – also die Abstände – bekannt ist. (vgl. Kromrey 2006, 242 ff.)

¹²³ Ibidem; 418

Zur Pragmatik raten Opp und Schmidt (1976) für den Fall, dass die Resultate die aus dem „Rechnen“ mit ordinalen Daten stammen, empirisch sinnvoll erscheinen – wenn zum Beispiel die ordinalen Variablen ‚allzu weit‘ von quantitativen Variablen ‚entfernt‘ sind.¹²⁴

Im Fall einer derartigen Auslegung der Datenbasis, könnten die ordinalskalierte Daten vor dem Hintergrund einer „subjektiven Intervallskalierung“ interpretiert werden. Die Ordinalskala erscheint dabei dem Teilnehmer als Skala mit gleichen Intervallschritten und einem – gedachten – Nullpunkt. Im vorliegenden Fall könnte somit eine nicht vorhandene Akzeptanz auf den Nullpunkt einer Skala projiziert werden, volle Akzeptanz auf den Höchstwert. Die Abstufung der Akzeptanzschritte könnte analog zu einer – subjektiv empfunden, gleichmäßigen – Abstufung von Intervallen betrachtet werden.

2.7.5 Statistische Mortalität

Unter „statistischer Mortalität“ wird allgemein – vor allem bei Langzeitstudien – die Dezimierung der Probandenanzahl im Vergleich Interventionsbeginn und Interventionsabschluss verstanden.

Der Begriff „Mortalität“ – ursprünglich demographisch genutzt – zeigt eine Relation zwischen der Anzahl sterbender Personen, zur Gesamtanzahl der Personen in der Population innerhalb eines bestimmten Zeitabschnitts auf. Im medizinischen Umfeld wurde mit „statistischer Mortalität“ das „Wegsterben“ der Probanden im Lauf einer langfristigen Behandlung bezeichnet.¹²⁵

Auf die Sozialwissenschaften übertragen wird darunter das Phänomen verstanden, dass am Anfang einer Intervention oder Maßnahme die Teilnehmeranzahl stets größer ist, als gegen Ende.

Im Rahmen dieser Untersuchung spielt die statistische Mortalität eine unwesentliche Rolle, da von vorneherein bei der Evaluation von PROFIL-TT feststand, dass nur durchgängige Datensätze die Zielkriterien erfüllen und berücksichtigt werden können (die die Population besteht nur Teilnehmer, die beide Teile von PROFIL-TT – Sprachkompetenz und Lehrkompetenz – besucht, sowie an allen drei Stufen der Evaluation teilgenommen haben; siehe Kapitel 2.7.1).

Bei unvollständigen Datensätzen, z.B. wenn eine Stufe der Evaluation ausgelassen worden ist, treten Störvariablen auf, die zu einer Verzerrung der Ergebnisse führen könnten.

2.8 Grundsätzliche Evaluationsvorgehensweisen

2.8.1 Formative Evaluation nach Scriven¹²⁶

Formative Evaluationen finden interventionsbegleitend statt und haben das Ziel, die evaluierte Maßnahme noch während der Durchführung zu verbessern. Diese laufende Verbesserung kann nur umgesetzt werden, wenn die

¹²⁴ Vgl. Opp K-D, Schmidt P (1976). 35

¹²⁵ Vgl. Wottawa H, Thierau H (2003). 125

¹²⁶ Vgl. Scriven M (1991). 168 f.

Möglichkeit besteht, während der noch nicht abgeschlossenen Intervention Änderungen einzusteuern und durchzusetzen. Aus diesem Grund werden formative Evaluationen zu mehreren Zeitpunkten während einer Intervention durchgeführt, bis die evaluierte Maßnahme den Vorstellungen der Projektverantwortlichen gerecht wird.

Eine Sonderform der formativen Evaluation ist die *preformative Evaluation*, in deren Rahmen das Maßnahmenkonzept untersucht wird. Bei solchen preformativen Evaluationen werden häufig strukturelle und konzeptionelle Fehler und Unstimmigkeiten innerhalb der Maßnahme aufgedeckt, die bei der späteren Implementierung zu großen Problemen führen können. Mittels preformativer Designs können derartige Schwächen erkannt und beseitigt werden. Trotz dieses offensichtlichen Nutzens werden derartige Evaluationen selten verwendet.

Eine weitere, sehr nützliche Art formativer Evaluationen – Scriven beschreibt sie als *die nützlichste Art formativer Evaluationen* – ist die „early-warning summative“. Die Vorgehensweise bei dieser Evaluationsart ist eigentlich summativ (siehe unten), mit dem Unterschied, dass hier kein bereits fertig implementiertes Projekt, sondern eine frühe Entwicklungsstufe – wobei die Entwicklung in diesem Fall noch nicht abgeschlossen sein sollte – evaluiert wird. Die summative Untersuchung anhand einer „Testversion“ einer Intervention in einem kleinen Rahmen liefert wertvolle Erkenntnisse über Stärken und Schwächen, die bei der Implementierung der fertig entwickelten Maßnahme sehr hilfreich sein können.

Robert Stake beschreibt den Unterschied zwischen formativer und summativer Evaluation mit folgenden Worten:

„When the cook tastes the soup, that’s formative; when the guests taste the soup, that’s summative.“¹²⁷

2.8.2 Summative Evaluation nach Scriven¹²⁸

Eine summative Evaluation wird in der Regel nach Beendigung der Maßnahme („evaluand“) durchgeführt, beziehungsweise bei fortlaufenden Programmen nach deren Stabilisation. Den primären Nutzen aus den Evaluationsergebnissen bei summativen Designs ziehen Entscheidungsgremien, die die Ergebnisse als Hilfe oder Unterstreichung ihrer Meinung verwenden. Die summative Evaluation sollte nicht mit „outcome-evaluation“ verwechselt werden, da im Rahmen einer „outcome-evaluation“ lediglich die Effekte einer Intervention untersucht werden, allerdings nicht die Prozesse innerhalb der Intervention. Sie kann sowohl *global*, wie *analytical* sein (siehe Kapitel 2.9.21 Global Evaluation und Kapitel 2.9.20 Analytic / analytical Evaluation).

Scriven beschreibt die summative Evaluation mit folgenden Worten:

„Where a summative evaluation is done [...], the aim is [...] to report on it, not to report to it.“¹²⁹

Diese Aussage macht den final bewertenden Charakter der summativen Evaluation sehr deutlich.

¹²⁷ Stake R zit. in Scriven M (1991).169

¹²⁸ Scriven M (1991). 340

¹²⁹ Ibidem

Trotz der weiten Verbreitung und Etablierung von Evaluationen mit formativem oder summativem Charakter, ist die strikte Trennung der beiden Evaluationsvorgehensweisen nicht unumstritten, da diese begriffliche Polarisierung „vielerlei Unschärfen und konzeptionelle [...] Unklarheiten“ aufweist.¹³⁰

2.8.3 Vergleich wesentlicher Merkmale von formativer und summativer Evaluation¹³¹

Zielgruppe

Bei formativen Evaluationsansätzen besteht die betreffende Zielgruppe hauptsächlich aus Personen und Personengruppen, die im direkten Umfeld der evaluierten Maßnahme stehen. Die Programm- Entwickler und – Manager haben dabei vorwiegend das Interesse, die Maßnahme in ihrer Effizienz und Effektivität zu steigern, beziehungsweise zu optimieren. Die Programmdurchführenden sind in erster Linie an einer Optimierung der Interventionsinhalte, sowie an einer stimmigen Umsetzung interessiert.

Bei summativen Evaluationen, die in erster Linie das Gesamtergebnis einer Intervention betrachten, haben als Auftraggeber vorwiegend Förderer und Sponsoren großes Interesse. Hintergrund ist dabei meist die Vergewisserung, dass Fördermittel sinnvoll und zweckmäßig eingesetzt worden sind. Zur Zielgruppe der Sponsoren und Förderer zählt auch – vor allem bei staatlich geförderten (Sozial-) Projekten – die „interessierte Öffentlichkeit“, die über die entsprechende Verwendung von Steuergeldern informiert werden will, außerdem Politiker, die über die ihnen anvertrauten Mittel Rechenschaft abgeben müssen.

Evaluation in einem solchen Umfeld hat zumeist einen rechtfertigenden und bilanzierenden Charakter.

Ausrichtung der Datenerhebung

Die Datenerhebung ist bei formativen Ansätzen stark analytisch ausgerichtet. Die Maßnahmenziele sowie die Maßnahme selbst werden auf der Mikroebene detailliert nach Problemen, Unstimmigkeiten und möglichen Verbesserungen untersucht. Anpassungen werden dementsprechend auch auf der Mikroebene eingesteuert.

Die Vorgehensweise bei summativen Datenerhebungen wird durch einen globalen Ansatz bestimmt. Die Maßnahme und die Effekte der Maßnahme werden als Ganzes betrachtet und auf der Makroebene interpretiert. Ein Schwerpunkt bei summativen Evaluationen liegt auf der Dokumentation der Evaluationsdurchführung und auf dem Gesamtergebnis. Die summative Evaluation ist – bezogen auf das Informationsraster – grobkörniger, einzelne Effekte werden nicht so stark gewichtet, der Fokus wird auf die Summe aller Einzelheiten gelegt.

Rollenverständnis von Evaluator und sonstigen Beteiligten

Programmentwickler und Programm- Ausführende werden in beiden Evaluationsformen aktiv eingebunden. Bei formativen Evaluationen besteht ihre Aufgabe darin, die Intervention durch die Inputs der Evaluatoren anzupassen und zu verbessern, die Maßnahme mit zu „formen“. Dieser Prozess setzt ein Maß an Flexibilität und Enga-

¹³⁰ Vgl. z.B. die Diskussion zwischen Scriven, Chen, Patton und Wholey; zit. in Hense J (2007)

¹³¹ Vgl. Wottawa H, Thierau H (2003). 64

gement bei den Projektmitarbeitern voraus, sie müssen sich der aktuellen Situation anpassen – und das immer wieder von neuem.

Bei summativen Evaluationen besteht die Aufgabe der Programmentwickler und –Ausführenden in erster Linie im Beschaffen von Daten, die dann für die zusammenfassende Evaluation am Ende ausgewertet und interpretiert werden können.

Die Rolle des Evaluators im Rahmen einer formativen Evaluation sollte – wie die der Projektverantwortlichen auch – sehr flexibel aufgefasst werden.

Durch die rekursiven Datenerhebungen im formativen Evaluationsprozess erfährt der Evaluator häufig neue Entwicklungen und Trends der untersuchten Maßnahme und muss darauf reagieren. Diese Flexibilität fordert eine enge Zusammenarbeit mit den Projektverantwortlichen, da sie die Anpassungen und Änderungen an der Intervention einsteuern und umsetzen.

Im Rahmen einer summativen Evaluation ist der Evaluator von den Projektverantwortlichen und –Mitarbeitern unabhängig. Dadurch dass die Daten am Ende der Intervention gesammelt aufbereitet und interpretiert werden, steht ein größerer zeitlicher Rahmen zur Verfügung. Die Planung der Evaluation wird dadurch wesentlich erleichtert. Aus den gesammelten Daten werden am Ende ein Resümee gezogen und Empfehlungen ausgesprochen.

Berichterstattung / Kommunikation der Evaluationsergebnisse

Die Berichtlegung ist im Rahmen formativer Evaluationen ein Prozess, der in der Regel informell und häufig auch auf verbalem („kurzem“) Weg stattfindet. So werden Zwischenergebnisse teilweise in Diskussionen vor den versammelten Prozessteilnehmern und den Beteiligten vorgestellt. Diese informelle Interaktion ist notwendig, um die Evaluationsergebnisse auf unbürokratischem Weg bei den Betroffenen einsteuern zu können. Die Berichterstattung findet bei formativen Evaluationen während des gesamten Evaluationsprozesses in vorher festgelegten Abständen, oder nach Bedarf statt. Schwerpunkte bei der formativen Berichterstattung werden in der Regel auf die Beziehungen zwischen Prozesselementen, Kontext, Prozess und Ergebnis gelegt.

Bei summativen Evaluationen findet die Berichterstattung in einem formalen Rahmen statt. Die Ergebnisse werden mittels exakter Berichte formuliert und an die zuständigen Projektmitglieder und –verantwortlichen kommuniziert. Diese Berichterstattung findet in der Regel erst nach Beendigung der untersuchten Intervention statt; ein regelmäßiges Reporting wie bei formativen Evaluationen findet nicht statt. Inhaltlich liegen die Schwerpunkte bei summativen Untersuchungen nicht im lokal- analytischen Bereich, sondern betreffen meist globale Implikationen für administrative Stellen, Management und Politik.

Rolle des Evaluators¹³²

Generell könnten die Rolle eines wissenschaftlichen Forschers und die eines Evaluators – hypothetisch und überspitzt formuliert – mittels zwei gegensätzlichen Polungen dargestellt werden. Der wissenschaftliche Forscher

¹³² Vgl. ibidem; 37

ist unabhängig und kann ungehindert, vollkommen selbstständig und frei nach der wissenschaftlich-empirischen Wahrheit suchen, wobei er keinerlei externen Zwängen oder Druck ausgesetzt ist.

Die Rolle des Evaluators hingegen ist in der Regel auftragsabhängig wodurch auch eine Abhängigkeit gegenüber dem Auftraggeber einer Evaluation besteht. Der Evaluator sollte sich an die Zielvorgaben und Schwerpunkte, die von seinen Auftraggebern gefordert werden, halten, da das letztendliche Entscheidungspotential bei ihnen liegt. Fachliche und organisatorische Kompetenz werden natürlich auch für die Rolle des Evaluators vorausgesetzt.

Der Evaluator sollte sich im Rahmen seiner Arbeit an eine Reihe von Regeln halten, um seine Aufgabe möglichst konfliktfrei ausüben zu können.

Eine unabhängige Forschung ist aus diesen Gründen häufig schwer zu realisieren.

Eine weitere Einschränkung, mit der sich ein Evaluator im Arbeitsprozess oftmals konfrontiert sieht, betrifft – vor allem bei auftragsgebundenen Evaluationen – die Kommunikationsfreiheit der Evaluationsergebnisse. Die Ergebnisse dürfen unter Umständen nicht publiziert werden, da diese Publikation – vor allem wenn die Evaluationsergebnisse eine negative Einfärbung aufweisen – den Interessen der Auftraggeber schaden könnten.¹³³

Generell muss der Evaluator pragmatisch handeln, er befindet sich häufig auf dem schmalen Grat zwischen empirischer Wissenschaft und praktischer Anwendungsorientierung. Der Evaluator muss „das Problem der ‚unwissenschaftlichen‘ Bewertung und die damit zusammenhängende Maximierung von Nutzen- Aspekten in seiner Arbeit unmittelbar aufgreifen“ (Wottawa, Thierau 2003, 37). Durch diese Verhaltensweise handelt er allerdings konträr zum Grundsatz der „Wertfreiheit“ der Wissenschaft.

Mit der Evaluation sollte das Bewusstsein einhergehen, dass die Tätigkeit des Evaluators nicht im engen Sinne wissenschaftlich, sondern meistens von den Auftraggebern abhängig ist. Das Ziel besteht deswegen darin, nicht absoluten Wahrheiten auf den Grund zu gehen, sondern bei der Gestaltung und Modellierung von gesellschaftlichen Programmen, Prozessen, etc. hilfreich zu unterstützen. Die Akzeptanz und Annahme dieser Unterstützung liegt dabei wieder bei den Auftraggebern, die sich die Frage stellen müssen, ob die Maßnahmen aus den Ergebnissen einer Evaluation umgesetzt werden sollen (bzw. können), oder nicht.

2.8.4 Interne Evaluation – Selbstevaluation

Interne Evaluationen werden von derselben Organisation vorgenommen, die auch das zu evaluierende Programm durchführt.¹³⁴

¹³³ Bei Lehrveranstaltungsevaluationen spielt dieser Punkt im Rahmen von Universitäts- oder Schul- Benchmarks eine wichtige Rolle, da mit einem schlechten Abschneiden im Rahmen einer Evaluation das Ansehen einer Bildungseinrichtung sinken kann.

¹³⁴ Handelt es sich dabei nicht nur um die gleiche Organisation, sondern kommen die Evaluatoren auch aus demselben Bereich oder derselben Abteilung, in der das zu evaluierende Programm stattfindet, handelt es sich um eine Selbstevaluation, ansonsten wird von einer „In-house-Evaluation“ gesprochen.

Die Vorteile der internen Evaluation bestehen in erster Linie in der Fachkenntnis des durchführenden Evaluators. Dieses „konkrete Situationswissen“¹³⁵ kann genutzt werden um Probleme und Schwierigkeiten im Evaluationsverlauf frühzeitig identifizieren und lösen zu können. Durch die organisatorische und persönliche Nähe des Evaluators zu den Programmverantwortlichen und Auftraggeber, besteht eine schnelle und meist unbürokratische Implementierungsmöglichkeit von formativen Änderungen. Diese inhaltliche und persönliche Nähe zum Programmumfeld und der Administration hat allerdings auch den Nachteil, die Evaluation mit lediglich geringem Abstand zum evaluierten Programm durchführen zu können. Diese oftmals mangelnde Objektivität kann verhindern, den Evaluationsgegenstand „global“ zu betrachten und die Untersuchung aufgrund möglicher persönlicher Konsequenzen eher rudimentär durchzuführen.¹³⁶

2.8.5 Externe Evaluation – Fremdevaluation

Die Vorteile der internen Evaluation sind gleichsam die Nachteile der externen, ebenso wie ihre Nachteile sich andererseits als Vorteile der externen Evaluation herausstellen.

Dadurch, dass die Evaluatoren nicht der Organisation der Mittelgeber oder der Durchführenden angehören, besteht in dieser Form grundsätzlich ein größerer Handlungsspielraum für die Evaluatoren. Diese Unabhängigkeit schafft die Voraussetzungen für eine Objektivität dem Untersuchungsgegenstand gegenüber. Damit sind allerdings häufig Schwierigkeiten bei der Informationsbeschaffung gekoppelt. Zum Beispiel kann mangelnde Akzeptanz auf der Seite der Evaluierten zu Abwehrreaktionen den Evaluatoren gegenüber führen.

Grundsätzlich sollte beim Hinzuziehen externer Evaluatoren davon ausgegangen werden, dass die Evaluation durch die vorhandene Methoden- und Instrumentenkompetenz und die Erfahrung in der Konzeption und Implementierung von Evaluationen professionell durchgeführt wird.¹³⁷

2.8.6 Evaluationsvorgehensweise bei PROFIL-TT

Wesentlich beim Design von PROFIL-TT ist eine Mischung aus formativer und summativer Evaluationsvorgehensweise, wobei der Schwerpunkt auf summativen Methoden liegt, da die Erhebung einerseits projektbegleitend – in den drei Stufen der Evaluation –, das Resümee allerdings hauptsächlich bilanzierend war, und nach Projektende zusammengefasst wurde. Ein direkter Eingriff in das Projekt war aufgrund der vornherein festgelegten Konzeption nur in geringem Maße möglich. Während der Evaluation wurde den Projektmitarbeitern sowie der Projektadministration regelmäßig Rückmeldung über die Evaluationsergebnisse gegeben, abschließend fand ein finales Reporting über das Gesamtprojekt statt.

Die Evaluation von PROFIL-TT war einerseits intern vom Institut für Psycholinguistik initiiert, wurde allerdings nicht direkt am Institut von Mitarbeitern, sondern im Rahmen dieser Arbeit durchgeführt.

¹³⁵ Stockmann R (2006). 263

¹³⁶ Vgl. ibidem; 74

¹³⁷ Vgl. ibidem

Das eigentliche Projekt, das angewendete Instrumentarium sowie die Ergebnisse werden im Kapitel 3 dieser Arbeit detailliert beschrieben.

2.9 Evaluationsmodelle

2.9.1 Übersicht

Evaluationen können anhand einer Vielzahl unterschiedlicher Kriterien charakterisiert werden. Aufgrund dieser Facettenfülle ist es bis jetzt noch nicht gelungen, ein Evaluationsmodell zu formulieren, das den Charakter einer eigenständigen Evaluationstheorie besitzt.¹³⁸

Die Unterscheidung der Modelle betrifft unterschiedliche Kriterien, die im Rahmen einer Evaluation auftreten, wesentlich sind und nach denen eine Evaluation durchgeführt werden kann.¹³⁹

Grundsätzlich wird dabei berücksichtigt, wie groß die Reichweite einer Evaluation sein sollte, ob einzelne Aspekte, Abschnitte eines Programms oder das komplette Programm im Fokus der Untersuchung stehen sollen.

Für die praktische Anwendung von Evaluationen hat diese Vielfalt zur Folge, dass häufig unterschiedliche Modelle miteinander vermischt werden, bis wiederum neue Modelle entstehen, die in spezifischen Kontexten den betreffenden Gegenstand genau beschreiben und bewerten können.¹⁴⁰ Die Tendenz, ein generell gültiges, globales Evaluationsmodell zu formulieren besteht allerdings weiterhin (z.B. der Ansatz von Worthen und Sanders zur managementorientierten Evaluation).¹⁴¹

Hense beschreibt die Vielzahl der Evaluationsmodelle als „Bündel deskriptiver, präskriptiver und/oder normativer Aussagen zu Bedingungen, Wirkungen und/oder Umsetzungsvarianten von Evaluation, die optimalerweise auf theoretischen und empirischen Ergebnissen der Evaluationsforschung basieren.“¹⁴²

Im folgenden Abschnitt sollen die wesentlichen Modelle und deren Eigenschaften dargestellt und erläutert werden.

¹³⁸ Vgl. Shadish, Cook, Leviton (1991); Alkin (2004); zit. in Hense J (2007)

¹³⁹ Die Charakterisierung von Evaluationsmodellen kann z.B. anhand folgender Kriterien getroffen werden: Design, Methodenauswahl, Zielsetzungen, Generalisierbarkeit von Ergebnissen versus Einzelfallanalysen, Verhältnis zwischen Evaluator und Auftraggeber, interne versus externe Evaluation, Selbst- versus Fremdevaluation, Wissenschaftlichkeit der Evaluation gegenüber Anwendungsorientierung und Verwendbarkeit der Ergebnisse (vgl. Rindermann H (2001). 10 f.; Hense J (2007))

¹⁴⁰ Vgl. Beywl W (2006). Evaluationsmodelle und qualitative Methoden. in Flick U (Hrsg.).(2006). 92 – 116

¹⁴¹ Vgl. Worthen BR / Sanders JR / Fitzpatrick JL (1997). 102

¹⁴² Hense J (2007)

2.9.2 Effizienzkriterienorientiert Evaluation (Rossi, Freeman, Lipsey)¹⁴³

Bei der *effizienzkriterienorientierten Evaluation* wird meist von externen Evaluatoren das Kosten- Nutzen- Verhältnis einer Veranstaltung oder Intervention untersucht. Der Aufwand, der für die Konzeption, Durchführung und Nachbereitung einer Maßnahme notwendig ist, wird dabei der Effektivität und Wirkung gegenübergestellt, die von dieser Maßnahme herrühren. Voraussetzung für eine effizienzkriterienorientierte Evaluation ist eine umfangreiche Wirkungsanalyse, die in erster Linie unter Verwendung von experimentellen Designs durchgeführt werden kann.¹⁴⁴

Bei der Wirksamkeitsprüfung von Interventionen stellt sich grundsätzlich die Frage, wie empirische Informationen über die Wirkungsweise einer Intervention gesammelt werden können und wie überprüft werden kann, ob eine solche Wirkung überhaupt vorhanden ist. Anhaltspunkte hierfür sind in der Regel Programmziele, anhand derer festgestellt werden kann, ob oder inwiefern eine Intervention zielgerichtet und somit effizient implementiert und durchgeführt wurde und ob dadurch eine Wirkung oder sogar einen Nutzen erzielt werden konnte.

Die Daten, die aus einer solchen Erhebung resultieren, entsprechen in der Regel wissenschaftlichen Untersuchungskriterien und weisen demzufolge eine hohe Generalisierbarkeit auf.

2.9.3 Versuchsplanorientierte Evaluation (Hager, Patry, Brezing)¹⁴⁵

Auch im Rahmen von *versuchsplanorientierten Evaluationen* wird auf die strenge Einhaltung methodischer Standards geachtet. Der Untersuchungsplan wird vor der Evaluation ausgearbeitet und während der gesamten Dauer eingehalten. Die Untersuchung betrifft – ähnlich wie bei der effizienzkriterienorientierten Evaluation – Wirkungen und Wirkfaktoren von lokal eingesetzten Interventionen. Die Einhaltung der Gütekriterien interne und externe Validität, Reliabilität und Präzision hat bei diesem Modell oberste Priorität.

2.9.4 Programmziel- / lernzielgesteuerte Evaluation (objective-based study)¹⁴⁶

In *programm-* oder *lernzielgesteuerten* Modellen werden anhand von meist experimentellen oder zumindest quasi-experimentellen Vorgehensweisen die angestrebten Programmziele mit den tatsächlichen Resultaten einer Intervention verglichen. Ausschlaggebend für einen solchen Vergleich ist die Formulierung der Programm- und Lernziele einer Intervention vor Beginn der Evaluation.

¹⁴³ Vgl. Rossi PH, Freeman HE, Lipsey MW (2004). 331 – 341

¹⁴⁴ Vgl. Stockmann R (2006). 101 – 106

¹⁴⁵ Vgl. Hager, Patry, Brezing (2000). Zusammenfassung. 47 – 61

¹⁴⁶ Vgl. Beywl W (2004). Artikel *Programmzielgesteuerte Evaluation*

2.9.5 Theoriegeleitete Evaluation (Chen)¹⁴⁷

Die Basis für *theoriegeleitete Evaluationen* beruht auf der Formulierung und Ausarbeitung eines logischen Modells vor Programmstart, das der Evaluation zugrunde liegt. Von diesem Modell abhängig ist die Konzeption der Evaluation, die Datenerhebung, -auswertung und -interpretation – somit die gesamte Steuerung.

Der Hintergrund dieses Modells besteht in der Annahme, dass es sinnvoller ist, die Arbeits- und Wirkungsweise einer Intervention zu verstehen, als zu untersuchen, ob ein Programm in nichtgeneralisierbaren und hochspezifischen Kontexten überhaupt einen Effekt erzielen kann. Die Grundlage für theoriegeleitete Evaluationen stellen „black-box evaluations“ (siehe Kapitel 2.9.18) dar.

2.9.6 Gegnerschaftsgesteuerte Evaluation (adversary Evaluation)¹⁴⁸

Die Besonderheit bei *gegnerschaftsgesteuerten Evaluationen* besteht darin, dass zwei Evaluationsteams, die bezüglich der zu evaluierenden Maßnahme unterschiedliche Auffassungen vertreten, dieselbe Maßnahme unabhängig voneinander untersuchen. Die beiden Evaluationsteams stellen ihre Ergebnisse nach Beendigung der Evaluation einer Kommission vor, die die Ergebnisse unter Miteinbeziehung aller Vor- und Nachteile bewertet, ein Resümee daraus zieht und eine Empfehlung für die weitere Vorgehensweise bezogen auf die Intervention ausspricht.

Dieser Modelltyp wird meist dann verwendet, wenn Vor- und Nachteile einer Intervention explizit herausgestellt werden sollen.

2.9.7 Entscheidungsgesteuerte Evaluation (Stufflebeam)¹⁴⁹

Entscheidungsorientierte Evaluationsansätze zielen in erster Linie auf die Beschreibung, Abgrenzung und Anwendung deskriptiver Informationen, um den Wert oder die Leistung eines Programms einschätzen oder beurteilen zu können. Einer entscheidungsorientierten Evaluation liegen meist zur untersuchten Intervention eine oder mehrere Alternativen vor. Um letztendlich eine Entscheidung für oder gegen das untersuchte Programm treffen zu können, sind detaillierte Aussagen über Ziele, Design, Implementierung und Ergebnisse dieser Intervention notwendig.

2.9.8 Beteiligteninteressengesteuerte Evaluation

Die *beteiligteninteressengesteuerte Evaluation* hat das Interesse, alle Anliegen und Konfliktthemen der beteiligten Personengruppen zu berücksichtigen. Die Praxisorientierung steht bei diesem Modell im Vordergrund, wobei

¹⁴⁷ Vgl. Chen HT (1990). 37 – 55 / Patton MQ (2002). 481 – 487

¹⁴⁸ Vgl. Owens T (1973) / Wolf RL (1979). 19 – 28

¹⁴⁹ Vgl. Stufflebeam DL, Shinkfield AJ (1985)

größtenteils qualitative Untersuchungsmethoden eingesetzt werden, um die Interessen möglichst aller involvierten Gruppen einzubeziehen.

Beteiligteninteressengesteuerte Evaluationen können in verschiedene Varianten differenziert werden. Naturalistische¹⁵⁰ und responsive Evaluationen sowie Fallstudieneinsätze¹⁵¹ fallen unter diese Kategorie.

Naturalistische Evaluation

Williams (1986) beschreibt die naturalistische Evaluation folgendermaßen:

„The qualitative perspective, defensibly obtained, can radically improve many evaluation efforts. More and more evaluators have begun to explore the naturalistic approach to see if and how it may be applied effectively in a variety of settings. Some have attempted evaluations that are exclusively naturalistic (including no quantitative methods) while others have combined naturalistic components (such as participant observations and repeated informal interviews) with more traditional tests, questionnaires, and structured interviews to create mixed-method studies“ (Williams DD (1986, 1)¹⁵²

Darin spielt er auf den starken Bezug der naturalistischen Evaluation auf die qualitative Evaluation an – er setzt sie nahezu gleich. Ferner können nach dieser Definition die Eigenschaften von naturalistischen Evaluationen an der Wirksamkeit ihrer eingesetzten Methoden gemessen werden.

Responsive Evaluation

Das responsive Evaluationsmodell zeichnet sich in erster Linie durch seine Flexibilität in der Durchführung aus. Das ursprünglich von Stake (1975) entwickelte Modell orientiert sich an den Interessen der Beteiligten und Involvierten, ohne einem vorher festgelegten Schema zu folgen. Das Modell reagiert flexibel auf die Anforderungen und Bedürfnisse der Zielgruppen.¹⁵³

2.9.9 Qualitätsentwickelnde Evaluation¹⁵⁴

Die qualitätsentwickelnde Evaluation ist ähnlich wie das Total Quality Management (TQM) ein Modell zu Qualitätsverbesserung und -sicherung.¹⁵⁵ Dieses Modell soll das Qualitätsbewusstsein und die Qualitätsentwicklung innerhalb einer Intervention vorantreiben und sichern. Die Evaluation in diesem Rahmen hat vor allem beratenden Charakter und soll das Qualitätsmanagement, das bei diesem Modell wesentlich ist, unterstützen.

¹⁵⁰ Vgl. Guba EG, Lincoln YS (1981)

¹⁵¹ Vgl. Stake RE (1995). 1 – 9

¹⁵² Williams DD (Hrsg.). (1986). 1

¹⁵³ Vgl. Stake RE (1975).

¹⁵⁴ Vgl. Rindermann H (2001). 11

¹⁵⁵ Vgl. Stockmann R (2006). 92 - 97

2.9.10 Nutzenorientierte Evaluation (Utilization-Focused Evaluation; Patton)¹⁵⁶

Die nutzenorientierte Evaluation ist ein generelles Evaluationsmodell, das den Praxisbezug und die Verwertungsorientierung der Evaluationsergebnisse fokussiert.

Bei diesem Modell empfiehlt Patton eine Orientierung an den Evaluationsstandards (siehe Kapitel 2.10), um quantitative und qualitative Methoden möglichst zielführend und ergebnisorientiert einsetzen und eine Verwertbarkeit der Ergebnisse sicherstellen zu können. Nutzenorientierte Evaluationen können formativ, summativ oder entwicklungsorientiert angelegt werden (siehe Kapitel 2.8.1 / 2.8.2), jeweils abhängig von der Zielsetzung der Evaluation.

2.9.11 CIPP-Modell (Stufflebeam)¹⁵⁷

Das CIPP-Modell – auch KIPP-Modell – (entwickelt von Daniel Stufflebeam) umfasst vier Evaluationsschritte und soll dazu beitragen, einen möglichst umfassenden Überblick über ein komplettes Programm – bezogen auf **K**ontext, **I**nputs, **P**rozess und das letztendliche **P**rodukt – vermitteln.

„Der Ansatz ist als umfassendes Rahmenmodell zur Steuerung der Evaluation unterschiedlichster Evaluationsgegenstände gedacht.“¹⁵⁸

Die einzelnen Elemente des CIPP-Modells beinhalten folgende Evaluationsabschnitte.

Das Umfeld einer Intervention, Maßnahme oder eines Programms wird mit dem Fokus auf die Rahmenbedingungen und zur Verfügung stehenden personellen und Sach-Mittel untersucht und bewertet.

Kontextevaluation

Der Sinn dieser Kontextevaluation besteht in einer Sensibilisierung der Programmverantwortlichen bezüglich der zunächst geplanten Programmziele und des -nutzens, sowie – im Falle einer späteren Umsetzung des geplanten Programms – in der Bewertung dieser Ziele.

Inputevaluation

Der nächste Schritt im Rahmen einer CIPP-Evaluation besteht in der Evaluation der Inputs eines Programms. Die Schwerpunkte bei dieser Inputevaluation liegen auf der Umsetzbarkeit, auf der Eruiierung alternativer Lösungsvorschläge und einer Abschätzung der Kosten- Nutzen- Planung.

Auch in diesem Prozessschritt dient die Evaluation in erster Linie der Beratung der Projektverantwortlichen bezüglich der zu verwendenden Ressourcen oder dem Einsatz alternativer Umsetzungsmöglichkeiten. Aufgabe

¹⁵⁶ Vgl. Patton MQ (1997). 378 – 379

¹⁵⁷ Vgl. Stufflebeam DL (1983). 117 – 141

¹⁵⁸ Hense J (2007)

des Evaluators darüber hinaus ist ein Vorschlag für die Vorgehensweise im Rahmen der Evaluation (Evaluationsplan).

Prozessevaluation

Die Prozessevaluation als Stufe des CIPP- Modells untersucht, überprüft und dokumentiert die Implementierung des Evaluationsplans, beziehungsweise der kompletten Maßnahme. Diese Untersuchung soll die Rechtfertigung einer Intervention vereinfachen und dabei helfen, „Effekte des Programms beurteilen und interpretieren zu können“.¹⁵⁹

Produktevaluation

Die eigentliche Evaluation des Produkts besteht im CIPP- Modell vor allem aus der Untersuchung der beabsichtigten und unbeabsichtigten Auswirkungen einer Maßnahme in kurz- und langfristiger Sicht. Wesentlich für die Auftraggeber und Programmverantwortlichen ist dabei der intendierte und nicht- intendierte Nutzen, der aus der Maßnahme für Teilnehmer und Beteiligte resultiert. Dieser Nutzen dient häufig als Argumentationshilfe zur Verteidigung oder Fortsetzung eines Programms.

2.9.12 Zielorientierte Evaluation

Die ziel-orientierte Evaluation vertritt den Ansatz, im Rahmen einer Evaluation die geplanten und intendierten Ziele herauszufinden, zu formulieren und klassifizieren. Nach Abschluss der Intervention und Datenerhebung wird dann festgestellt, ob oder inwiefern die empirischen Ergebnisse mit den zuvor definierten Ziele übereinstimmen.

Das Gegenmodell zu diesem Ansatz von Ralf Tyler¹⁶⁰ ist das Modell der goal-free evaluation von Scriven, der die zielorientierte Evaluation stets kritisiert hat.

2.9.13 Goal-free evaluation (Scriven)¹⁶¹

Die „Zielfreiheit“ dieses Evaluationsmodells bezieht sich im Rahmen einer Maßnahme darauf, dass zur Evaluation keine formulierten Programmziele – im Sinne eines Soll- Ist- Vergleichs – verwendet werden, sondern dass sich Evaluatoren ausschließlich mit den interventionsbedingten Effekten eines Programms beschäftigen, die auch tatsächlich beobachtet werden können. Die von den Programmverantwortlichen intendierten und nicht- intendierten Wirkungen einer Maßnahme fließen in den Evaluationsprozess nur insofern ein, wie sie auch tatsächlich als realer Effekt auftreten und beobachtet werden können.

¹⁵⁹ Ibidem

¹⁶⁰ Vgl. z.B. Tyler R (1949)

¹⁶¹ Vgl. Scriven M (1973). 319-328

Dieses Modell von Michael Scriven widerspricht dadurch vielen Evaluationsmodellen. Scriven geht in seiner Forderung der „Zielfreiheit“ so weit, dass er von jeglicher Interaktion mit Programmdurchführenden abrät.

Dieser häufig kritisierte, radikale Ansatz begründet sich aus einem Trend zu Beginn der 1970er Jahre, in dem Evaluation mit der Durchführung eines Soll- Ist- Vergleichs synonym verwendet wurde und soll Evaluatoren sensibilisieren, die Auswirkungen eines Programms global zu untersuchen und sich von den festgelegten Zielen nicht beeinflussen zu lassen.

2.9.14 Partizipative Evaluation¹⁶²

Die Zusammenarbeit zwischen Evaluator und Beteiligten ist in diesem Modell sehr eng. Im Rahmen von partizipativen Evaluationen lassen sich zwei unterschiedliche Ausprägungen unterscheiden. Die Differenzierung zwischen den beiden Ansätzen bezieht sich vor allem auf die unterschiedlichen Hintergründe der Kooperation zwischen Evaluator und Interessengruppen. Partizipative Evaluation wird zumeist mit zielgruppen-orientierter Evaluation synonym verwendet.

Practical participatory evaluation¹⁶³

In der *practical participatory evaluation* bezieht sich die Zusammenarbeit vor allem auf eine effizient gestaltete Anwendung der Evaluationsergebnisse. Durch eine stärkere Miteinbeziehung von entscheidungsbefugten Personengruppen in den Evaluationsprozess soll die Voraussetzung für eine einfachere und wirkungsorientierte Umsetzung der Evaluationsergebnisse geschaffen werden.

Transformative participatory evaluation¹⁶⁴

Bei dieser Form der partizipativen Evaluation soll die Teilhabe an der „Demokratisierung sozialen Wandels“¹⁶⁵ forciert und die Stakeholder in ihrer Position gestärkt werden.

2.9.15 Cluster-Evaluation (Haubrich)¹⁶⁶

Bei Cluster-Evaluationen wird das Evaluationsprojekt in mehrere, kleinere Projekte unterteilt, die sich anhand lokaler Vorgehensweisen oder verschiedener Teil- Projektziele unterscheiden. Diese Vorgehensweise wird meistens angewendet, wenn ein komplexer und umfangreicher Evaluationsgegenstand untersucht werden soll, dessen Ziele eine zu große Streuung aufweisen, um sie gesammelt verfolgen und bewerten zu können.

¹⁶² Vgl. Beywl W (2004). Artikel *Partizipative Evaluation*

¹⁶³ Vgl. Beywl W (1999)

¹⁶⁴ Vgl. ibidem

¹⁶⁵ Ibidem

¹⁶⁶ Vgl. Haubrich K (2001)

2.9.16 Deskriptive Evaluation – Explorative Evaluation – Hypothesenprüfende Evaluation¹⁶⁷

Deskriptive Evaluationen haben das Ziel, den untersuchten Gegenstand, die Zielgruppe und das Umfeld, in dem die Intervention stattfindet möglichst genau zu beschreiben.

Die reine Deskription als Evaluationsgrundlage wird allerdings vielfach kritisch aufgefasst, da es als fragwürdig gilt, ob ein Gegenstand oder eine Maßnahme nur aufgrund einer detaillierten Beschreibung eingeschätzt oder bewertet werden kann. Uhl (1999) gliedert an die deskriptive Evaluation, explorative und hypothesenprüfende Evaluationen an. Demnach besteht in der Deskription die unterste methodologische Stufe wissenschaftlichen Vorgehens. Die methodisch nächste Stufe – die explorative Evaluation – hat als Zielsetzung, neue Phänomene aufzudecken und Anstöße zur Entwicklung neuer Hypothesen zu geben. Die explorative Dimension gilt dabei als Zwischenstufe zur hypothesenprüfenden Forschung, die einheitlich orientiert ist und eine streng methodische Vorgehensweise fordert.

2.9.17 Askriptive Evaluation¹⁶⁸

Askriptive Evaluationen dienen dazu, einer Intervention oder einem Gegenstand eine Wertigkeit oder einen Nutzen zuzuteilen. Die Evaluationsergebnisse werden allerdings nicht zur Verbesserung oder Optimierung der Maßnahme verwendet.

Der Begriff der askriptiven Evaluation wurde von Scriven geprägt, der darin eine weitere Differenzierung von formativen und summativen Evaluationen sieht.

Scriven führt als Beschreibung für das Modell der askriptiven Evaluation die retrospektive Bewertung der Rekrutierung von Soldaten für den Vietnamkrieg an. Ein Modell, das einen solchen Evaluationszweck erfüllt, kann weder formativ sein – die Maßnahme ist bereits abgeschlossen¹⁶⁹ – noch können die Evaluationsergebnisse einerseits zu einer Fortführung der Intervention und andererseits zu einem Abbruch führen; das Modell ist somit weder formativ noch summativ und müsse deswegen askriptiv sein.¹⁷⁰

Hense geht auf dieses Beispiel ein und begründet dadurch die seiner Meinung nach vorhandene Schwäche des polarisierten Formativ/Summativ-Ansatzes.

¹⁶⁷ Vgl. Uhl A (1999)

¹⁶⁸ Hense J (2007)

¹⁶⁹ Anm. d. Verfassers: Unter Umständen könnte ein solcher Prozess auf aktuell Maßnahmen wie z.B. die aktuelle Auseinandersetzung im Irak oder in Afghanistan generalisiert werden, was einen formativen oder einen summativen Charakter rechtfertigen würde.

¹⁷⁰ Scriven M (2004)

*„Vor allem die Verwischung von intendierten Evaluationszwecken (vor der Evaluation) und der Nutzung von Evaluationsergebnissen (während oder nach der Evaluation) wird hier evident, aber auch die unklare Differenzierung zwischen Konzept und Durchführung einer Maßnahme“.*¹⁷¹

Durch eine mögliche Generalisierung der Evaluationsergebnisse auf aktuelle Regierungsmaßnahmen könnte sich der askriptive Ansatz von Scriven tendenziell auch in einen formativen wandeln.

2.9.18 Black-box Evaluation

Black-box Evaluationen bewerten Wirkungen und Effekte von Interventionen, ohne die interne Funktionsweise oder die genaue Beschaffenheit der Intervention zu untersuchen. Grundsätzlich wird die Intervention als abgeschlossenes System betrachtet, das nach außen beobachtbare Wirkungen als Resultat hat.¹⁷² Für die Umsetzung von Black-box Evaluationen werden meist experimentelle oder zumindest quasi-experimentelle Untersuchungsmethoden angewendet. Auf diesem Evaluationsmodell basiert die theorieorientierte Evaluation – als Gegenmodell, die White-Box Evaluation.

Ein wesentlicher Vorteil dieser Vorgehensweise besteht darin, dass sämtliche Effekte, die aus einer Intervention resultieren, untersucht werden. Die Schwäche besteht darin, dass nicht immer festgestellt werden kann, ob ein Effekt interventionsbedingt und intendiert ist, oder von nicht-intendierten Faktoren abhängt.

2.9.19 Experten-orientierte Evaluation

Die Beurteilung eines Programms im Rahmen einer experten-orientierten Evaluation hängt von der Einschätzung, dem Wissen und dem Urteilsvermögen von Spezialisten ab, die in dem untersuchten Gebiet eine starke Reputation besitzen. Dieses Modell hat einerseits den Vorteil, dass es mit einem relativ geringen Aufwand realisiert werden kann und einen guten Überblick, sowie eine genaue Einschätzung des untersuchten Gegenstands liefern kann. Andererseits jedoch kann die Beurteilung – basierend auf einzelnen Expertenmeinungen – auch stark subjektiv ausgelegt sein und wesentliche Aspekte von Interventionen bei der Evaluation vernachlässigen (vgl. *Gutachtenmodell* Stockmann 2006, 237).

2.9.20 Analytic / analytical Evaluation (Scriven)¹⁷³

Beim analytischen Evaluationsansatz wird das Untersuchungsobjekt – analog zur Cluster-Evaluation (siehe Kapitel 2.9.15) – systematisch in einzelne Teilbereiche untergliedert, die wiederum jeder für sich separat untersucht wird. Trifft man an dieser Stelle eine Evaluations-Differenzierung auf Micro- und Macroebene, kann man den analytischen Ansatz der Micro-Evaluation zuordnen. Generell kann man die Vorgehensweise in zwei Aspek-

¹⁷¹ Hense J (2007)

¹⁷² Vgl. Love AJ (2003)

¹⁷³ Vgl. Scriven M (1991). 54 f.

te unterteilen, Evaluation von Komponenten und Evaluation von Dimensionen – wobei zwischen diesen beiden Ansätzen auch Mischformen – zum Beispiel die Programmevaluation – existieren.

Im Rahmen der Komponentenevaluationen („component evaluation“) werden die einzelnen Bestandteile einer Maßnahme oder einer Intervention evaluiert; die Evaluation von Dimensionen („dimensional evaluation“) beschäftigt sich in erster Linie mit der Bewertung von „aspects“.

2.9.21 Global Evaluation (Scriven)¹⁷⁴

Der analytischen Evaluation steht die „global evaluation“ gegenüber. Im Rahmen einer globalen Evaluation wird der Gesamtcharakter bzw. die Komplettpformance einer Maßnahme evaluiert – im Gegensatz zur analytischen Evaluation handelt es sich bei der globalen Variante um eine Evaluation auf Macro- Ebene.

In pragmatischer Hinsicht sind globale Evaluationen schneller umzusetzen und in der Regel unkomplizierter zu bearbeiten als analytische Evaluationen, die sehr detailliert vorgehen.

2.9.22 Alternative Evaluationsmodelle

Als Gegenbewegung zu den in den 1970er Jahren häufig verwendeten Modellen (vorwiegend zielorientierte Ansätze), gelten alternative Evaluationsmodelle. Sie werden in erster Linie als Kritik an den in dieser Zeit verwendeten Modellen gesehen.

Hense beschreibt die Merkmale von alternativen Evaluationsmodellen folgendermaßen:

„Kennzeichen sind u.a. die Aufwertung der Rolle der Stakeholder im Evaluationsprozess, Methoden- und Wertepluralismus, qualitative Methoden, Präferenzierung sekundärer Evaluationszwecke und teils eine konstruktivistisch beeinflusste erkenntnistheoretische Fundierung.“¹⁷⁵

Empowerment evaluation (Fetterman)¹⁷⁶

Der Evaluator zieht sich im Rahmen dieses Modells stark zurück und lässt die Evaluation von den Programmbeteiligten konzipieren und durchführen. Neben dieser Eigenverantwortung spielt der Evaluator eine beratende Rolle und ist somit nur am Rande in den Evaluationsprozess involviert.

Dieses Modell bezieht sich im praktischen Einsatz wesentlich auf die Vorgehensweise bei der Selbstevaluation (siehe Kapitel 2.8.4) und soll die häufig auftretende Differenz und mangelnde Akzeptanz des Evaluators den Projektdurchführenden und -verantwortlichen gegenüber durch dieses „empowerment“ egalisieren.

¹⁷⁴ Vgl. ibidem; 177 f.

¹⁷⁵ Hense J (2007)

¹⁷⁶ Vgl Fetterman D (2007)

2.9.23 Modelle in der Evaluation PROFIL-TT

In die Evaluation von PROFIL-TT sind verschiedene Ansätze aus unterschiedlichen Evaluationsmodellen integriert worden (Modellmix). Der Einsatz von deskriptiven Techniken in Verbindung mit askriptiven und zielorientierten Ansätzen wird im Design und der Datenauswertung deutlich. Die Bezugnahme auf Projekt- und die Lernziele der einzelnen Module spielt im Design der Erhebungsinstrumente ebenso eine wesentliche Rolle. Von Beginn des Projekts an wurde die Durchführung und Umsetzung der Evaluation PROFIL-TT geplant und jeder Schritt mittels eines genauen Zeitplans festgelegt.

2.10 Evaluationsstandards¹⁷⁷

2.10.1 Übersicht

Das Ziel einer Evaluation von Programmen¹⁷⁸ besteht in erster Linie in der Bestimmung und Feststellung der Qualität. In diesem Rahmen bietet Evaluation die Möglichkeit, Hinweise zur Verbesserung zu geben und dadurch die Implementierung von Maßnahmen und Interventionen hilfreich zu unterstützen.

Die Rolle des Evaluators ist dabei vor allem von der interessensgebundenen Informationssammlung für seine Auftraggeber sowie alle anderweitig von der Evaluation betroffenen Personengruppen zur Effizienz- und Kosteneinschätzung eines Programms geprägt.

Eine valide Richtlinie zur Durchführung von Evaluationen bieten die *Program Evaluation Standards* des *Joint Committee on Standards for Educational Evaluation*. Diese Standards wurden von der Deutschen Gesellschaft für Evaluation (DeGEval; <http://www.degeval.de>) übersetzt und angepasst, und sollen die Wissenschaftlichkeit bei Konzeption und Durchführung von Evaluationen sicherstellen.

Bei der deutschen Übersetzung und Interpretation wurden die ursprünglichen, amerikanischen Standards geringfügig abgeändert und vereinzelt angepasst.

Der folgende Teil der Arbeit soll die Standards des Joint Committee on Standards for Educational Evaluation darstellen und näher erläutern.¹⁷⁹

Das Joint Committee on Standards for Educational Evaluation differenziert die dreißig Evaluationsstandards in vier Oberkategorien:

- Utility Standards (Nützlichkeitsstandards)

¹⁷⁷ Beywl W / Widmer T (2006)

¹⁷⁸ Verwendungsinfo: „Programm“ umfasst in der Verwendung des *Joint Committee on Standards for Educational Evaluation* Evaluationsgegenstände aller Art (Programme, Projekte, Materialien [Maßnahmen, Interventionen]) aus den Bereichen Erziehung, Aus- und Weiterbildung (ibidem; 28)

¹⁷⁹ Eine Übersicht über alle Standards des Joint Committee on Standards for Educational Evaluation befindet sich im Anhang I.

- Feasibility Standards (Durchführbarkeitsstandards)
- Propriety Standards (Korrektheitsstandards)
- Accuracy Standards (Genauigkeitsstandards)

2.10.2 Hintergrund der Evaluationsstandards

Die Standards dienen als Richtlinie für die Durchführung einer konsistenten und aussagekräftigen Evaluation und geben Empfehlungen dafür ab.

„Die mit den Standards gestellten Anforderungen sind oft sehr anspruchsvoll und konkurrieren miteinander. Eine unter realen Bedingungen stattfindende Evaluation vermag demzufolge kaum alle Standards gleichzeitig in vollem Umfang zu erfüllen.“¹⁸⁰

Die allgemeine Zielsetzung der Standards besteht darin, Evaluatoren Richtlinien und unterstützende Hilfestellungen zu bieten. Diese Unterstützung wendet sich allerdings auch an all diejenigen, die in eine Evaluation involviert oder davon betroffen sind, ein qualitativ möglichst hochwertiges Evaluationsprojekt zu realisieren.

Die Standards sollen darüber hinaus dazu dienen, die Qualität einer Evaluation für Auftraggeber besser einschätzen zu können und Transparenz zu schaffen.

Die dritte Zielsetzung, die von den Evaluationsstandards beabsichtigt wird, betrifft die Schulung, Aus- und Weiterbildung von Evaluatoren. Die Standards sollen als Grundstock für die Qualifizierung von Evaluatoren dienen, und somit das „Basiswerkzeug“ für fähiges Fachpersonal im Bereich der Evaluation bieten.

Grundsätzlich sollten die Standards nur in einer situationsgebundenen Abhängigkeit eingesetzt werden, bei der für jeden spezifischen Kontext entschieden werden muss, welche Standards angewendet werden können. Die Anwendung und die Entscheidungen, welche Standards berücksichtigt werden und bei welchen das Augenmerk mehr auf die Pragmatik gesetzt wird, sollten dabei begründet und für Außenstehende nachvollziehbar sein.

Die Evaluationsstandards sollen Evaluatoren in wichtigen Fragestellungen und Aspekten einer Evaluation unterstützen. Für die meisten Standards kann eine funktionale Zuordnung zu den Phasen einer Evaluation getroffen werden.

Mittels der Standards kann entschieden werden, ob es überhaupt zweckdienlich und sinnvoll ist, eine Evaluation durchzuführen; sie bieten generell Hilfestellungen bei der Eingrenzung des Evaluationsgegenstandes und der Umsetzung sämtlicher Phasen einer Evaluation – angefangen von der Konzeption, über die Informationsgewinnung und -auswertung bis hin zum Reporting. Ferner werden durch die Standards administrative Angelegenheiten – beispielsweise die Ausarbeitung eines Evaluationsvertrags, wichtige Hinweise zum Evaluationsmanagement sowie die personelle Planung und die allgemeine Budgetierung – erklärt und reglementiert.

¹⁸⁰ Beywl W / Widmer T (2006). 10

2.10.3 Nützlichkeitsstandards¹⁸¹

Das Joint Committee on Standards for Educational Evaluation legt in sieben Standards Kriterien für die Nützlichkeit einer Evaluation fest.

Generell wird in diesen Standards proklamiert, dass eine Evaluation in ihrer Gesamtheit die verschiedenen betroffenen Zielgruppen erreichen sollte, die die Evaluation entweder in Auftrag gegeben oder ein bestimmtes Interesse an der Durchführung und den Ergebnissen haben, oder einen – wie auch immer gearteten – sonstigen Nutzen aus der Evaluation eines Programms ziehen können.

Die Nützlichkeit von Evaluationen bezieht sich hauptsächlich auf den Informationsgehalt, den eine Evaluation für den Nutzer haben sollte, den richtigen Zeitpunkt sowie die Wirksamkeit.

N1 Ermittlung der Beteiligten und Betroffenen¹⁸²

In der Regel sind an Evaluationen mehrerer Personengruppen beteiligt oder auf unterschiedliche Weise involviert. Die erste Aufgabe eines Evaluators sollte es somit sein, diese unterschiedlichen Personengruppen zu identifizieren – Stakeholder, alle sonstigen an der Evaluation beteiligen oder auf irgendeine Weise von ihr betroffenen Personen – und festzustellen, welche Informationen sie sich von der Evaluation des vorliegenden Programms erwarten. Eine Vernachlässigung oder sogar Diskriminierung von beteiligten Personengruppen kann dazu führen, dass Informationsblockaden gegenüber dem Evaluator aufgebaut werden. Unter derartigen Störungen können die allgemeine Akzeptanz der Evaluation und letztendlich die Umsetzung der Ergebnisse leiden.

Ein zielführende Vorgehensweise bei der Identifizierung der Beteiligten eines Evaluationsprojekts geht zunächst von der Ermittlung von Führungspersonlichkeiten aus, mit deren Unterstützung der Evaluator dann – vergleichbar mit einem „Schneeballeffekt“ – weitere Beteiligte herausfinden kann. Als nächster Schritt sollte mit den einzelnen Betroffenen Kontakt aufgenommen und mit ihnen die persönliche Relevanz des vorliegenden Evaluationsprojekts besprochen und festgehalten werden.

Vor allem bei großen Evaluationsprojekten gibt es eine Vielzahl von involvierten Personen, was diesen Prozess sehr langwierig und Ressourcen- beanspruchend gestalten kann. Zudem wird es für den Evaluator – bei einer Vielzahl von Meinungen, Interessen und Wünschen – nahezu unmöglich, all diesen Erwartungen gerecht zu werden. Der Kontakt mit zu vielen beteiligten Personen kann im Extremfall die Durchführung einer Evaluation verhindern. In den Standards wird deswegen empfohlen, die einzelnen Beteiligten und deren Präferenzen nach ihrer Wichtigkeit für den Evaluationsprozess zu priorisieren. Diese Einteilung sollte mit dem Auftraggeber der Evaluation abgesprochen und abgestimmt sein. Von einer Differenzierung zwischen Auftraggeber und den übrigen Beteiligten wird allerdings abgeraten, der Auftraggeber sollte ebenso als Nutzer einer Evaluation gelten, wie andere Beteiligte.

¹⁸¹ Vgl. *ibidem*; 51 – 90

¹⁸² Vgl. *ibidem*; 53 – 57

N2 Glaubwürdigkeit des Evaluators¹⁸³

Ein grundlegender Anspruch, der an einen Evaluator gestellt wird, ist persönliche Fachkompetenz. Er sollte einen Überblick über die Methodenvielfalt einer Evaluation haben und diese Methoden auch zielgerichtet und unter Berücksichtigung von Qualitätsstandards einsetzen. Eine ausreichende Qualifikation hierfür sollte vorausgesetzt werden können. Ferner schafft ein kommunikatives und offenes Verhältnis des Evaluators gegenüber dem Auftraggeber sowie sonstigen Beteiligten eine Atmosphäre, die als Nährboden sowohl für eine gute Zusammenarbeit, als auch für die Akzeptanz und Glaubwürdigkeit des Evaluators als Person und der Evaluation als Werk dieser Person wesentlich ist.

Die Glaubwürdigkeit kann durch diese regelmäßige und offene Kommunikation mit den Hauptbeteiligten – zum Beispiel durch eine realitätsnahe und technisch- qualitativ hochwertige Beschreibung der Vorgehensweise bei der Evaluation – gesichert werden. Erfahrungen und Referenzen eines Evaluators auf dem entsprechenden Untersuchungsgebiet fördern die Glaubwürdigkeit und das Standing.

Im Rahmen dieses Standards wird nahe gelegt, dass – sollte es einem Evaluator nicht gelingen seinen Auftraggebern Glaubwürdigkeit zu vermitteln – der Evaluationsauftrag nicht angenommen, beziehungsweise die Evaluation von einer anderen Person durchgeführt wird.

N3 Umfang und Auswahl der Informationen¹⁸⁴

Bei der Art und Beschaffenheit der erhobenen Informationen spielt die Relevanz für den Auftraggeber eine wichtige Rolle, da der Nutzen einer Evaluation wesentlich von der Verwendbarkeit der Ergebnisse abhängt. Die Ergebnisse sollten zweckmäßig und für die Nutzer einer Evaluation hilfreich sein. Hierfür ist es notwendig, dass der Evaluator die Evaluationsziele der Auftraggeber kennt. Allerdings sollte auch darauf geachtet werden, dass nicht nur Daten erhoben werden, die explizit vom Auftraggeber gewünscht wurden, sondern es sollten generell die Dimensionen *„Effektivität, unerwünschte Nebenwirkungen, Kosten, Bezugnahme auf die Bedürfnisse der Lernenden, Sinnhaftigkeit der dem Programm zugrunde liegenden Annahmen und Wertvorstellungen und Durchführbarkeit“* (Handbuch der Evaluationsstandards 2006, 65) erhoben werden. Diese globale Sicht auf einen Evaluationsgegenstand sichert dem Evaluator die Unabhängigkeit der Untersuchung.

Da im Rahmen eines Evaluationsprojekts häufig nicht alle Fragen beantwortet werden können – diese Tatsache sollte vor der Evaluation auch an die Verantwortlichen und Beteiligten kommuniziert werden– besteht die Aufgabe des Evaluators auch darin, in Absprache mit dem Auftraggeber Untersuchungsschwerpunkte zu selektieren und diese wichtigen und hochrelevanten Inhalte zu fokussieren. Diese Selektion verhindert oftmals die Erhebung von irrelevanten Daten. Der Evaluator sollte bei der Entwicklung realistischer Erwartungen an die Evaluationsergebnisse die Beteiligten unterstützen.

Persönliche „Untersuchungsvorlieben“ des Evaluators sollten in diesem Prozess ausgeblendet und nicht berücksichtigt werden.

¹⁸³ Vgl. *ibidem*; 59 – 63

¹⁸⁴ Vgl. *ibidem*; 65 – 69

Die zu Beginn der Evaluation festgelegten Untersuchungsschwerpunkte sollten im kompletten Verlauf des Programms verfolgt werden, wobei auch – vor allem im Rahmen formativer Evaluationen – dem Anspruch von Flexibilität Genüge geleistet werden sollte.

Diese Themenaktualisierung findet häufig aufgrund mangelnder Kommunikation zwischen den Parteien nicht statt. Ein weiteres Problem, bei der Umsetzung dieses Standards, besteht in der Wahl der Datenerhebungsinstrumente. So werden – zumeist aus finanziellen Gründen – bereits vorhandene Instrumente für die Informationssammlung verwendet, die allerdings die notwendigen Informationen nicht lückenlos abfragen können.

N4 Feststellung von Werten¹⁸⁵

Die Feststellung von Werten dient als Hintergrund für jede Ergebnisinterpretation. Ausschlaggebend dabei ist die Einteilung der Inhalte nach der jeweiligen Nützlichkeit.

Für den Evaluator stellen sich bei der Feststellung der Werte zwei Fragen:

- Wer nimmt die Werturteile vor?
- Nach welchen Kriterien werden Werte interpretiert?

In diesem Schritt der Evaluation ist es wichtig, dass der Evaluator und der Auftraggeber gemeinsam ein Verfahren auswählen, um Informationen Werte zuweisen zu können. Diese Auswahl sollte transparent sein und begründet werden, sowie unterschiedliche Möglichkeiten, die zur Interpretation der Evaluationsdaten und zur Wertzuweisung verwendet werden können, einbeziehen. Interpretationsalternativen – zum Beispiel separate Interpretation durch verschiedene Evaluatoren – können die Wertbeurteilung unter Umständen hilfreich ergänzen. Aus den Alternativen zur Wertbeurteilung sollte die für den Auftraggeber das nützlichste Resultat gewählt werden.

N5 Klarheit des Berichts¹⁸⁶

Ein Evaluationsbericht kann auf unterschiedliche mediale Weise – je nach Zielsetzung der Evaluation – gehalten, bzw. abgegeben werden. Bei formativen Evaluationen wird in der Regel hoher Wert auf ein fortlaufendes Reporting gelegt, was häufig auch in mündlicher Form und informell stattfinden kann. Im Rahmen summativer Evaluationen sollte das Reporting einen formellen, schriftlichen Charakter haben.

Eine Grundvoraussetzung bei jeder Form der Berichterstattung ist die Klarheit und Transparenz der Inhalte, wobei die Genauigkeit der Ergebnisse nicht darunter leiden darf. Evaluationsberichte sollten in erster Linie verständlich und glaubwürdig abgefasst sein.

Der Evaluator sollte sich im Vorfeld der Evaluation Gedanken über die Form der Ergebnispräsentation machen.

In der Evaluationspraxis werden häufig zwei Reports verfasst, ein technischer und ein deskriptiver. Bei beiden sollten allerdings die Prämissen der Verständlichkeit – dies trifft vor allem auf Fachbegriffe zu – und der Fokus auf wesentliche Kernaussagen eingehalten werden. Die Methodologie sollte erwähnt, allerdings nicht zu stark

¹⁸⁵ Vgl. ibidem; 71 – 76

¹⁸⁶ Vgl. ibidem; 77 – 80

hervorgehoben werden, da der Schwerpunkt eines Evaluationsberichts auf den Ergebnissen liegen sollte. Eine Zielgruppenanalyse ist beim Verfassen des Reports meist hilfreich, da, basierend darauf eine Einschätzung bezüglich Hintergrund- und Kontextwissen oder einer eventuell notwendigen, mehrsprachigen Ausarbeitung getroffen werden kann.

In vielen Fällen wird seitens der Auftraggeber vor der Veröffentlichung eine finale Abstimmung des Evaluationsberichts gewünscht.

N6 Rechtzeitigkeit und Verbreitung des Berichts¹⁸⁷

Evaluationsergebnisse können für die Beteiligten nur dann von Nutzen sein, wenn sie zum richtigen Zeitpunkt kommuniziert werden. Voraussetzung hierfür ist, dass die Datenerhebung, -auswertung und -interpretation im Rahmen eines zu Beginn der Evaluation erstellten Zeitplans stattfindet – und auch eingehalten wird / werden kann.

Dieser zeitliche „Evaluationsfahrplan“ dient sowohl dem Evaluator als auch dem Auftraggeber als Orientierung und informiert über den aktuellen Stand der Untersuchung. Sofern möglich, ist es ratsam, Zeitfenster für unvorhersehbare Probleme sowie Zeitpunkte für Zwischenreports festzulegen.

Der Evaluator sollte neben einem Zeitplan auch – zusammen mit dem Auftraggeber – einen „Verbreitungsplan“ ausarbeiten, in dem beschrieben wird, welche Beteiligten zu welchem Zeitpunkt auf welche Weise welche Ergebnisse zur Verfügung gestellt bekommen.

N7 Wirkung der Evaluation¹⁸⁸

Die Wirkung („impact“) einer Evaluation hat einen wesentlichen Einfluss auf das weitere Vorgehen der Nutzer. Die Aufgabe des Evaluators besteht in diesem Rahmen in der Unterstützung aller Beteiligten bei der Ableitung von Maßnahmen, die aus den Evaluationsergebnissen resultieren, bis hin zur Implementierung der erarbeiteten Verbesserungsvorschläge; die Umsetzung solcher – meist weit reichender struktureller – Änderungen stößt häufig auf Widerstand, den der Evaluator abbauen sollte. Eine klare Trennung der Kompetenzen von Evaluator und Auftraggebers ist in diesem Prozess wesentlich, da es in erster Linie die Aufgabe des Auftraggebers ist, zu entscheiden, welche Konsequenzen aus den Evaluationsergebnissen gezogen werden; der Evaluator kann in diesen Prozess als Berater miteinbezogen werden.

Die Wirkung und der eigentliche Nutzen eines Evaluationsprojekts sollte allen Beteiligten innerhalb der Konzeptionsphase offen gelegt werden. Diese Kommunikation sollte nicht nur den theoretischen Nutzen betreffen, sondern auch die praktische Umsetzung der Evaluationsergebnisse.

¹⁸⁷ Vgl. *ibidem*; 81 – 86

¹⁸⁸ Vgl. *ibidem*; 87 – 90

2.10.4 Durchführbarkeitstandards

„Die Durchführbarkeitsstandards sollen sicherstellen, daß eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewußt ausgeführt wird.“¹⁸⁹

Evaluationen finden in der Regel nicht im Labor, sondern in der Praxis statt. Dabei ist keine Evaluation wie die andere; die Grundvoraussetzungen, intendierte und nicht- intendierte Variablen sowie die Zielgruppen differieren. Der Evaluator muss sich auf diese sich ständig ändernden Rahmenbedingungen und Voraussetzungen anpassen, die Durchführbarkeitsstandards bieten dabei hilfreiche Richtlinien.

D1 Praktische Verfahren¹⁹⁰

Um die Ergebnisse einer Evaluation möglichst realitätsnah halten zu können, ist es wichtig, dass die Instrumente, mit denen in der Praxis Daten erhoben werden, die Intervention möglichst wenig – am besten überhaupt nicht – stören oder beeinträchtigen.

Untersuchungsbedingte Eingriffe in eine Maßnahme führen zwangsläufig dazu, dass sich der Ablauf oder die Rahmenbedingungen ändern. Die Wirkung kann sich durch derartige Eingriffe ebenso verändern. Die Wahl von geeigneten Methoden, die zudem realistisch eingesetzt werden können, ist für die zielgerichtete Umsetzung notwendig. Sofern die Möglichkeit besteht sollten Evaluationsuntersuchung in den Rahmen von „Routineprozessen“ eingebettet werden, eine Störung der gesamten Maßnahme lässt sich auf diese Weise stark eingrenzen.

Ressourcen und Zeitraum, die für die diskrete Umsetzung der Evaluation nötig sind, sollten so gewählt werden, dass eine Umsetzung auch in der Praxis möglich ist. Zum Beispiel ist es unmöglich eine Evaluation mit zuwenig Personal in zu engen Zeitfenstern ergebnisorientiert umzusetzen; die Ergebnisse in solchen Fällen sind in der Regel nicht zu verwenden; der Nutzen für die Auftraggeber kann dadurch stark dezimiert werden.

Für den Evaluator ist es prinzipiell ratsam, eine Evaluation so praktisch wie nötig und so theoretisch wie möglich umzusetzen. Die Richtlinie besteht in der Kombination von Pragmatik und Anwendungsorientierung. Wissenschaftlich- theoretische Methoden können nur sinnvoll in einem Rahmen verwendet werden, den die praktische Umsetzung gestattet. Die Genauigkeit der Datenerhebung sollte allerdings auch unter diesen Gesichtspunkten nicht außer Acht gelassen werden.

Um eine praktische Umsetzbarkeit der Vorgehensweise sicherstellen zu können, besteht die Möglichkeit, eine preformative Evaluation durchzuführen, wobei Erhebungsinstrumente und Vorgehensweise zunächst in einer Testumgebung angewendet werden, um den gesamten Versuchsapparat zu prüfen (vgl. Kapitel 2.8.1).

D2 Politische Tragfähigkeit¹⁹¹

Durch den bewertenden Charakter, durch den „Evaluation [...] Entscheidungen zur Umverteilung von Ressourcen und Macht herführen kann“ (ibidem; 99) besteht bei jeder Evaluation die generelle Gefahr, dass aus dem

¹⁸⁹ Ibidem; 91

¹⁹⁰ Vgl. ibidem; 93 – 98

¹⁹¹ Vgl. ibidem, 99 – 104

einschätzenden Instrument ein machtpolitisches Werkzeug wird, der von unterschiedlichen Interessengruppen in verschiedenen Machtpositionen manipuliert und missbraucht werden könnte.

Vor diesem Hintergrund ist es vor Beginn einer Evaluation wichtig, dass sich der Evaluator einen umfassenden Überblick über Machtstrukturen und Interessenverteilungen der Beteiligten und Auftraggeber verschafft. Diesem Überblick sollte eine Neutralität gegenüber allen und Gleichberechtigung aller beteiligten Interessensgruppen folgen. Durch eine verstärkte Zuneigung zu einer bestimmten Interessengruppe wäre die Akzeptanz der Evaluationsergebnisse nicht mehr gesichert.

Durch regelmäßige Statusberichte und Zwischenreports an alle beteiligten Interessengruppen kann der Evaluator die Reaktionen der Beteiligten einschätzen und die Akzeptanz der Ergebnisse bei der Zielgruppe stärken.

D3 Kostenwirksamkeit¹⁹²

Das Ziel einer jeden Evaluation ist es einen Nutzen für Auftraggeber und Beteiligte zu schaffen. Dieser Nutzen kann auf unterschiedliche Weise festgestellt werden. Auf der obersten Ebene ist eine Maßnahme dann effizient, wenn die Summe aller aufgewendeten Ressourcen, kleiner ist, als der „net effect“ einer Evaluation.

Die zweite Möglichkeit, den Nutzen einer Intervention zu bestimmen, besteht in der Untersuchung der Wirksamkeit. Kann anhand von Maßnahmen eine – eventuell – positive Wirkung nachgewiesen werden, oder nicht. Im Rahmen der Effizienzsteigerung einer Intervention wird häufig versucht, die gute Performance einer Maßnahme beizubehalten – vielleicht sogar noch zu steigern –, wobei die finanziellen und nicht- finanziellen Aufwendungen reduziert werden.

Grundsätzlich ist eine derartige Kosten- und Aufwandsschätzung einer Intervention sehr aufwändig und kompliziert, da die Einschätzungen von verschiedenen Stellen stark differieren können.

Mit Hilfe einer klar strukturierten und übersichtlichen Budgetplanung, in der alle vorkommenden Positionen einer Evaluation aufgeführt sind und einer Sammlung und Gegenüberstellung des Evaluationsnutzens zu diesen Kosten, kann die Kostenwirksamkeit einer Evaluation begründet und nachgewiesen werden.

Diese Kosten- Nutzen- Analyse sollte vor Beginn jeder Evaluation durchgeführt werden, da mit dieser Hilfe die Sinnhaftigkeit der Durchführung einer Evaluation gestärkt werden kann. Außerdem sollte bei der Methodenauswahl stets dieser Kosten- Nutzen- Faktor berücksichtigt werden, da komplizierte – und kostenintensive – Vorgehensweisen unter Umständen auch zu effizienteren Resultaten führen können.

2.10.5 Korrektheitsstandards

„Die Korrektheitsstandards sollen sicherstellen, daß eine Evaluation rechtlich und ethisch korrekt durchgeführt wird und dem Wohlergehen der in die Evaluation einbezogenen und auch der durch die Ergebnisse betroffenen Personen gebührende Aufmerksamkeit widmet.“¹⁹³

¹⁹² Vgl. ibidem, 105 – 109

¹⁹³ Ibidem; 111

An umfangreichen Evaluationsprojekten ist häufig eine große Zahl von Personen involviert. Der Schutz der Rechte von Einzelpersonen – gesetzliche Regelungen zur Privatsphäre und der Schutz der Persönlichkeit – muss dabei unbedingt berücksichtigt werden.

K1 Unterstützung der Dienstleistungsorientierung¹⁹⁴

Dieser Standard entfällt bei der Deutschen Gesellschaft für Evaluation.

Programme und Interventionen sollten an den Interessen der Teilnehmer ausgerichtet sein, damit für diese Zielgruppe ein höchstmöglicher Nutzen daraus gezogen werden kann. Dieses Ziel sollte sichergestellt sein. Evaluationen leisten zur Qualitätsverbesserung und -sicherung einen wesentlichen Beitrag.

Informationssammlung und regelmäßige Berichterstattung gegenüber den beteiligten Personengruppen im Sinne der Dienstleistungsorientierung sind geboten.

K2 Formale Vereinbarungen (=F1)¹⁹⁵

In den formalen Vereinbarungen einer Evaluation sollten grundsätzliche Rahmenbedingungen, Vorgehensweisen und organisatorische Angelegenheiten diskutiert und im Rahmen einer vertraglichen Einigung seitens der Auftraggeber und Evaluatoren festgehalten werden. Dieser klassische Evaluationsvertrag sollte auch generelle Personalangelegenheiten und -verantwortlichkeiten regeln. Ferner werden darin die Häufigkeit und Art des Ergebnisreportings, die Vorgehensweise bei der Datenerhebung und -auswertung und die Finanzierung diskutiert und beschlossen. Die formalen Vereinbarungen einer Evaluation beinhalten auch die Abstimmung des Evaluationsplans.

Dieser Vertrag sollte während des gesamten Evaluationsprozesses von allen beteiligten Interessengruppen eingehalten, Änderungen dokumentiert und gegebenenfalls aktualisiert werden.

K3 Schutz individueller Menschenrechte (=F2)¹⁹⁶

Im Rahmen von Evaluationsstudien müssen – neben den generell geltenden Menschenrechten¹⁹⁷ – Rechte eingehalten werden, die sich spezifisch auf Personen beziehen, die an Evaluationen teilnehmen. Solche Regeln sind zum einen in Gesetzen manifestiert und werden dazu von moralisch-ethischen Grundsätzen abgeleitet.

Die Teilnehmer sollten freiwillig und zwanglos an der Untersuchung teilnehmen und die sanktionsfreie Möglichkeit haben, die Teilnahme an einer Evaluationsstudie zurückzuziehen. Die von ihnen gemachten Aussagen sollten ausnahmslos dem Informationsschutz – und somit einer vollkommenen Anonymisierung – unterstehen.

¹⁹⁴ Vgl. *ibidem*; 113 – 117

¹⁹⁵ Vgl. *ibidem*; 119 – 124

¹⁹⁶ Vgl. *ibidem*; 125 – 130

¹⁹⁷ Grundgesetz der Bundesrepublik Deutschland, Artikel 1, §1 „*Die Würde des Menschen ist unantastbar. Sie zu achten und zu schützen ist Verpflichtung aller staatlichen Gewalt.*“ Dieses grundlegende Menschenrecht – wie alle anderen auch – darf keinesfalls im Rahmen von Evaluationen vernachlässigt werden.

Die Aufklärung der Untersuchungsteilnehmer über diese Rechte fällt in den Aufgabenbereich des Evaluators und sollte von ihm klar kommuniziert werden.

Die Einhaltung dieser gesetzlichen und ethisch-moralischen Richtlinien ist wesentlich für eine allgemeine soziale Akzeptanz und ein sicheres Standing von Evaluationen.

Bei der Evaluation von Aus- und Weiterbildungsmaßnahmen ist generell darauf zu achten, dass alle Personengruppen die Möglichkeit haben, an diesen Maßnahmen teilzunehmen, was z.B. die Bildung einer – von der Maßnahme nicht betroffene – Kontrollgruppe häufig unmöglich macht.

K4 Human gestaltete Interaktion (=F2)¹⁹⁸

Auch im Rahmen von Evaluationsstudien gelten die Regeln für den zwischenmenschlich- interaktiven Bereich, die die Gefühle und Selbstachtung von Personen schützen. Diese Regeln müssen unbedingt eingehalten und sollten auch in interkulturellen Kontexten beachtet werden.

Eine regelmäßige, freundliche und informative Kommunikation erleichtert darüber hinaus die Durchführung einer Evaluation erheblich.

K5 Vollständige und faire Einschätzung (=F3)¹⁹⁹

Eine Orientierung bei der Evaluationsvorgehensweise an Feedbackregeln sorgt dafür, den evaluierten Gegenstand in seiner Gesamtheit zu betrachten. Die Darstellung sowohl der positiven, als auch der negativen Aspekte eines Programms in ausgewogener Form, sowie der intendierten und nicht- intendierten Zielsetzungen erfüllt diesen Anspruch auf Vollständigkeit. Meinungen und Sichtweisen aus anderen Perspektiven – möglichst vor der Veröffentlichung der Evaluationsergebnisse eingeholt – können eventuell übersehene Aspekte komplettieren. Eine Beeinflussungen der Ergebnisse, zum Beispiel insofern, dass negative Aspekte eines Programms seitens der Auftraggeber unterschlagen werden und im Reporting nicht auftauchen – führt in vielen Fällen zu Akzeptanzschwächen gegenüber der Evaluation.

K6 Offenlegung der Ergebnisse (=F5)²⁰⁰

Beteiligte und Betroffene einer Evaluation erwarten in der Regel, über die Ergebnisse einer Evaluationsstudie informiert zu werden. Vor Evaluationsbeginn sollte festgelegt werden, wie die Ergebniskommunikation während oder nach der Evaluation gestaltet werden soll. Grundsätzlich sollte die Information nach den Grundsätzen der „Direktheit, Offenheit und Vollständigkeit“ (ibidem; 143) erfolgen, um die Glaubwürdigkeit der Evaluation zu unterstützen. Mangelnde Information kann bei den Betroffenen die Auswirkung haben, dass Ergebnisse falsch interpretiert werden und daraus irrtümliche Schlussfolgerungen, im schlimmsten Falle sogar ungeeignete Maßnahmen abgeleitet werden. Dies würde auf ganzer Linie dem Grundsatz der Evaluation – einen Nutzen zu bringen – widersprechen.

¹⁹⁸ Vgl. Beywl W / Widmer T (2006). 131 – 135

¹⁹⁹ Vgl. ibidem; 137 – 141

²⁰⁰ Vgl. ibidem; 143 – 148

Der Evaluator sollte deswegen die aus der Evaluation resultierenden Informationen vor der Publikation ebenso kontrollierten können, wie die Auftraggeber; eine Kontrolle, ggf. Zensur sollte allerdings nur im allgemeinen Einvernehmen und gut begründet stattfinden.

K7 Deklaration von Interessenskonflikten (in F4)²⁰¹

Evaluationen werden durchgeführt, um Informationen zu sammeln, die Entscheidungen herbeiführen, unterstützen oder rechtfertigen sollen. Dieser Entscheidungsprozess wird häufig von sehr unterschiedlichen Interessen geleitet – siehe auch Standard D2 Politische Tragfähigkeit.

Diese – meist differierenden Interessen – sollten vom Evaluator transparent gemacht werden. Eine Vermittlung der verschiedenen Interessenvertreter kann dem Evaluationsprozess nur förderlich sein, da eine mögliche Verzerrung, Lenkung oder Verfälschung der Evaluationsergebnisse zum Vorteil einer bestimmten Personengruppe bei offen kommunizierten Interessen häufig ausgeschlossen werden kann. Durch eine Veröffentlichung der Evaluationsergebnisse können auch Evaluatoren -bezogene Interessen eingedämmt werden. Sollen die Interessenskonflikte so schwerwiegend sein, dass sie vom Evaluator – oder wegen des Evaluators – nicht kanalisiert werden können, empfiehlt sich die Durchführung einer Meta- Evaluation (Siehe Kapitel 2.10.6 – G12).

Vor allem bei internen oder Selbst- Evaluationen ist es ratsam, die Evaluatoren hierarchisch zu entgliedern, da eine Einflussnahme auf die Evaluation seitens Vorgesetzter oder Kollegen die objektive Umsetzung erschweren kann.

K8 Finanzielle Verantwortlichkeit (in D3)²⁰²

Evaluationsprojekte sollten nicht auf eine sorgfältige Buchführung und Dokumentation sämtlicher finanzieller Ausgaben und Aufwendungen seitens des Evaluators verzichten, um den Verdacht von Finanzmittelmissbrauch von vorneherein auszuräumen und dadurch das Ansehen der Evaluation und die Ergebnisse zu stärken.

Generell ist es ratsam, die Eck- Kostenpunkte einer Evaluation mit dem Auftraggeber abzustimmen und auf umfangreichere Posten – zum Beispiel Personal – explizit und detailliert einzugehen. Die Kommunikation der Ausgaben im Evaluationsbericht fördert diese Kostentransparenz und somit wiederum die Akzeptanz der Ergebnisse.

2.10.6 Genauigkeitsstandards

„Die Genauigkeitsstandards sollen sicherstellen, daß eine Evaluation über die Güte und/oder die Verwendbarkeit des evaluierten Programms fachlich angemessene Informationen hervorbringt und vermittelt.“²⁰³

²⁰¹ Vgl. ibidem; 149 – 154

²⁰² Vgl. ibidem; 155 – 158

²⁰³ Ibidem; 159

Die wichtigste Frage, die sich ein Evaluator bei einer Evaluation stellen sollte ist, ob die Instrumente mit denen er Daten erhebt auch das messen, was im Interessensmittelpunkt der Untersuchung steht, ob diese Daten auch der Realität entsprechen und angemessen informativ sind. Bezogen auf die Informationsauswertung sollte dieser logische Zusammenhang von Datenmaterial und Untersuchungsziel klar erkennbar sein und dargestellt werden.

G1 Programmdokumentation²⁰⁴

Grundlegend für die Durchführung einer Evaluation ist die genaue Kenntnis über die zu evaluierende Maßnahme. Dabei ist nicht nur die letztendliche Form der Durchführung ausschlaggebend, sondern auch die ursprünglich angedachte Konzeption.

Die Beziehung, in der die Programminhalte mit den Wirkungen stehen, wird durch eine klare und detaillierte Dokumentation oftmals erleichtert. Um diese Programmbeschreibung möglichst genau und realitätsgetreu zu gestalten, sollte der Evaluator auf unterschiedliche Informationsquellen zurückgreifen. Die Programmadministration sowie die Auftraggeber können diesen Prozess unterstützen, es ist allerdings aus Sicht des Evaluators auch ratsam, das gesamte Programm und die Rahmenbedingungen selbst zu beobachten und zu analysieren. Bei dieser Untersuchung fallen häufig auch Diskrepanzen zwischen der Programmkonzeption und der eigentlichen Umsetzung auf.

G2 Kontextanalyse²⁰⁵

Die Rahmenbedingungen einer Evaluation, alle Variablen, die nicht den direkten Ablauf und die Inhalte betreffen, können für die Ergebnisinterpretation von Interesse sein.

Eine detaillierte Beschreibung dieses Programmumfelds – am besten anhand verschiedener Informationsquellen – gilt deswegen für den Evaluator als wichtige Orientierung, mittels der er störende Einflüsse auf Evaluationen feststellen, dokumentieren und in die Ergebnisauswertung einbeziehen kann.

G3 Beschreibung von Zielen und Vorgehen²⁰⁶

Eine Evaluation definiert sich in der Regel über Ziele und Intentionen einer Maßnahme. Gelegentlich laufen die Ziele, die mit Hilfe einer Maßnahme verfolgt werden, unter unterschiedlichen, individuellen Betrachtungsweisen auseinander. Diese Differenz sollte von Beginn an festgestellt und diskutiert werden. Für die Evaluation von Vorteil – allerdings nicht immer möglich – ist eine Egalisierung dieser verschiedenen Interessen und Zielsetzungen. Grundvoraussetzung – sollten solche Zielkonflikte existieren – ist jedoch, dass der Evaluator diese unterschiedlichen Zielsetzungen mit den Betroffenen diskutiert und den Prozess dokumentiert.

²⁰⁴ Vgl. ibidem; 161 – 165

²⁰⁵ Vgl. ibidem; 167 – 170

G4 Verlässliche Informationsquellen²⁰⁷

Für Auftraggeber und Beteiligte einer Evaluation ist es in der Regel von Interesse, auf welche Weise die vorliegenden Evaluationsergebnisse zustande gekommen sind. Eine Benennung und Beschreibung der verwendeten Informationsquellen (Befragung von Einzelpersonen, Gruppen, Unterrichtsmaterial, etc.) ist deshalb ratsam.

Ferner ist für die Akzeptanz der Evaluation die Beschaffenheit der Untersuchungsgruppe (Population / Beschaffenheit von Stichproben, Änderung dieser Stichproben, etc.) wichtig, da darauf die Glaubwürdigkeit der Evaluation basiert. Wenn derartige Informationen vom Evaluator nicht geliefert werden, kann die Glaubwürdigkeit der Evaluationsergebnisse in Frage gestellt werden.

Neben der Dokumentation der Erhebungszielgruppe sollten auch die Erhebungsverfahren und -instrumente beschrieben werden. Beim Verwenden bereits vorhandener Daten und Instrumente ist meistens eine Qualitätsprüfung dieser Werte notwendig.

Benutzt der Evaluator mehrere unterschiedliche Informationsquellen, sollte die Daten auf Vergleichbarkeit oder Gleichwertigkeit hin untersucht werden, da ansonsten dieser Vergleich nicht stichhaltig ist. Ebenso sollte grundsätzlich darauf geachtet werden, dass alle Informationsquellen nur zu einem bestimmten Grad stichhaltig sind und nach unterschiedlicher Güte der erhobenen Information eingeschätzt werden können. Der Unterschied von statistisch „eher harten“ zu statistisch „eher weichen“ Daten, sollte in der Ergebnisinterpretation auch berücksichtigt werden.

G5 Valide Informationen²⁰⁸

Die Validität von Informationen bei Evaluationsuntersuchungen stellt den qualitativen Wert und die Glaubwürdigkeit der Dateninterpretation sicher. Dabei wird unter Validierung der Sammel- Prozess von Nachweisen zur Unterstützung dieser Glaubhaftigkeit verstanden. In diesem Evaluationsstandard wird gefordert, dass ein Validierungsverfahren folgende Kriterien beinhaltet:

- 1) eine Analyse, welche Art von Information mittels welchen Verfahrens beschafft wird
- 2) eine Dokumentation dieses Verfahrens
- 3) eine Begründung, warum dieses Verfahren mit welcher Intention und Zielsetzung eingesetzt wurde

Die Validität im Rahmen von Evaluation ist immer kontext- und situationsbezogen, dabei ist es wichtig, dass die Schlussfolgerungen aus den erhobenen Daten valide sind, damit sichergestellt werden kann, dass die Informationen zur Beantwortung der Fragestellung einer Evaluation vorhanden sind. Um diese Validität garantieren zu können, werden häufig mehrere Datenerhebungsverfahren parallel eingesetzt und verschiedene Standards zu diesen Untersuchungen hinzugezogen. Der Vergleich von Belegen in fachlich und methodisch ähnlichen Evaluationen kann ebenso ein Hinweis auf Validität sein. Diese Methode ist – sofern sie im entsprechenden Kontext

²⁰⁶ Vgl. ibidem; 171 – 174

²⁰⁷ Vgl. ibidem; 175 – 178

²⁰⁸ Vgl. ibidem; 179 – 186

angewendet werden kann – ein sicheres Verfahren, um die Validität von Untersuchungsergebnissen sicherstellen zu können.

G6 Reliable Informationen²⁰⁹

Die Reliabilität einer Untersuchung „bezieht sich auf den Grad der Konsistenz der Informationen“ (ibidem; 187).

Um Reliabilitäten im Rahmen einer Datenerhebung bestimmen zu können, ist es wesentlich, nicht- intendierte Abweichungen feststellen zu können. Solche Abweichungen können entweder durch Störeinflüsse oder Zufallsfehler bei der Erhebung oder durch nachvollziehbare, erklärbare Systemeigenschaften ausgelöst werden.

Bei einigen etablierten Verfahren ist eine Anfälligkeit gegenüber bestimmten Fehlern, die die Reliabilität beeinträchtigen können, bekannt. Bei diesen Untersuchungsmethoden sollten die Ergebnisse explizit auf derartige Schwachstellen hin überprüft werden.

Ebenso wie bei der Validität empfiehlt es sich auch vor dem Hintergrund der Reliabilität – sofern die Möglichkeit besteht – Verfahren zu verwenden, die unter Reliabilitäts- Gesichtspunkten bereits erfolgreich eingesetzt wurden. Dabei sollten allerdings zwischen unterschiedlichen, kontextbezogenen Arten von Reliabilität, die unterschiedliche Messfehler ausdrücken können, differenziert werden; die aufgrund solcher Vergleiche gezogenen Schlussfolgerungen sind häufig nicht reliabel.

Reliable Informationen müssen darüber hinaus nicht analog auch valide sein, ebenso wie valide Informationen nicht unbedingt ein hohes Maß an Reliabilität besitzen müssen. Reliabilität und Validität bedingen sich nicht grundsätzlich.

G7 Systematische Informationsüberprüfung²¹⁰

Eine Fehleranfälligkeit beim Datenhandling besteht in jedem Stadium der Datenerhebung, -auswertung und -interpretation. Je mehr Prozessschritte die erhobenen Daten durchlaufen, desto höher ist die Wahrscheinlichkeit, dass einzelne Datensätze fehlerhaft sind. Fehler im Datenmaterial können zu schwerwiegenden Verzerrungen bei der Ergebnisinterpretation und dadurch zu einer stark geminderten Aussagekraft der Untersuchungsergebnisse führen, weswegen die Datenqualität unbedingt garantiert und abgesichert werden sollte. Durch regelmäßige Kontrollen der Datenbasis können solche Fehler entdeckt oder vermieden werden. Sollten innerhalb der Datenbasis vereinzelt Unstimmigkeiten und Unregelmäßigkeiten auftauchen empfiehlt es sich, die betroffenen Daten bis an die Datenbasis zurückzuverfolgen und einer intensiven Kontrolle zu unterziehen. Ebenso sollten häufig auftretende Fehler und typische Probleme im Datenhandling in Qualitätssicherungsschleifen fokussiert werden.

Die Kontrolle sollte sich auch auf computergestützte Datenanalyseverfahren beziehen, da auch in diesem Bereich vereinzelt Inkonsistenzen oder logische Fehler auftreten können.

²⁰⁹ Vgl. ibidem; 187 – 192

²¹⁰ Vgl. ibidem; 193 – 197

G8 Analyse quantitativer Informationen²¹¹

Die Aufgabe einer quantitativen Analyse besteht im Zusammenfassen, Ordnen und Interpretieren quantitativer Daten. Vor der eigentlichen „Hauptanalyse“ ist es ratsam, Testanalysen anhand fiktiver, allerdings nach Struktur und Beschaffenheit den empirischen Daten nachempfunderer Test- Daten durchzuführen, um das Datenhandling besser einschätzen und eventuell auftauchende „Ausreißer“ erkennen zu können. Eine Visualisierung der erhobenen und ausgewerteten Daten – vor allem im Bereich deskriptiver Statistik – erleichtert die Dateninterpretation und führt zu einer anschaulichen Aufbereitung der erhobenen Informationen. Sofern es finanzielle Mittel und die Zeit zulassen empfiehlt sich außerdem die Verwendung mehrerer Analysemethoden um die Stichprobenkonsistenz zu sichern.

Grundsätzlich sollte ein Analyse- und Darstellungsverfahren ausgewählt werden, das der Evaluation und den Daten angemessen ist und der Zielgruppe entspricht, die einen Nutzen aus den Ergebnissen ziehen sollte.

Qualitative und quantitative Daten sollten sich stets ergänzen – sofern die Möglichkeit besteht, beide Datentypen zu erheben.

G9 Analyse qualitativer Informationen²¹²

Bei qualitativen Daten und Informationen handelt es sich um verbale Ausdrücke, deren Analyse in der Regel narrative Texte zum Ergebnis hat. Diese Analyse kann häufig nicht mittels eines vorgefertigten Schemas durchgeführt werden, sondern muss interaktiv und kontextbezogen durchgeführt werden. Eine fachgemäße Analyse erkennt man an der Ableitung einiger Kategorien, die die erhobenen Daten gemäß der Fragestellung der Evaluation dokumentieren. Die Kategorien entsprechen dabei den Richtlinien von Validität und Reliabilität.

Für diese Analyse sollte ein geeignetes Verfahren gewählt werden, um die qualitativen Daten angemessen zusammenfassen zu können. Die Kriterien für die Kategorien sollten eine für die Evaluation hohe Aussagekraft besitzen und prägnant formuliert sein. Durch eine mehrfache Analyse desselben Datenmaterials von verschiedenen Evaluatoren kann die Aussagekraft der Analyse gesteigert werden. Fließt in die Analyse eine Differenzierung der Quellen nach Glaubwürdigkeit, Fachkompetenz oder Vereinnahmung mit ein, hat dies in der Regel eine Intensivierung der Aussagekraft zur Folge.

Solche qualitativen Analysen sind meist sehr aufwändig und kostenintensiv, wobei auch hier der generelle Kosten- Nutzen- Grundsatz einer Evaluation beachtet werden sollte.

G10 Begründete Schlussfolgerungen²¹³

Schlüsse und Interpretationen aus erhobenen Daten müssen begründet und logische Evidenzketten nachgewiesen werden, um die Akzeptanz der Erhebung und deren Ergebnisse zu sichern. Bei Interpretationen, die nicht vollständig bewiesen und argumentativ fundiert werden können stellt sich häufig die Frage nach einer eventuellen Fehlerhaftigkeit der Daten und Ergebnisse. Die Interpretationen und Schlüsse sollten sich nur auf „die Situatio-

²¹¹ Vgl. ibidem; 199 – 204

²¹² Vgl. ibidem; 205 – 210

²¹³ Vgl. ibidem; 211 – 214

nen, Zeiträume, Personen, Zusammenhänge und Zwecke“ (ibidem; 211) beziehen, die die Evaluation betreffen. Ein Feedback zu den Interpretationen von unterschiedlichen Programmmitwirkenden ist in diesem Rahmen oft hilfreich.

Auch bei den Schlussfolgerungen gilt eine valide und genaue Informationsbasis als Voraussetzung für eine gute Dateninterpretation und ein stimmiges Ergebnis.

G11 Unparteiische Berichterstattung²¹⁴

Das Reporting einer Evaluation sollte so neutral wie möglich durchgeführt werden. Diese Neutralität kann von Evaluatoren jedoch nur erreicht werden, wenn verschiedene Faktoren, die die Berichterstattung von Evaluationen verzerren oder verfälschen können, berücksichtigt und vermieden werden. Beteiligte mit stark divergierenden Interessen versuchen häufig auf die Evaluationsergebnisse und die Berichterstattung darüber Einfluss zu nehmen. Ein Evaluator sollte diese Gefahr frühzeitig erkennen und entsprechende Maßnahmen dagegen treffen, damit die Berichterstattung der Evaluation nicht in ein bestimmtes – meist vom Auftraggeber gewünschtes Ziel – gelenkt wird.

Generell sollte die Berichterstattung sowohl positive wie auch negative Aspekte berücksichtigen und über den Evaluationsgegenstand global berichten. Ein Vergleich mit Evaluationsberichten zu ähnlichen Untersuchungsgegenständen kann bei der Umsetzung unparteiischer Berichterstattung durch Referenzergebnisse behilflich sein.

G12 Meta-Evaluation²¹⁵

Eine Meta-Evaluation wird in der Regel dann durchgeführt, wenn die Ergebnisse, Vorgehensweisen und Instrumente einer Evaluation überprüft, kontrolliert oder bewertet werden sollen. Im Evaluationsprozess auftretende Fehler oder Unstimmigkeiten können durch eine solche Maßnahme aufgedeckt und ausgeräumt werden. Der Nutzen dieser Kontrollfunktion ist in vielen Fällen wesentlich.

Meta-Evaluationen werden häufig bereits formativ und begleitend zur Evaluierung der Konzeption und Implementierung der Evaluation eingesetzt; die Entscheidung für die Durchführung einer Meta-Evaluation sollte in diesem Fall zeitnah zu Beginn der eigentlichen Evaluation fallen. Die summative Funktion von Metaevaluation besteht in der anschließenden Bewertung der Evaluation, der Bewertung und Einschätzung der Ergebnisse, Interpretationen und Schlussfolgerungen sowie Umsetzung der Empfehlungen.

Wie bei Evaluationen generell ist auch bei Meta-Evaluationen eine Unabhängigkeit der Evaluatoren sehr wichtig, wobei die Kompetenzen der Meta-Evaluatoren so weit reichen sollten, Mitglieder des Evaluationsteams auszutauschen – zum Beispiel bei Nichterfüllung entsprechender Aufgaben oder schwerwiegendem Fehlverhalten.

Die Akzeptanz einer Evaluation kann durch die fachgerechte Durchführung einer Meta-Evaluation stark erhöht werden, bei Zweifeln an den Ergebnissen der Meta-Evaluation besteht ferner die Möglichkeit der Evaluation der

²¹⁴ Vgl. ibidem; 215 – 218

²¹⁵ Vgl. ibidem; 219 – 224

Meta- Evaluation. Diese Vorgehensweise kann beliebig – natürlich in Abhängigkeit von finanziellen und personellen Ressourcen – fortgesetzt werden.

2.10.7 Zusammenfassung der Standards

Die oben genannten Evaluationsstandards bieten die Möglichkeit, Evaluationen mit einem entsprechenden Nutzen für alle Beteiligten und auf hohem fachlichem Niveau durchzuführen. Andererseits ist es aufgrund der breiten Fächerung und Vielzahl der Standards nahezu unmöglich, in einer praktisch ausgerichteten Evaluation alle Richtlinien einzuhalten und alle Standards zu erfüllen. Manche Standards können aufgrund des Kontexts und der Rahmenbedingungen nur bedingt oder gar nicht eingehalten werden.

Das Joint Committee on Standards for Educational Evaluation gliedert deswegen die Standards in funktionale Gruppen ein und weist diese gewichtet den einzelnen Phasen der Evaluation zu. So können die Standards ihren Anwendungsschwerpunkt auf die erste Phase einer Evaluation beziehen, in der entschieden wird, ob die Durchführung einer Evaluation überhaupt möglich oder sinnvoll ist. In der nächsten Phase bieten sie Hilfestellungen, das Evaluationsproblem einzugrenzen und zu formulieren. Bei den Prozessschritten Evaluationsplanung, Datenerhebung, -auswertung und -interpretation, sowie Berichterstattung und Evaluationsorganisation kommen unterschiedliche Standards zum Tragen, wobei einige auch mehrfach und in unterschiedlichen Phasen der Evaluation herangezogen werden können.

Evaluationen sollen generell nützlich und an den Informationsbedürfnissen der Nutzer ausgerichtet (*Nützlichkeit*), realistisch, gut durchdacht und diplomatisch sein, dabei kostenbewusst durchgeführt werden (*Durchführbarkeit*), rechtlich und ethisch korrekt ablaufen und dem Wohlergehen der in die Evaluation einbezogenen und von den Ergebnissen betroffenen Personen Aufmerksamkeit schenken (*Korrektheit*) und als letztes über die Güte und/ oder die Verwendbarkeit eines evaluierten Programms fachlich angemessene Informationen hervorbringen (*Genauigkeit*). Nur wenn diese Metakriterien eingehalten werden, kann eine Evaluation die essentielle Zielsetzung – Bewertung und Einschätzung einer Maßnahme oder Intervention – erfüllen.

2.10.8 Durchführungsentscheidung einer Evaluation²¹⁶

Unter bestimmten Grundvoraussetzungen sollte eine Evaluation nicht durchgeführt werden. Im folgenden Abschnitt werden einige dieser Voraussetzungen beschrieben.

Wenn bereits vor Evaluationsbeginn feststeht, dass der erwünschte Informationszuwachs durch die Evaluation als unwesentlich eingeschätzt werden kann, da zum Beispiel eine Evaluation der Maßnahme bereits durchgeführt wurde und diese – abgesehen von den bereits vorhandenen Ergebnissen – keine neuen, relevanten Erkenntnisse zu erwarten sind, empfiehlt es sich, auf eine erneute Evaluation zu verzichten. Von einer Evaluation sollte auch dann abgesehen werden, wenn absehbar ist, dass die zur Verfügung stehenden personellen und finanziellen Res-

²¹⁶ Vgl. Stockmann R (2006). 284

sources für die geplante Umsetzung nicht ausreichen werden oder die Validität und Reliabilität der Ergebnisinterpretationen aufgrund von maßnahmenbezogenen Einflüssen nicht gewährleistet werden können.

Neben diesen pragmatischen Gründen, eine Evaluation nicht durchzuführen, gibt es zahlreiche politische Umstände, unter denen es ratsam ist, dieses Werkzeug nicht einzusetzen. Besteht seitens der Auftraggeber lediglich eine mangelnde Akzeptanz für die Ergebnisse, ist die Wahrscheinlichkeit, dass die von den Evaluationsergebnisse abgeleiteten Maßnahmen nicht implementiert werden, sehr hoch, was eine Evaluation überflüssig machen würde.

Sollte eine Evaluation als nachträgliches Rechtfertigungsinstrument für bereits getroffene Entscheidungen eingesetzt werden, wird von einer Durchführung ebenso abgeraten.

2.10.9 Anwendung der Evaluationsstandards im Projekt PROFiL-TT

Im Rahmen der Evaluation des Projekts PROFiL-TT wurde versucht, den Evaluationsstandards des Joint Committee on Standards for Educational Evaluation gerecht zu werden und diese generell einzuhalten. Dennoch war es im Rahmen der vorliegenden Evaluation nicht möglich alle Standards gleichermaßen einzubeziehen und zu berücksichtigen.

„Das Joint Committee ist sich bewusst, daß nicht alle Standards in allen Evaluationen gleichermaßen anwendbar sind. Es muß fachkundig entschieden werden, welche Standards für die jeweilige Situation anwendbar sind. Die Nutzer der Standards sollten die Sachdienlichkeit jedes Standards für einen bestimmten Kontext sorgfältig prüfen und daraufhin entscheiden, welchen Standards die größte Bedeutung beigemessen werden soll.“ (Handbuch der Evaluationsstandards 2006, 27)

Bei der Evaluation des Projekts PROFiL-TT wurde dieser Prämisse Folge geleistet und alle Standards auf Sachdienlichkeit und pragmatischen Einsatz hin überprüft und entsprechend angewendet. Die Pragmatik steckt dabei – wie häufig bei Evaluationen – einen Bewegungsrahmen für den Evaluator ab, der einen gewissen Spielraum zur Verfügung stellt. Grundsätzlich wurde bei dieser Evaluation darauf geachtet, dass die Ergebnisse aus der Datenerhebung durch die Einhaltung der Korrektheits- und Genauigkeitsstandards valide und reliabel sind, sich aber analog dazu auch in den Grenzen von Machbarkeit und Effizienzorientierung bewegen.

Als Instrument zur Datenerhebung wurde deswegen auf das Heidelberger Inventar zur Lehrevaluationsforschung zurückgegriffen.

2.11 Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE) nach Rindermann (1994)

2.11.1 Übersicht

Mit dem multifaktorellen Modell der Lehrveranstaltungsqualität, dem Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE), das 1994 entwickelt und seit dieser Zeit einem kontinuierlichen Verbesserungsprozess

unterzogen wurde, ist es Dr. Heiner Rindermann gelungen, ein Instrument zu generieren, das eine weit reichende, aussagekräftige, valide und reliable Bestimmung und Bewertung von Lehrveranstaltungen erlaubt.²¹⁷

Das Modell basiert grundlegend auf Befragungen von Studierenden und Lehrenden, die an unterschiedlichen Universitäten und Fachhochschulen in verschiedenen Studiengängen durchgeführt wurde. Der Schwerpunkt dieser Befragungen liegt dabei auf Kriterien, nach denen Lehre bewertet werden kann und anhand derer sich „gute Lehre“ definieren lässt.

Diese so gewonnene Datenmenge wurde anhand statistischer Methoden auf Validität, Reliabilität und mittels Faktorenanalysen kontrolliert und empirisch fundiert.

2.11.2 Dimensionen und Grundstruktur

Im folgenden Teil der Arbeit sollen die einzelnen Hauptfaktoren „Dozent“, „Studierende“, „Rahmenbedingungen“ und „Lehrerfolg“ des HILVE- Inventars vorgestellt und die Beziehungen in denen sie zueinander stehen erläutert werden.

Die Interaktion dieser vier Bedingungsfaktoren ist für die Güte von Lehrqualität ausschlaggebend. Nur wenn zwischen den Komponenten ein harmonisierendes Zusammenspiel herrscht, kann die Zielvariable „Lehrerfolg“ erreicht werden.

„Erst das Zusammenspiel eines fachlich, didaktisch und kommunikativ kompetenten Dozenten mit vorgebildeten, interessierten und engagierten Studenten sowie günstigen Rahmenbedingungen in Form von ausreichend Zeit, Platz, adäquater Anforderungshöhe und inhaltsangemessenen Veranstaltungsformen [...], kann Lehrerfolg mit hoher Wahrscheinlichkeit bewirken.“ (Rindermann 2001, 63)

Nach Rindermann sind somit die Faktoren Dozent, Studierender und Rahmenbedingungen für die Zielgröße „Lehrerfolg“ ausschlaggebend.

Die Beziehung zwischen den einzelnen Faktoren ist allerdings nicht statisch, sondern flexibel und in unterschiedlicher Ausprägung, abhängig vom jeweiligen Kontext, wechselseitig bedingend.

Die Wirkung der einzelnen Faktoren ist nicht separat, sondern erfolgt meist gebündelt und in der Kombination mehrerer Einzelfaktoren miteinander. Oftmals sind diese Beziehungen auch bedingend miteinander verkettet²¹⁸.

²¹⁷ Vgl. Rindermann H, Amelang M (1994)

²¹⁸ Rindermann führt folgende Beziehungen als Beispiele an:

Dozenten können Veranstaltungsteilnehmer z.B. durch den Einsatz geeigneter Lehrkonzepte motivieren, wodurch ihre Beteiligung an der Veranstaltung gesteigert wird. Andererseits können auch Prüfungen über den Inhalt der Veranstaltung die Beteiligung steigern. Die Methodenkompetenz des Dozenten wirkt wiederum auf die Anzahl der Teilnehmer, wogegen sich eine überfüllte Maßnahme – eine negative Rahmenbedingung – andererseits für den Lehrerfolg – z.B. durch die Zunahme von Störungen oder eine eingeschränkte Interaktionsmöglichkeit – negativ auswirken können. (vgl. Rindermann 2001, 64)

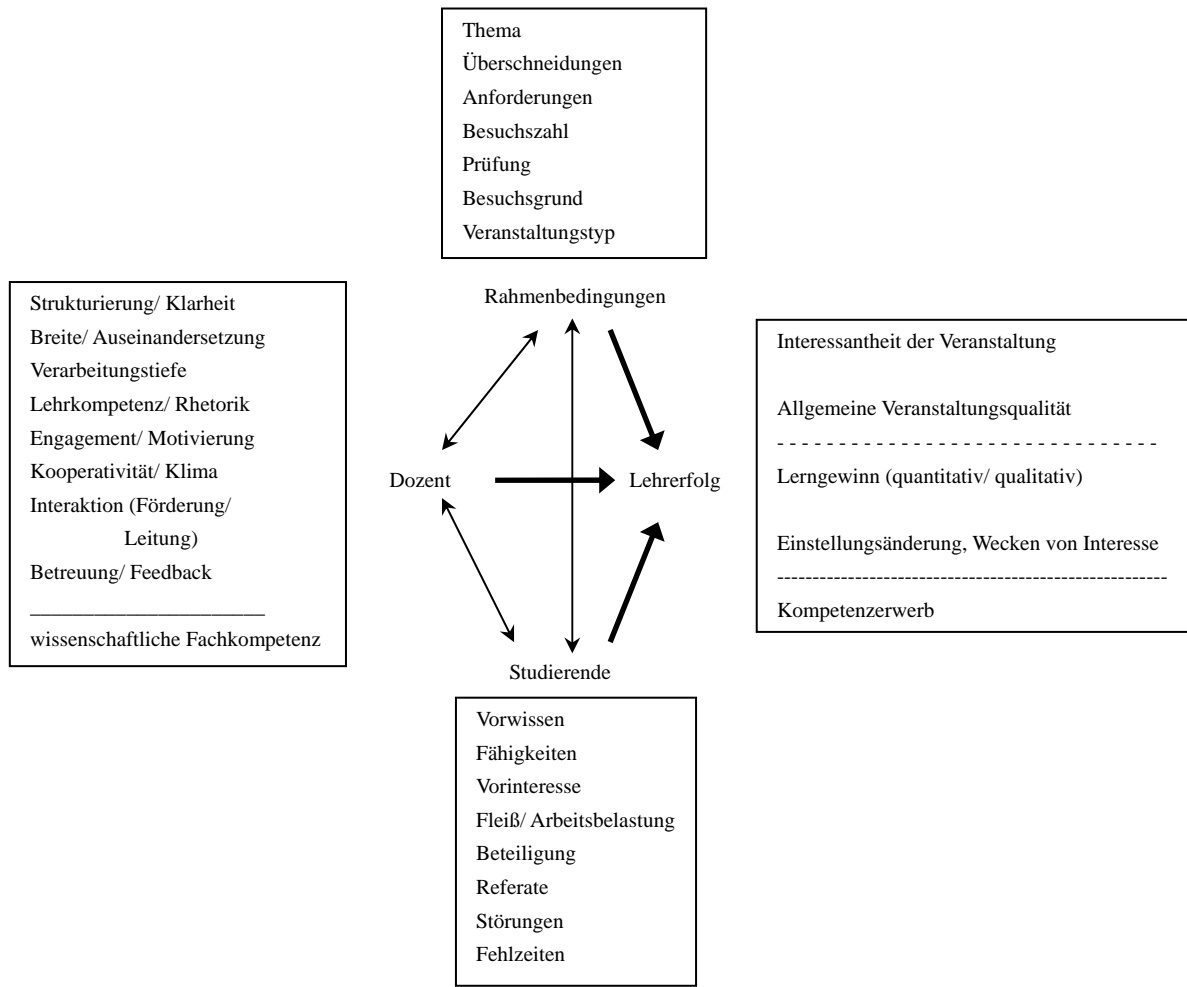


Abbildung 1: Multifaktorielles Modell der Lehrveranstaltungsqualität²¹⁹

Aus der graphischen Darstellung des multifaktorellen Modells der Lehrveranstaltungsqualität geht hervor, dass die drei Bedingungsfaktoren (Dozent, Studierende, Rahmenbedingungen) auf die „Produkt- Variable“ Lehrerfolg bezogen werden. Eine Interaktion findet dabei zwischen den Bedingungsfaktoren „Dozent“, „Studierende“ und „Rahmenbedingungen“ statt, die wiederum einzeln oder gebündelt auf den Lehrerfolg Einfluss nehmen.

Der eigentliche Lehrerfolg kann mittels positiver Beurteilungen der untersuchten Veranstaltungen sowie durch das Erreichen des beabsichtigten Effekts bei den Teilnehmern eingeschätzt werden.

2.11.2.1 Dozent

Um ihr Wissen bestmöglich weitergeben zu können, benötigen Dozenten eine Reihe von Fähigkeiten, die sich auf soziale, wie auch auf pädagogisch- methodische Eigenschaften beziehen.

Eine klare Struktur gilt als zentrales Merkmal guter Lehre, da dadurch der Verständnisprozesses gefördert und die kognitive Strukturierung erleichtert werden. In einer Studie von Weinert²²⁰ konnte ein positiver Zusammen-

²¹⁹ Rindermann H (2001). 64

²²⁰ Weinert FE (1996). 1 – 12

hang zwischen einer klaren Strukturierung und objektiver studentischer Leistung festgestellt werden. Aus Untersuchungen von Rindermann²²¹ ging hervor, dass sowohl aus der Sicht von Teilnehmern – in diesem Fall Studenten –, als auch aus der Sicht von Dozenten, Struktur als wichtigstes Kriterium zur Beurteilung von Lehre empfunden wird.

Die Möglichkeit, den vermittelten Stoff anhand von praktischen Beispielen in unterschiedliche Kontexte einzubinden gilt ebenso als Merkmal guter Lehre. Rindermann bezeichnet diesen dozentenbezogenen Faktor als „Breite/ Auseinandersetzung“. Je besser ein Transfer auf andere Kontexte gelingt, desto leichter fällt die Rezeption, Verarbeitung, Wiedergabe sowie die praktische Umsetzung des Wissens.

Bei der Rezeption von Wissen kann generell nach aktiver und passiver Stoffaufnahme differenziert werden. Während bei der aktiven Stoffaufnahme die Wissensverarbeitung und der Transfer auf ein vertieft werden, tritt bei der passiven Stoffaufnahme eher der Effekt eines schnelleren Vergessens auf. Durch die Einbindung der Inhalte in unterschiedliche Kontexte, gelingt es höhere kognitive Ziele zu erreichen, die Teilnehmer zu aktiver Mitarbeit zu aktivieren sowie ein flexibles Wissensgerüst aufzubauen, das nicht an einen bestimmten Kontext gebunden ist.

Lehrkompetenz ist die Fähigkeit zu lehren. Diese Fähigkeit kann in eine Vielzahl von Merkmalen unterteilt werden. Zum einen zeichnet sich Lehrkompetenz generell durch didaktische Fähigkeiten, Professionalität und Sicherheit in der Vorbereitung oder in der Expertise, komplizierte Fachinhalte verständlich machen zu können aus, andererseits durch rhetorisches Ausdrucksvermögen, klare Artikulation und abwechslungsreiche Intonation.

Weiter gefasst können allerdings sämtliche, den Dozenten betreffende Faktoren, die „gute Lehre“ voraussetzt unter dem Begriff „Lehrkompetenz“ vereint werden.

Ein weiterer Faktor, der wesentlich zur Güte der Lehre beiträgt ist das Engagement des Dozenten, beziehungsweise die Motivationsfähigkeit den einzelnen Teilnehmern gegenüber. Wirkt ein Dozent desinteressiert, überträgt sich diese Einstellung häufig auf die Teilnehmer, was wiederum die Motivation, sich auch selbstständig und außerhalb der Lehreinheiten mit den inhaltlichen Themen auseinanderzusetzen reduziert und somit den Lernprozess beeinträchtigt.

Verkörpert der Dozent authentisches Interesse – beispielsweise durch verbale, nonverbale oder paraverbale Ausdrucksmittel – ist die Leistungsbereitschaft und Motivation der Teilnehmer sowohl innerhalb, als auch im Rahmen von persönlichen Vorbereitungen wesentlich stärker ausgeprägt.

Neben den didaktischen Eigenschaften eines Dozenten, wirken sich auch *soziale Faktoren* wie Freundlichkeit, Kooperation und Kommunikation positiv auf die Motivation und Mitarbeit der Teilnehmer aus. Rindermann fasst diese Faktoren unter den Überbegriffen Kooperativität und Klima zusammen.

Durch den geplanten Einsatz von Interaktionen, z.B. Moderation von Diskussionen oder Förderung von Gruppenarbeiten, kann die Unterrichtseffektivität für die einzelnen Teilnehmer gleichfalls erhöht werden.

²²¹ Rindermann H (1996). 43

Betreuung und Feedback sind nicht nur für den kompletten Kursablauf wichtig; sie sind ein probates Mittel zur Qualitätssicherung der Beiträge im Kurs. Auf sozialer Ebene stärkt Feedback den zwischenmenschlichen Umgang von Dozent und Kursteilnehmer.

Die *wissenschaftliche Fachkompetenz* ist, neben den bereits oben angesprochenen Faktoren, eine zentrale Voraussetzung für Lehrqualität. Dieser Faktor kann nach Rindermann in der Regel nicht durch Studenten beurteilt werden, da der Wissenserwerb erst durch den Dozenten stattfindet. Das Wissen ist somit bei den Teilnehmern noch nicht vorhanden, wodurch eine solche Beurteilung schwer fallen würde.²²²

2.11.2.2 Studierende

„Gute Lehre“ hängt neben dem Dozenten, dem wichtigsten Bedingungsfaktor, auch von den Kompetenzen, Einstellungen und Handlungen der Studierenden ab.

Der Lernprozess kann beispielsweise durch unterschiedliche, ungeeignete Arten von *Vorwissen* negativ beeinflusst werden. Die Ergebnisse vorheriger Bildungserfahrung, die in der Schule, Universität oder im Selbststudium gesammelt wurden, können zu gering oder unangemessen sein, um dem Verlauf des Kurses folgen zu können. In dieser Situation muss entweder der Dozent eingreifen und die Wissensbasis, die als Voraussetzung für den Kurs gilt, schaffen – was für diejenigen, die bereits einen höheren Grad an Vorwissen besitzen unter Umständen als langwierig und demotivierend empfunden werden kann – oder der einzelne Teilnehmer sorgt durch Selbststudium dafür, den inhaltlichen Anforderungen des Kurses genügen zu können.

Gegensätzlich dazu besteht die Möglichkeit, dass Teilnehmer bereits ein zu hohes Maß an Vorwissen besitzen. In diesem Fall können aufgrund der inhaltlichen Anforderungen Störungen auftreten; im Extremfall kann ein zu hohes Maß an Vorwissen zum vorzeitigen Abbruch der Bildungsmaßnahme führen.

Um die negativen Auswirkungen von unadäquatem Vorwissen zu vermeiden, sollte jeder Dozent die Inhalte seiner Bildungsmaßnahmen an die Zielgruppen kommunizieren, die Teilnehmer sollten sich an die Anforderungen halten, beziehungsweise aktiv an der Erfüllung arbeiten.

Der Lehrprozess kann auch nur dann zu günstigen Lernresultaten führen, wenn die Veranstaltungsteilnehmer auf kognitiver und auf nicht- kognitiver Ebene die Voraussetzungen der Teilnahme erfüllen. Das Maß kognitiver Fähigkeiten kann mittels aussagekräftiger, im besten Fall normierter Abschlussnoten (Abiturnote oder Studienabschlussnote) bestimmt werden. Nicht- kognitive Fähigkeiten zeigen sich im Rahmen der Veranstaltung, beispielsweise im Präsentationsverhalten, in der Kommunikationsfähigkeit oder in der Gruppenintegration einzelner Teilnehmer. Nur wenn diese kognitiven und nicht- kognitiven Fähigkeiten als Grundlage ausreichend ausgeprägt sind, können in diesem Bereich die Voraussetzungen für „gute Lehre“ erfüllt sein.

Das *Vorinteresse* – die Beschäftigung mit dem Thema aus eigenem Antrieb – zählt Rindermann „zu den wichtigsten Bedingungs- und Biasvariablen“,²²³ da dadurch die Darstellung der Lehrbeurteilung verzerrt werden kann, obwohl diese Variable keinen direkten Bezug zu den eigentlichen Abläufen der Veranstaltung hat. Für eine

²²² Vgl. Rindermann H (2001). 69

²²³ Ibidem; 70

hohe Lernmotivation und einen erfolgreichen Lernprozess ist das Vorinteresse, das die Teilnehmer der Veranstaltung und den Kursinhalten entgegenbringen essentiell.

Für Lernerfolg sind in der Regel auf studentischer Seite drei Bestimmungsfaktoren notwendig. Abgesehen von den beiden bereits erwähnten, *Fähigkeit* und *Vorwissen*, spielt darüber hinaus die eingebrachte *Lernzeit zur Vor- und Nachbereitung* der Kursinhalte außerhalb der Kurszeiten eine zusätzliche Rolle. Der Fleiß, der seitens der Veranstaltungsteilnehmer zum Wissenserwerb aufgewendet wird, kann die persönliche Arbeitseinstellung und -haltung und somit das Interesse an der Veranstaltung und den Inhalten widerspiegeln.

Der Grad der Beteiligung an den Lehrinhalten hängt auch vom Typ und den Unterschieden der verschiedenen Typen von Lehrveranstaltungen ab. So variiert zum Beispiel die Beteiligung an Veranstaltungen, die mit einer Prüfung abgeschlossen werden – z.B. Seminare – und Vorlesungen, bei denen das erworbene Wissen nicht mittels Prüfung oder Test am Ende abgefragt wird.

Interaktiver Unterricht mit studentischer Beteiligung ist nach Dunkin und Barnes²²⁴ die Voraussetzung für die Entwicklung höherer kognitiver Prozesse, zum Beispiel kritisches Denken und Hinterfragen von Sachthemen. Ebenso wird Wissenserwerb und effektives Lernen stark durch interaktive Auseinandersetzung gefördert (nach Seel)²²⁵.

In vielen studentischen und universitären Veranstaltungen dienen Referate oftmals als Fundament zur Wissensvermittlung im Kurs. Ein Referat soll die aktive und intensive Beschäftigung mit einem speziellen Themengebiet nachweisen und darüber hinaus den Dozenten entlasten. Als Synergieeffekt für den Kursteilnehmer besteht durch Referate die Möglichkeit, Präsentationstechniken und die didaktische Aufbereitung von Lehrinhalten zu trainieren.

Da Referate meist der Kernpunkt studentischer Lehrveranstaltungen bilden, besteht hierbei eine nicht unwesentliche Verantwortung zur Inhaltsvermittlung – was die hohe Korrelation zwischen Qualitätsbeurteilung der Referate und der Beurteilung der Veranstaltung erklärt.²²⁶

Faktoren, die, ausgehend vom Teilnehmer, für „gute Lehre“ hinderlich sind, sind ferner akustische (Unruhe, Gespräche, etc.) und organisatorische *Störungen* (Essen und Trinken im Kurs, Zeitung lesen, mangelnde Pünktlichkeit, usw.), die teilweise allerdings auch durch externe Faktoren (Wetter, öffentliche Verkehrsmittel, ...) bedingt sein können.

Der Umgang mit diesen Störungen ist einerseits von den Rahmenbedingungen (Teilnehmerzahl, Besuchsmotivation, Verbindlichkeit) und andererseits von der sozialen und didaktischen Kompetenz des Dozenten abhängig.

Fehlzeiten von Veranstaltungsteilnehmer können Hinweise auf Unstimmigkeiten bei den Rahmenbedingungen oder beim Lehrverhalten des Dozenten geben.

Abgesehen von den bereits erwähnten Faktoren, besteht darüber hinaus auch eine starke Abhängigkeit von externen – nicht kontrollierbaren – Bedingungen, zum Beispiel Krankheit.²²⁷

²²⁴ Vgl. Dunkin MJ, Barnes J (1986). 392 – 431

²²⁵ Vgl. Seel A (1997). 98 – 113

2.11.2.3 Rahmenbedingungen

Der dritte Hauptfaktor, von dem „gute Lehre“ abhängt, wird über die Rahmenbedingungen einer Veranstaltung definiert.

Vor allem bei universitären Veranstaltungen besteht die Problematik, dass dem Dozenten das *Veranstaltungsthe- ma* häufig durch den Lehrplan oder die Konzeption vorgegeben ist, sodass darauf aus Dozentsicht nur ein geringer Einfluss genommen werden kann.

Ein interessantes und für die Teilnehmer relevantes Thema wirkt sich in der Regel positiv auf die Motivation der Kursteilnehmer aus. Ebenso vom Thema abhängig sind die Mitarbeit und das Engagement in der Veranstaltung. Bei ansprechenden Themen ist das Interesse dem Kursverlauf zu folgen in der Regel höher.

Überschneidungen zwischen unterschiedlichen Veranstaltungen begründen sich vor allem aus einer ungenügenden Curriculumsorganisation und führen meist zu einer Dezimierung der Teilnehmerzahl, beziehungsweise bei obligatorischen Veranstaltungen zu Störungen und der Zunahme von Fehlzeiten. Überschneidungen innerhalb einer Veranstaltung können einerseits den Zweck der Wiederholung und Stoffvertiefung erfüllen, andererseits können solche *Redundanzen* auch dazu führen, die Motivation der Teilnehmer zu senken.

Die Anforderungen, die an die Veranstaltungsteilnehmer in Bezug auf *Stoffschwere*, *Stoffumfang* und *Vermittlungsgeschwindigkeit*, gestellt werden, sind ein wesentliches Indiz zur Beschreibung von Lehrveranstaltung. Beachtlich hierbei ist eine negative Korrelation zur Selbsteinschätzung des Lernerfolgs und der Variable „*Breite / Auseinandersetzung*“ des Lernstoffes. Eine positive Korrelation konnte von Rindermann in HILVE- Studien mit Arbeitshaltung (Fleiß) der Veranstaltungsteilnehmer festgestellt werden.²²⁸ Eine Beziehung zwischen der Anforderungsbeurteilung mit der Person des Dozenten und dem individuellen Lehrverhalten konnte in diesen Studien auch festgestellt werden.

Ein weiterer Faktor innerhalb der Rahmenbedingungen einer Lehrveranstaltung ist die *Anzahl der Teilnehmer*.

Große Teilnehmerzahlen sind kommunikativen Unterrichtsformen nicht zuträglich, da eine große Gruppe beispielsweise moderierte Diskussionen verkompliziert und interaktive Lehrstile aus Gründen der Pragmatik erschwert. Ähnliches gilt für überfüllte Vorlesungen (Platzmangel, Probleme bei der akustischen und visuellen Wahrnehmung). Zwischen der Besuchszahl und Beteiligung konnte Rindermann eine negative Korrelation nachweisen. Hohe Besuchszahlen haben häufig ein geringeres Maß an Beteiligung zur Folge, können allerdings auch ein Indikator für eine hohe Qualität der Lehre sein (genau wie eine geringe Teilnehmerzahl auf mangelnde Qualität schließen lassen kann).

²²⁶ Vgl. Rindermann H (2001). 71

²²⁷ Vgl. ibidem

²²⁸ Vgl. Rindermann H (1997). 216 – 234

Ist eine *Prüfung* Bestandteil einer Veranstaltung, hat das ebenso eine geringere Beteiligung zur Folge, dafür ist das persönliche Engagement der Kursteilnehmer stärker ausgeprägt, die Teilnehmer besitzen ein höheres Maß an Lernmotivation. „*Prüfungen stellen indirekt ein zentrales didaktisches Instrument dar*“²²⁹.

Der *Besuchsgrund* ist - ebenso wie das Vorinteresse – eine wichtige Biasvariable, wodurch die Beurteilung der Lehre verzerrt werden kann, ohne dass ein direkter Zusammenhang mit den eigentlichen Abläufen innerhalb einer Veranstaltung offensichtlich ist.

Wird ein Kurs hauptsächlich wegen einer bestimmten Obligation – zum Beispiel Ablegen einer Prüfung, Anwesenheitspflicht oder der Erwerb eines Zertifikats – besucht, erfolgt durch die Veranstaltungsteilnehmer in der Regel eine kritischere Beurteilung des Kurses als bei Veranstaltungen, die „aus freiem Willen“ und eigenem Antrieb heraus besucht werden.

Der Veranstaltungstyp kann in unterschiedliche Grade an Beteiligung differenziert werden. Hierbei gilt in der Regel, dass in sozial- interaktiven Lehrformen (Seminare, praxisorientierte Kurse, Arbeitsgruppen, etc.) eine Einstellungsänderung und höhere kognitive Ziele leichter erreicht werden können, als in Veranstaltungstypen, die auf monologischen Vortragssituationen basieren (Vorlesungen, Vorträge, etc.).²³⁰

2.11.2.4 Lehrerfolg

Als Zielgröße für die Messung und Beurteilung „guter Lehre“ und Produktvariable aus den Bedingungsfaktoren „Dozent“, „Studierende“ und „Rahmenbedingungen“ gilt der Lehrerfolg.

In diese Variable fließen die Faktoren *Interessantheit einer Veranstaltung*, *allgemeine Veranstaltungsqualität*, *Lerngewinn*, *Wecken von Interesse* sowie der eigentliche *Kompetenzerwerb* mit ein.

Die *Interessantheit* einer Veranstaltung wurde von Rindermann als Effektvariable des Dozentenhandelns und – in etwas schwächerem Maße – des studentischen Vorinteresses konzipiert. Indikatoren hierfür sind abwechslungsreiche Unterrichtsgestaltung, der Wechsel von Sozialformen und Medien sowie das Einbinden des Vorinteresses der Teilnehmer (z.B. bei Praxisbezügen und Beispielen). Die *Interessantheit* einer Veranstaltung hat Einfluss auf die allgemeine Veranstaltungsqualität, den Lernerfolg und die Einstellungsänderungen der Veranstaltungsteilnehmer.

Allgemeine *Veranstaltungsqualität* liefert eine globale, sehr grobkörnige Aussage über die Beschaffenheit von Lehrveranstaltungen – „overall course“ im angloamerikanischen. Vor allem bei der Beurteilung von Dozenten empfiehlt es sich spezielle Dozentenskalen (siehe Kapitel 2.11.2.1) hinzuzuziehen.

Als wichtigstes Merkmal „guter Lehre“ gilt der qualitative und quantitative *Lerngewinn* – also der durch die Veranstaltung intendierte und somit beabsichtigte Wissenszuwachs. Für die Kursteilnehmer ist dieser Faktor ausschlaggebend, da im Lernzuwachs in der Regel der Hauptgrund für den Besuch einer Veranstaltung besteht.

²²⁹ Crooks TJ (1988)

²³⁰ Vgl. Dunkin MJ, Barnes J (1986). 392 – 431

Wünschenswert ist hierbei ein grundlegendes und breiteres Verständnis der Veranstaltungsinhalte und Themenschwerpunkte.

Der Lerngewinn wird dabei auf das einzelne Individuum bezogen (Vergleich des Wissensstand vor der Veranstaltung mit dem Stand nach der Veranstaltung), nicht zwischen den einzelnen Veranstaltungsteilnehmer (A ist besser als B). Hierfür existiert ein geringer Zusammenhang zwischen erreichten Klausurergebnissen und Lernselbsteinschätzungen.

Eine *Einstellungsänderung* führt bei Kursteilnehmer in der Regel zu Motivationssteigerungen, die eine selbstständige Beschäftigung mit den Kursinhalten mit sich bringen. „Gute Lehre weckt Neugier und Interesse“.²³¹

Der *Kompetenzerwerb* verweist auf einen – mittels Prüfung oder Test feststellbaren – Lernerfolg. Darüber hinaus können auch Kompetenzen in nicht explizit den Kursinhalten entsprechenden Bereichen erlangt werden. In der Regel sind dabei nichtkognitive Fähigkeiten und Eigenschaften betroffen, die der Kursteilnehmer durch den Veranstaltungsbesuch indirekt vermittelt bekommt. Soziale Kompetenz, Präsentationstechniken oder die Integration in gruppendynamische Prozesse, sowie Durchsetzungsfähigkeit, verbales, nonverbales und paraverbales Ausdrucksverhalten, der Umgang mit und die Anwendung von Lern- und Arbeitsstrategien können als nicht-fachspezifische Kompetenzen erworben werden. Diese Kompetenzen können generell und losgelöst vom ursprünglichen Kontext eingesetzt werden.

Grundsätzlich ist der Methodeneinsatz dozentenabhängig. Gewisse Einschränkungen, bezogen auf das Kursziel (Seminar, Paukkurs, Vorlesung) und ausgehend von der Organisation und Konzeption der Veranstaltung, sind dabei nur schwer zu vermeiden und führen zu unterschiedlichen Erreichungsgraden der einzelnen Ziele.

Die Itemliste des HILVE- Inventars nach Rindermann ist generell in den einzelnen Hauptfaktoren erweiterbar und kann – je nach projektiertem Einsatz – dementsprechend angepasst werden.²³²

2.11.3 HILVE-II: Dimensionen und Items²³³

Bei HILVE-II handelt es sich um die revidierte und angepasste Form des Heidelberger Inventars zur Lehrveranstaltungsevaluation. Die Struktur der überarbeiteten Form wurde im Rahmen von umfangreichen Studien an der Fachhochschule Nürnberg, der Universität Magdeburg und dem Ausbildungsinstitut für Logopädie in Reichenau untersucht.

Wie bereits die erste Form des HILVE, basiert auch das HILVE-II auf den Grundannahmen der Qualität und Güte von Lehre. Die wesentlichen Bedingungsfaktoren „Dozent“, „Studierende“, „Rahmenbedingungen“ und „Lehrerfolg“ werden dabei in eine Vielzahl von Hauptkriterien integriert, von denen einzelne Items abgeleitet sind.

Im folgenden Teil der Arbeit werden die einzelnen Dimensionen und Items dargestellt und erläutert.

²³¹ Rindermann H (2001). 74

²³² Vgl. ibidem; 73 ff.

²³³ Das komplette Heidelberger Inventar ist befindet sich im Anhang II.

Rindermann empfiehlt für die Einschätzung der einzelnen Items im HILVE-II eine siebenstufige Skala, die ein Spektrum von vollkommener Zustimmung bis vollkommener Ablehnung der vorgegebenen Items erlaubt. Bei der Abfrage einzelner Items wird diese Siebenfachskalierung den entsprechenden Inhalten angepasst (z.B. bei den Fragen zum *Kurstempo*, „viel zu langsam“ bis „viel zu schnell“, oder der *Schwere des Stoffes*, „viel zu leicht“ bis „viel zu schwer“).

Den vier Hauptfaktoren werden im HILVE-II folgende Dimensionen und Items zugeordnet.

Struktur

Struktur beschreibt den Aufbau und die Organisation einer Veranstaltung. In diesem Rahmen kann bewertet werden, ob eine Veranstaltung für die Teilnehmer logisch und nachvollziehbar gestaltet wurde und ob der Teilnehmer mit der Durchführung und der Umsetzung des inhaltlichen Konzepts zufrieden war.

Auseinandersetzung

Der HILVE-II nimmt mit *Auseinandersetzung* Bezug auf die Art und Weise, wie im Rahmen der Veranstaltung die Lerninhalte vermittelt werden. Der Praxisbezug, die Anschaulichkeit der verwendeten Beispiele sowie die allgemeine Relevanz der Inhalte werden in den Items abgefragt.

Verarbeitung

Auch der Punkt *Verarbeitung* bezieht sich in erster Linie auf die Veranstaltungsinhalte. Die Fragestellungen zielen hauptsächlich auf die Vorgehensweise bei der Vermittlung der Lehrinhalte. Es soll hinterfragt werden, ob und inwiefern Veranstaltungsteilnehmer aktiv in den Prozess der Inhaltsvermittlung eingebunden werden und für sie die Möglichkeit besteht, diesen kritisch begleiten.

Lehrkompetenz

Lehrkompetenz beschäftigt sich mit der pädagogischen und methodischen Vorgehensweise des Dozenten. Dabei soll einerseits die „technische“ Seite der Stoffvermittlung – verbaler Ausdruck und eine angemessene Zusammenfassung der Inhalte im Veranstaltungsverlauf – bewertet und andererseits die Fähigkeit, komplexe Themen verständlich vermitteln zu können, eingeschätzt werden. Die Beschäftigung des Dozenten mit den Lehrinhalten im Vorfeld der Veranstaltung und deren intensive Vorbereitung fallen gleichfalls unter Lehrkompetenz.

Dozentenengagement

Die Motivationsfähigkeit eines Dozenten den Teilnehmern gegenüber soll mit *Dozentenengagement* beurteilt werden. Das Engagement, das ein Dozent in die Vermittlung der Lehrinhalte steckt, die Euphorie und Begeisterung bei der Umsetzung, sind für diesen Punkt wesentlich. Die Teilnehmer können zudem bewerten, inwiefern bei ihnen der Eindruck entstanden ist, ob oder wie viel dem Dozenten daran gelegen ist, Lerninhalte an die Veranstaltungsteilnehmer zu vermitteln.

Klima

Im Rahmen einer Lehrveranstaltung bezieht sich *Klima* im Heidelberger Inventar auf die Stimmung, die zwischen den Kursteilnehmern und dem Dozent herrscht und von beiden Seiten beeinflusst werden kann. Die Teilnehmer sollen mit dieser Dimension die Freundlichkeit, Kooperativität und Aufgeschlossenheit des Dozenten bewerten.

Interessantheit

Wie bereits im oberen Teil erwähnt, spielt die *Interessantheit* einer Veranstaltung, und das damit verbundene Interesse der Teilnehmer an den Inhalten, in der Lehrveranstaltungsevaluation eine wesentliche Rolle. Mit *Interessantheit* soll die Art und Weise beurteilt werden, in der die Veranstaltung in ihrer Gesamtheit für die Teilnehmer aufschlussreich und ansprechend durchgeführt wurde. Darüber hinaus kann der Verlauf einer Lehrveranstaltung – bedingt durch die *Interessantheit* – beurteilt werden.

Thema

Die Kategorie *Thema* beinhaltet im HILVE-II einerseits, inwiefern – oder ob überhaupt – sich die Veranstaltungsteilnehmer vor dem Kurs mit den Inhalten beschäftigt haben und ob die Kursinhalte für die Teilnehmer relevant sind. Der Schwerpunkt bei der Bewertung des Themas liegt auf dem Maßnahmenkontext und zielt auf eine von Dozent und Veranstaltung unabhängige Beurteilung.

Redundanz

Redundanz kann, wie im oberen Teil der Arbeit beschrieben, unterschiedliche Folgen für den Kursverlauf und die Teilnehmer haben. Die Teilnehmer sollten in diesem Rahmen bewerten, in welchem Maße ihre themenbezogenen Vorkenntnisse für die Veranstaltung angemessen waren. Auch der Vergleich mit Überschneidungen von anderen Kursen bzw. Modulen fließt in dieses Item mit ein.

Anforderungen

Überforderung oder *Anforderungen* beziehen sich auf vier Schwerpunkte. Die Stoffschwere, den Stoffumfang, das Kurstempo und die Schwierigkeitsstufe der Anforderungen allgemein. Die Teilnehmer haben mit diesen Items die Möglichkeit zu bewerten, ob der Lerninhalt für sie auf dem richtigen Niveau und ob die Geschwindigkeit, in der dieser Inhalt dargeboten wurde, für sie angemessen war oder ob die Stoffmenge noch von ihnen verarbeitet werden konnte. Aus technischer Sicht ist der Optimalwert der Skala für diese Items zentral.

Lernen- quantitativ und qualitativ

Die Qualität und Quantität des *Lernens* wird im HILVE-II separat bewertet. Grundlegend ist hierfür eine Selbsteinschätzung der Teilnehmer, bezogen einerseits auf den Lernumfang, der in der Veranstaltung vermittelt wird, andererseits auf die Verwertbarkeit der vermittelten Informationen. Die Bewertung bezieht sich auf den Vergleich des Wissenstandes *vor der Maßnahme* mit dem Wissenstand *nach der Maßnahme*. Die Beurteilung der Sinnhaftigkeit und Effizienz der Lehre soll in diesen beiden Itemgruppen im Vordergrund stehen..

Betreuung

Betreuung bezieht sich sowohl auf die Veranstaltung selbst, als auch auf das Umfeld.

Im Rahmen der Veranstaltung soll gemessen werden, ob ein Dozent den Veranstaltungsteilnehmern Rückmeldung gibt, beziehungsweise ob diese Rückmeldung für die Teilnehmer nützlich ist. Außerhalb der Veranstaltung bezieht sich die Betreuung auf die Präsenz des Dozenten im Kurs (zum Beispiel bei inhaltlichen und organisatorischen Fragen vor oder während der Veranstaltung).

Referate

Referate sind vor allem im Rahmen studentischer Veranstaltungen ein weit verbreitetes Medium, Lerninhalte in Seminaren zu vermitteln. Kursteilnehmer können mit diesen Items die im Seminar vorgetragene Präsentationen und Referate (bezogen auf den sprachlichen Ausdruck, den inhaltlichen Aufbau, die fachliche Tiefe und Anwendbarkeit der Themen) beurteilen. Darüber hinaus wird die fachliche Unterstützung des Dozenten implizit bewertet. Dieser Aspekt fließt allerdings nicht in die Itembildung mit ein.²³⁴

Fleiß/ Arbeitshaltung

Fleiß und *Arbeitshaltung* wird im Heidelberger Inventar in selbstständige Vor- und Nachbereitung der Veranstaltungsinhalte und in den Aufwand unterteilt, der von den Teilnehmern für diese Vor- und Nachbereitung aufgewendet wird. In diesem Rahmen soll das Engagement, das die Teilnehmer der Veranstaltung entgegenbringen, erhoben werden.

Interaktionsmöglichkeiten

Mit *Interaktionsmöglichkeiten* wird bewertet, inwiefern die Veranstaltungsteilnehmer im Kurs die Gelegenheit haben, sich aktiv am Kursgeschehen (zu Beispiel mit Fragen und Beiträgen) zu beteiligen. Interaktivität bezieht sich hierbei auch auf die Qualität bei der Durchführung und Steuerung von Diskussionen.

Beteiligung

Die *Beteiligung* erhebt im HILVE-II ähnlich wie Fleiß und Arbeitshaltung das Engagement der Veranstaltungsteilnehmer, allerdings nicht mit dem Fokus auf die Vor- und Nachbereitung der Inhalte, sondern auf die Veranstaltung selbst.

Kommunikativen Unterrichtsformen

Kommunikative Unterrichtsformen werden vorwiegend in Seminaren oder kleineren Veranstaltungen verwendet und werden auch vorwiegend in diesem Rahmen erhoben. Anders als bei *Interaktionsmöglichkeiten* soll mit *Kommunikativen Unterrichtsformen* die Möglichkeit beurteilt werden, Diskussionen oder Gruppenarbeiten in einer Veranstaltung abzuhalten. Der Bewertungsfokus liegt vor allem auf der angemessenen Quantität von Diskussionen und dem ausreichenden Einsatz von kommunikativen Lehrformen.

²³⁴ Dieses Item wurde bei der Evaluation von PROFiL-TT entfernt, da Referate im eigentlichen Sinn keinen Bestandteil der Veranstaltung darstellen.

Anomie

Im ursprünglichen Sinne bedeutet *Anomie* ein „Zustand der Gesetz- oder Regellosigkeit“²³⁵. Übertragen auf den Ablauf einer Lehrveranstaltung umfasst Anomie, durch Teilnehmer ausgelöste Störungen – wie zum Beispiel Unruhe im Seminar und vermehrte Abwesenheit.

Interessenförderung

Die *Interessenförderung*, ähnlich wie der Lernerfolg eine wesentliche Zielsetzung im Rahmen von Weiterbildungsveranstaltungen, betrachtet einen globalen Aspekt des Veranstaltungsbesuchs. Im Gegensatz zur *Interessantheit* wird damit die Auswirkung der Veranstaltung auf die allgemeine Beschäftigung mit den Veranstaltungsinhalten bewertet. Eine gelungene Interessenförderung im Rahmen eines Train-the-Trainer Seminars hätte beispielsweise die Auswirkung, dass sich die Veranstaltungsteilnehmer – im vorliegenden Fall selbst Trainer – nach dem Besuch der Maßnahme mehr für trainingsbezogene Inhalte interessieren und diese Inhalte motiviert ein- und umsetzen.

Allgemeineinschätzung

Als Fazit über die gesamte Veranstaltung wird mit *Allgemeineinschätzung* eine Gesamtbeurteilung durchgeführt. Diese Kompletteinschätzung ist weit gestreut und zielt vor allem auf die Gesamtwirkung der Veranstaltung. Fokussiert wird dabei die Einschätzung, inwiefern die Teilnahme an der Maßnahme einen persönlichen Nutzen bewirken konnte. Dieser Nutzen wird mittels einer Gesamtnote bestimmt, die aufgrund der intensiven Reflexion der Inhalte, Dozenten- und Teilnehmer- bezogenen, sowie der Rahmenbedingungen im bisherigen Verlauf der Befragung eine Grobeinschätzung der Maßnahme liefern kann.

Fähigkeiten / Zusatzfragen

Unter dem Punkt *Fähigkeiten* wird im HILVE-II das Ergebnis von bereits abgelegten Tests, Prüfungen – z.B. die Abitur- oder Studienabschlussnote – oder anderen Arten bereits absolvierter Qualifizierungen und Zertifizierungen erhoben. *Zusatzfragen* beinhalten in der Regel Identifikationsmöglichkeit der Teilnehmer, anhand derer die Ergebnisse von Teilnehmern in Quer- oder Längsschnittstudien anonym verglichen werden können. Darüber hinaus bietet sich hier die Möglichkeit biographische und demographische Daten der Teilnehmer abzufragen, die Auskünfte über Kontext- und Hintergrundvariablen geben können.²³⁶

Offene Fragen

Die Teilnehmer haben bei der Beantwortung der *offenen Fragen* die Möglichkeit, quantitative Angaben und Kommentare, Verbesserungsvorschläge und Kritik an der Veranstaltung, sowie an den Erhebungsmethoden und -instrumenten auszudrücken. Diese Möglichkeit ist einerseits für die Teilnehmer wichtig, da dadurch der Freiraum geschaffen wird, unabhängig von den standardisierten Erhebungsmethoden der restlichen Befragung abzuweichen und eigene Inputs zu liefern, andererseits auch für die Erhebung selbst, da mittels qualitativer Aussa-

²³⁵ Martens H (1996).

²³⁶ Vgl. Rindermann H (2001). 387

gen die Daten der quantitativen Erhebung verifiziert und unterstützt, erklärt oder auch falsifiziert werden können.

2.11.4 Anwendung der HILVE-II Dimensionen und Items auf die Evaluation von PROFIL-TT

Mit dem HILVE-II hat Heiner Rindermann ein umfassendes, valides und reliables Inventar zur Beurteilung von Lehrveranstaltungen entwickelt, das eine sehr detaillierte retrospektive Betrachtung und Beurteilung von Lehrveranstaltungen ermöglicht.

Will man die Dimensionen und Items des HILVE-II auf „On-Going-“ und formative Varianten von Evaluationen übertragen, muss das Inventar angepasst werden. Um einen detaillierten Vergleich der Inhalte und Fragestellungen über einen größeren Zeitraum zu erhalten – eine Kombination von prospektiver sowie mehrstufiger retrospektiver Betrachtung bietet sich im vorliegenden Fall besonders an –, ist die Verwendung von synkretistischen Verfahren (vgl. Kapitel 1.8.6) ratsam und wurde in der Evaluation von PROFIL-TT angewandt.

Die Items für die prospektive Befragung wurden dabei vor dem Hintergrund von Einschätzungen, bzw. Erwartungen adaptiert – um eine Vergleichbarkeit mit den retrospektiven Befragungen gewährleisten zu können.

Vereinzelte wurden Items des HILVE-II an das Projekt angepasst. Die Items zu Struktur, Auseinandersetzung, Redundanz, Anforderungen, quantitatives Lernen, Interaktivität, Betreuung, Beteiligung sowie kommunikative Unterrichtsformen wurden in alle drei Erhebungsphasen integriert. Die Möglichkeit zur Eingabe von Anregungen, Anmerkungen oder Kritik als Freitext bestand ebenso in allen drei Evaluationsphasen.

Die Dimensionen Allgemeinschätzung und Dozentenbewertung (auf jedes Modul einzeln bezogen) finden sich lediglich in den beiden Erhebungsphasen nach der Intervention wieder.

3 Projekt PROFiL-TT

3.1 Projektbeschreibung

Das Projekt PROFiL-TT ist eine Weiterbildungsveranstaltung am Institut für Psycholinguistik der Ludwig-Maximilians-Universität München und soll bayerischen Hochschuldozenten („train-the-trainer“) die Möglichkeit bieten, den Horizont ihrer Lehrexpertise zu erweitern und neue, innovative Lehrmethoden – vor dem Hintergrund universitäres Know-how in Wirtschaftsunternehmen zu tragen – vermittelt zu bekommen.

Das Konzept von PROFiL-TT besteht aus mehrtägigen Modulen mit den Schwerpunkten Sprach- und Lehrkompetenz, die von Mitarbeitern des Instituts für Psycholinguistik (unter der Leitung von Prof. Dr. Gerd Kegel) und des Lehrstuhls für Pädagogische Psychologie (unter der Leitung von Prof. Dr. Jochen Gerstenmaier) trainiert werden. Die einzelnen Module Rhetorik, Stimmbildung und Sprecherziehung, Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen, kohärenter Sprachgebrauch und Lehrkompetenz stecken dabei die Meilensteine für eine umfassende Vorbereitung und Durchführung von Weiterbildungsveranstaltungen ab. Aus pragmatischen Gründen findet PROFiL-TT in der vorlesungsfreien Zeit statt.

Die Voraussetzung für die Anmeldung und kostenlose Teilnahme an PROFiL-TT besteht – abgesehen von einer Dozententätigkeit an einer bayerischen Universität oder Hochschule – einerseits in einer bereits aktuellen, bzw. geplanten Tätigkeit als Referent, Dozent oder Weiterbildner und andererseits in einem Auswahlgespräch mit Dozenten und Projektverantwortlichen von PROFiL-TT.

PROFiL-TT ist für die Teilnehmer kostenlos und wird vom Bayerischen Staatsministerium für Wissenschaft, Forschung und Kunst sowie dem Europäischen Sozialfonds gefördert. Die erfolgreiche Teilnahme an PROFiL-TT wird mit einem Trägerzertifikat bestätigt.

Eine geplante Gruppengröße von acht bis zehn Teilnehmern, soll eine individuelle Betreuung der Teilnehmer ermöglichen.

3.1.1 Projekthintergrund

Der Hintergrund für das Projekt PROFiL-TT besteht in dem gewünschten Synergieeffekt zwischen dem wissenschaftlichen Know-how der bayerischen Fachhochschulen und Universitäten einerseits und dem ökonomischen Bedarf an qualitativ hochwertiger Weiterbildung der bayerischen Industrie-, Dienstleistungs- und Verwaltungsmittelpunkte auf der anderen Seite.

Die Wirtschaft in Bayern ist auf diese qualitativ hochwertige Informationsvermittlung angewiesen, um sich auf dem aktuellen Stand der Technik halten zu können und im internationalen Vergleich den Anschluss nicht zu verlieren.

PROFiL-TT will eine fundierte Ausbildung für universitäre Weiterbildner und Dozenten bieten, damit diese Zielgruppe ihr Wissen möglichst effektiv in der Wirtschaft multiplizieren kann. Dieser Effekt soll mit einem hohen wissenschaftlichen Standard in der Vermittlung von Fachinhalten und qualitativ anspruchsvoller Informationsvermittlung verbunden werden.

„Die Fachdozenten und -dozentinnen müssen [...] praxisorientiert auf die spezifischen rhetorischen, pädagogischen und didaktischen Anforderungen bei externen Zielgruppen vorbereitet werden. Darin liegt die Aufgabe von PROFiL-TT.“²³⁷

3.1.2 Modularer Aufbau

PROFiL-TT ist in seiner Gesamtheit modular konzipiert und umfasst zwei wesentliche Teile, Lehrkompetenz und Sprachkompetenz. Diese beiden Teile setzen sich aus einzelnen Schulungsmodulen von unterschiedlichem Umfang zusammen. Die Kombination der Module in den beiden Teilen soll die zentralen Leistungsbereiche speziell auf die Anforderungen der Informationsvermittlung gegenüber externen Kunden zugeschnitten sein. Methodisch wird dabei ein Mix aus fundierter Theorie und angewandter Praxis eingesetzt, wobei der Praxisanteil in jedem Modul einen überwiegenden Anteil beansprucht.

Die beiden Themenschwerpunkte können unabhängig voneinander gebucht werden, sind allerdings inhaltlich stark miteinander verknüpft, sodass sie als Ganzes betrachtet werden können.

Sprachkompetenz

Der Fokus im Teil Sprachkompetenz liegt auf Möglichkeiten zur differenzierten Introspektion und der gezielten Lenkung des Spracheinsatzes – Sprache als Werkzeug der Informationsvermittlung – und ist in Wort und Bild von zentraler Bedeutung für die Vermittlung von Wissen.

„Den Kern der Rhetorik bilden der Redner und seine Wirkung sowie die Entfaltung der Stimme. Gezielter Medieneinsatz und Techniken zur Gestaltung komplexer Vorträge runden den kompetenten und souveränen Auftritt ab.“²³⁸

Der Teil Sprachkompetenz untergliedert sich in die Module *Rhetorik, Stimmbildung und Sprecherziehung, Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen* und *Kohärenter Sprachgebrauch*.

Lehrkompetenz

Im Teil Lehrkompetenz sollen den Kursteilnehmern Techniken und Methoden vermittelt werden, die es ihnen ermöglichen, Lehrveranstaltungen mit weniger *trägem Faktenwissen* und mehr *praktisch einsetzbarem Handlungswissen* zu füllen.

Lehrkompetenz in PROFiL-TT unterteilt sich in die Teilgebiete *Problemorientiertes Lernen, Kooperatives Lernen, Gestaltung von Lernumgebungen* und *Evaluation von Lernumgebungen*. Die Ziele und Inhalte dieser Module werden im folgenden Abschnitt näher erklärt.

²³⁷ PROFiL-TT. <http://www.profil-tt.lmu.de> (Aufgerufen: 03. Januar 2007)

²³⁸ Ibidem

3.1.3 Projektziele

Als umfassendes Projektziel hat sich PROFiL-TT eine Ausbildung gesetzt, aus der eine effiziente und zielorientierte Vermittlung von Informationen resultiert – die Grundlage hochwertiger Weiterbildung.

Bestandteil dieser Zielsetzung ist die Vermittlung von rhetorischen, pädagogischen und didaktischen Fähigkeiten, mittels derer universitäres Fachpersonal befähigt werden soll, Fachinhalte praxisbezogen und kundenorientiert zu vermitteln. Ebenso will PROFiL-TT auf diese Weise das Netzwerk zwischen Universitäten und Unternehmen und die berufliche Flexibilität der Teilnehmer stärken. Hierfür sind die Mitarbeiter fortlaufend bemüht, neue Ausbildungsmodule für PROFiL-TT zu entwickeln und die bereits bestehenden Module den praktischen Umständen entsprechend zu modifizieren.

Darüber hinaus hat PROFiL-TT die Zielsetzung, Kriterien für die Ausbildung von universitären Trainern zu entwickeln und Qualitätsrichtlinien für Lehrpersonal zu formulieren.

Das Fernziel von PROFiL-TT besteht im Aufbau eines Kompetenzzentrums für „universitär fundierte externe Weiterbildung“²³⁹. Das Konzept könnte auf diese Weise generalisiert und an beliebigen Universitäten und Bildungseinrichtungen übernommen werden.

3.1.4 Modulziele – inhaltliche Schwerpunkte

Neben den allgemein gefassten „globalen“ Zielen des gesamten Projekts PROFiL-TT hat jedes einzelne der fünf Module eigenständig abgesteckte Ziele. Im folgenden Abschnitt werden diese Ziele vorgestellt und beschrieben, welche Methoden eingesetzt werden, diese zu erreichen.

Modul Rhetorik

Das Hauptziel des Moduls Rhetorik ist die Förderung sprachlicher Fähigkeiten. Diese Förderung bezieht sich im monologischen Bereich auf die Selbstdarstellung der Teilnehmer. Im Rahmen dieses Moduls sollen die Kursteilnehmer mehr über ihre rhetorische Wirkung anderen gegenüber erfahren und sich ihrer dabei eingesetzten Körpersprache bewusster werden. Außerdem sollen in diesem Modul Kriterien zur Sprechweise, Möglichkeiten, eine monologische Kommunikationssituation einzuleiten und eine – mittels schlüssiger Argumentation geführte – Überzeugung fokussiert werden.

Im dialogischen Bereich soll das Hauptaugenmerk des Moduls Rhetorik auf dem sprachlichen Agieren in Gruppen liegen. Wesentlich hierbei sind auf der rezeptiven Seite, Techniken, mit Fragen umzugehen, bzw. darauf zu reagieren, auf der proaktiven Seite stehen Fragetechniken im Vordergrund. Weitere Schwerpunkte im dialogischen Rhetorikbereich sollen auf den Umgang mit Emotionen im Kursgeschehen gelegt werden, sowie auf rhetorische Lenkungs- und Leitungstechniken – Möglichkeiten ein Gruppengespräch zu beginnen, diese Gespräche zu moderieren, notfalls zu unterbrechen und letztendlich zu beenden.

²³⁹ Ibidem

Generell sollen im Modul Rhetorik „Hinweise zu solider Vorbereitung und erfolgreicher Durchführung von Präsentationen vermittelt“, sowie „die Koordination von Visualisierung, Verbalisierung und sprecherischem Ausdrucksverhalten“²⁴⁰ trainiert werden.

Die dafür gewählte Methodik wird in diesem Modul flexibel, allerdings jeweils zu Beginn angekündigt, deduktiv oder induktiv eingesetzt. Im methodischen Vorgehen sollen die Schwerpunkte auf kurzen klassischen Lehrvorträgen, Rollenspielen – verbunden mit Videofeeds –, generellen Übungen sowie Diskussionen über rhetorische Verhaltensweisen – im Kontext oder personalisiert – liegen.

Um diese Ziele möglichst vielschichtig einhalten zu können, ist vor den einzelnen Modulen eine Einschätzung der Teilnehmer bezüglich ihrer Fähigkeiten, Fertigkeiten und Möglichkeiten notwendig.

Modul Stimmbildung und Sprecherziehung

Das Modul Stimmbildung und Sprecherziehung ist mit einem halben Tag relativ kurz und soll in erster Linie zur Sensibilisierung und Motivation der Teilnehmer dienen. Die Sensibilisierung bezieht sich vor allem darauf, wie jemand spricht – betont, unbetont, akustisch verständlich – und wie mittels stimmlichen Einsatzes Aufmerksamkeit bei der Zuhörerschaft erlangt werden kann.

Ein weiteres Ziel besteht darin, eine Bewusstheit für Stimmhygiene zu schaffen.

Methodisch ist das Modul in zwei Teile untergliedert. Im ersten, analytisch- theoretischen Teil soll die sprecherisch- stimmliche Wirkung anhand von Beispielen untersucht werden, im zweiten, praktischen Teil, soll Übungsraum für die Teilnehmer geschaffen werden, in dem sie die analysierten Techniken und deren Wirkungen testen und anwenden können.

Atmung, Betonung und Stimmwirkung fließen – analog zum Modul Rhetorik – ergänzend mit ein.

Im chronologischen Ablauf – das Modul Stimmbildung und Sprecherziehung folgt im Kursdesign auf den Teil Lehrkompetenz – findet an dieser Stelle ein inhaltlicher und methodischer Bruch statt. In den Teilgebieten der Lehrkompetenz wird versucht, *den Dozent für den Seminarteilnehmer* in den pädagogisch-psychologischen Grundlagen von Lernen und Instruktion „fit“ zu machen, im Modul Stimmbildung und Sprecherziehung hingegen soll der *Kursteilnehmer von PROFiL-TT durch die Seminarinhalte* „fit“ gemacht werden.

Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen

Die Ziele des Moduls Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen beziehen sich auf die Erarbeitung von Kriterien und die Ausbildung von Fertigkeiten für die zielgruppengerechte, professionelle PowerPoint- Präsentation einerseits und andererseits auf die Herausarbeitung von Kriterien für die zielgruppen-gerechte Gestaltung ansprechender Kursunterlagen.

Die Funktionen redebegleitender Visualisierungen und die dabei zu beachtenden Regeln stehen im ersten Teil dieses Moduls im Mittelpunkt. Hierfür sollen die Grundsätze der Präsentation – Ziele, Mittel, Probleme, Erläuterung und die Demonstration verschiedener Präsentationsstile – und die Vorbereitung auf die Präsentation – organisatorische und technische Aspekte – im Rahmen von Gruppenarbeiten und anhand von praxisbezogenen Real-

²⁴⁰ Ibidem

situationen der Teilnehmer untersucht und vermittelt werden. Im Mittelpunkt dieses Moduls soll stets der Bezug auf die Praxis, auf den Alltag der Veranstaltungsteilnehmer und ihre realen Aufgaben- und Problemstellungen stehen.

Weitere inhaltliche Schwerpunkte liegen auf der Gestaltung von Folien, den unterschiedlichen Arten der Hörerführung, der praktischen und kontextbezogenen Anwendung von PowerPoint, sowie der Kombination von PowerPoint mit anderen Visualisierungsmöglichkeiten.

Im zweiten Themenbereich soll vorwiegend auf die Gestaltung der Teilnehmerunterlagen eingegangen werden, auf die Funktionalität und die formalen Möglichkeiten, die hieraus entstehen können, um Kursunterlagen während der Veranstaltung als Unterstützung und nach der Veranstaltung zu Zwecken der Nachhaltigkeit und der Wiederholung einsetzen zu können.

Modul Kohärenter Sprachgebrauch

Das Ziel des Moduls kohärenter Sprachgebrauch besteht in der flexiblen, offenen und skalierbaren Gestaltung von Vorträgen, die einerseits kontextbezogen, andererseits wissensorientiert sein sollen.

Präzisiert wird dieses Ziel durch den Versuch, den Teilnehmern durch Methoden der Rekurrenz Sicherheit zu vermitteln, durch vertikale Integration von Inhalten, Orientierung und durch Unvollständigkeit eine dynamische Spannung zu erzeugen.

In diesem Modul wird versucht, Ansätze und Strategien zu erarbeiten und vorzustellen, mittels derer umfassende und komplizierte Sachverhalte klar und nachvollziehbar erklärt und nachvollzogen werden können. Die Kohärenz bildet hierfür ein wesentliches Strukturmerkmal, da der Zusammenhang der Rede für die Hörerführung ausschlaggebend ist. Der Schwerpunkt in diesem Modul liegt auf der flexiblen Kontrolle der Kohärenz. Die Teilnehmer sollen lernen, durch sprachliche und inhaltliche Analogien der Zuhörerschaft Anhaltspunkte und Hilfestellungen für die leichtere Rezeption von Informationen zu bieten. Ausschlaggebend für eine solche Vernetzung sind „u.a. die Entwicklung von Bezeichnungsspielraum, Techniken der Wiederaufnahme, das Setzen von Orientierungs- und Wendepunkten und Werkzeuge zur Steuerung thematischer Ebenen.“²⁴¹

Methodisch soll im Modul kohärenter Sprachgebrauch vorwiegend mit simultaner Selbstreflexion gearbeitet werden, der Einsatz von Videokameras ist dabei nicht geplant.

Modul Lehrkompetenz

Der Hintergrund des Moduls Lehrkompetenz basiert auf empirischen Befunden aus der Instruktionspsychologie und der Psychologie zum Wissenserwerb. Dieser soll aus der kognitiven Perspektive der Bedeutung des Vorwissens, insbesondere seiner Organisation und Elaboration, betrachtet werden.

Das Modulziel, die Lehrkompetenz in Fort- und Weiterbildung zu fördern, soll unter Verwendung von starken, domänenspezifischen Lernstrategien, mittels der Passung von Konstruktion und Instruktion – je mehr Aktivität von den Veranstaltungsteilnehmern gefordert wird und je individualistischer die Lerninhalte sind, desto besser ist der

²⁴¹ Ibidem

Lernerfolg für die Teilnehmer –, sowie mittels der Einbettung von Wissen in problemorientierte Anwendungskontexte, erreicht werden.

Unter Instruktion wird in diesem Zusammenhang der Aufbau von flexiblen, mentalen Modellen verstanden, die unter anderem cognitive maps – vor allem durch die Visualisierung von Wissen –, propositionale semantische Netze, Induzierung elaborativer Strategien – insbesondere durch ausgearbeitete Lösungsbeispiele (Lückentexte) –, Individualisierung mittels neuer Medien oder Apprenticeship-orientierte Ansätze beinhalten.

Kooperation wird aus der sozialkognitiven Perspektive betrachtet. Sozial geteiltem (shared) und verteiltem (distributed) Wissen wird hierbei wesentliche Bedeutung zugeschrieben.²⁴² Darüber hinaus ist die Strukturisierung durch Ressourcenverteilung, sowie durch Training von Skriptvorgaben in diesem Zusammenhang essentiell. Ein weiterer Schwerpunkt liegt auf zentralen Wirkmechanismen – der Aktivierung von Ko-Konstruktionsprozessen.²⁴³

Das Modul Lehrkompetenz ist in sich in vier zentrale Submodule gegliedert, *problemorientiertes Lernen*, *kooperatives Lernen*, *Gestaltung von Lernumgebungen* und *Evaluation von Lernumgebungen*.

Problemorientiertes Lernen

Der Schwerpunkt in diesem Teil liegt auf dem Transfer von Wissen und problemorientiertem Lernen. Die Probleme, die in Verbindung mit Wissenstransfer auftreten, sollen mittels verschiedener Aufgaben exakt untersucht werden. Beispiele sollen die Kriterien von problemorientiertem Lernen illustrieren und die Anwendung der wichtigsten Elemente zeigen. Methodisch soll dabei analytisch, reflexiv sowie interaktiv vorgegangen werden. Die Ergebnisse der Problemfelduntersuchung sollen vor dem Kurs präsentiert und danach diskutiert werden.

Kooperatives Lernen

Wesentlich in diesem Modul sind die Vermittlung und das Erkennen von Merkmalen kooperativen Lernens. Den Teilnehmern soll die Möglichkeit gegeben werden, kooperatives Lernen aktiv in Lernumgebungen zu integrieren und damit praxisbezogen zu arbeiten. Auch in diesem Teilbereich soll ein Ineinandergreifen von Theorie und Praxis als Basis für die Wissensvermittlung dienen. Gruppenbrainstormings und alltagsbezogene Beispiele sollen diesen Prozess unterstützen und durch gemeinsame Analysen und Erarbeiten von Lösungsansätzen vorantreiben.

Gestaltung von Lernumgebungen

Dieses Modul soll als Schnittstelle zwischen problemorientiertem und kooperativem Lernen fungieren. Wesentlich dabei ist die praktische Implementierung der Inhalte der ersten beiden Module. Hintergrund hierfür sollte der Arbeitsalltag der einzelnen Kursteilnehmer sein.

²⁴² Ausschlaggebend in diesem Rahmen ist gemeinsam gleiches Wissen, als Wissens- und Verständnisbasis zu schaffen.

²⁴³ Z.B. Geben und Erhalten von Erklärungen, Stellen und Beantworten von Fragen, Geben und Erhalten von Feedback.

Evaluation von Lernumgebungen

Im letzten Teil des Moduls Lehrkompetenz sollen innovative Unterrichtsformen bewertet werden. Hierfür werden Bewertungsstrategien erarbeitet, die für Lernumgebungen und Lernformen (bezogen auf Wissen und Lernprozesse) notwendig sind.

Der Teil Lehrkompetenz ist für eine Dauer von eineinhalb Tagen konzipiert. Die Themengebiete in diesem Modul bauen grundsätzlich aufeinander auf und steigen in ihrer inhaltlichen Komplexität an.

Methodisch wird durchgängig durch alle Module die Vorgehensweise *Elaboration*, dann *Präsentation*, danach *Reflexion* und am Ende *Diskussion* angewendet. Diese Vorgehensweise setzt allerdings eine Beschäftigung, bzw. Einarbeitung der Seminarteilnehmer in die entsprechenden Themen im Vorfeld voraus (z.B. durch Einlesen in, im Seminarvorfeld verteilte Texte).

Das generelle Ziel dieses Moduls Lehrkompetenz ist, den Dozent „fit“ zu machen, die Seminarteilnehmer „fit“ zu machen.

3.1.5 Zielgruppe / Teilnahmevoraussetzungen

PROFiL-TT richtet sich in erster Linie an Dozenten, die an bayerischen Universitäten oder Hochschulen Lehrtätigkeiten ausüben und die Absicht haben, diese Tätigkeit in einen ökonomischen Rahmen zu verorten, beziehungsweise auf universitätsexterne Gebiete auszuweiten. Diese Dozenten sollten bereits vor dem Beginn von PROFiL-TT externe Weiterbildungen konzipiert oder durchgeführt haben, zumindest sollte ein konkretes Anliegen, in absehbarer Zukunft eine universitäts- externe Weiterbildungstätigkeit aufzunehmen bestehen. Ein weiteres Auswahlkriterium für die Teilnahme ist der Marktbedarf des entsprechenden Fachwissens des Dozenten. Für den Besuch entstehen dem Teilnehmer keine Kosten, allerdings nehmen alle Teilnehmer vor der endgültigen Aufnahme in das Projekt an einem obligatorischen Auswahlgespräch teil, in dem sie über die Hintergründe, Zielsetzungen und Motivation ihrer Teilnahme an PROFiL-TT Auskunft geben sollen.

Eine Zertifizierung der Teilnehmer findet dann statt, wenn PROFiL-TT regelmäßig besucht (alle fünf Module) wurde, sowie eine Teilnahme an der Evaluation erfolgt ist.

3.1.6 Mitarbeiter

Die Mitarbeiter von PROFiL-TT bestehen, wie bereits im oberen Teil der Arbeit erwähnt, einerseits aus Dozenten und Personal des Instituts für Psycholinguistik, andererseits aus Dozenten vom Lehrstuhl für Pädagogische Psychologie, beide an der Ludwig-Maximilians-Universität München.

Grundsätzlich erfolgt keine fixe Zuordnung im Sinne von Dozent A hält immer Modul A. Die Module sind so konzipiert worden, dass sie dozentenunabhängig – die geeignete Qualifizierung vorausgesetzt – gehalten werden können. Die Projektleitung und -koordination ging vom Institut für Psycholinguistik aus.

Sprachkompetenz

Der Projektleiter *Prof. Dr. Gerd Kegel* lehrt den Einsatz von Präsentationstechniken sowie Gestaltung von Kursunterlagen. Die Forschungsschwerpunkte von Prof. Dr. Kegel liegen auf Spracherwerb, Spracherwerbsstörungen, Sprachentstehung, Sprache und Kognition, Rhetorik, Sprache und Zeitverarbeitung, E-Lernen und E-Kommunikation

Dr. Andreas Hendrich koordiniert das Projekt PROFiL-TT und leitet das Modul *kohärenter Sprachgebrauch*. Die Forschungsschwerpunkte von Dr. Hendrich liegen auf Sprache und Kognition, komplexer Sprachverarbeitung, Hypertext sowie Sprache und Körper.

Cornelia Gericke und *Imme Kerwin* betreuen die Module Rhetorik und Stimmbildung und Sprecherziehung, *Stefanie Vits*, das Modul Rhetorik und *Astrid Wittenberger* Stimmbildung und Sprecherziehung

Lehrkompetenz

Verantwortlich für die Inhalte des Teilgebiets Lehrkompetenz (pädagogisch-psychologische Grundlagen des Lernens) ist *Prof. Dr. Jochen Gerstenmaier*. Die wissenschaftlichen Forschungsschwerpunkte von Prof. Dr. Gerstenmaier sind Psychologie des Lernens und der Instruktion, Wissenserwerb, pädagogisch-psychologische Beratung und darüber hinaus career counseling.

Dr. Susanne Günther und *Marija Kulis* unterstützen Prof. Dr. Gerstenmaier bei der Durchführung dieses Moduls.

3.1.7 Träger / Förderer

Getragen wird das Projekt PROFiL-TT vom Institut für Psycholinguistik zusammen mit dem Lehrstuhl für Pädagogische Psychologie. Die Ansprechpartner sind Prof. Dr. Gerd Kegel als Projektleiter und Dr. Andreas Hendrich als Projektkoordinator.

Die Implementierung der Pilotstudie PROFiL-TT am Institut für Psycholinguistik basiert auf der erfolgreichen Entwicklung und Durchführung des Programms PROFiL – Professionell in der Lehre, einem rhetorisch- pädagogischen Weiterbildungsprogramm, das als Zielgruppe den universitären Dozenten- Nachwuchs – im Gegensatz zu PROFiL-TT allerdings universitätsintern konzipiert – fokussiert.

Die finanzielle Förderung von PROFiL-TT wird größtenteils vom Bayerischen Staatsministerium für Wissenschaft, Forschung und Kunst und dem Europäischen Sozialfonds getragen. Nach den Zyklen im Jahr 2005 endete diese Förderung, worauf das Projekt eingestellt wurde.

3.2 Evaluation PROFiL-TT

Das im letzten Kapitel näher beschriebene Pilotprojekt PROFiL-TT soll im Rahmen dieser Dissertationsarbeit evaluiert werden. Im folgenden Teil werden die Vorgehensweise, die angewandten Methoden und Instrumente, Hintergründe sowie der komplette Ablauf und die Ergebnisse der Evaluation von PROFiL-TT genau erläutert und dargestellt.

In die Evaluation von PROFiL-TT sind verschiedene Ansätze aus unterschiedlichen theoretischen Evaluationsmodellen integriert worden. Der folgende Teil beschäftigt sich mit diesem Methodenmix.

Grundsätzlich wurde als Evaluationsansatz eine Mischung aus *formativen* und *summativen Methoden* (Kapitel 2.8.1 / 2.8.2) gewählt. Die Evaluation generell fand interventionsbegleitend statt, wobei eine Einflussnahme auf den Interventionsablauf und die -inhalte nur bedingt realisierbar war. Diese mangelnde Einflussnahme lag in erster Linie daran, dass das Projekt bereits in der Konzeptionsphase auf ein vorgegebenes Spektrum möglicher Inhalte festgelegt wurde, das – personal- und ressourcenbedingt – nicht beeinflusst werden konnte. Die summativen Gesichtspunkte der Evaluation von PROFiL-TT bestehen vorwiegend aus einer zusammenfassenden Einschätzung der Veranstaltungsteilnehmer zum Nutzen und der Effektivität von PROFiL-TT. Die Evaluation des Pilot-Projekts hatte in erster Linie den Hintergrund, die Effizienz der Maßnahme dem Förderkreis offen zu legen und somit die Zielerreichung zu kontrollieren.²⁴⁴ Die ausschlaggebende Zielsetzung bestand somit darin, *über*, und nicht *an* PROFiL-TT zu berichten.²⁴⁵ Weicht man von der strengen Einteilung in formativ und summativ ab, besteht auch die Möglichkeit, die vorliegende Evaluation askriptiven Ansätzen zuzuordnen, da dem untersuchten Gegenstand eine Wertigkeit sowie ein Nutzen „zuschrieben“ werden soll, allerdings bedingt durch die Konzeption und die Rahmenbedingungen dieser Maßnahme durch die Evaluation weder verändert, noch optimiert werden kann.²⁴⁶

Für die Evaluation wurde zu Beginn der Untersuchung ein *Versuchsplan* erstellt, der die Abläufe und Aktivitäten innerhalb der Evaluation genau reglementiert. In diesem Versuchsplan wurde festgelegt, dass die Evaluation über ihre gesamte Laufzeit (alle Veranstaltungszyklen im Jahr 2005) in drei Stufen (Pre-, Post- und Retrospektions-Erhebung) durchgeführt wird, wann diese Stufen beginnen sollen und bis zu welchem Zeitpunkt die einzelnen Erhebungszeiträume abgeschlossen sein sollten (siehe Kapitel 3.2.3). Als Basis für die Erhebung wurde die revidierte Form des Heidelberger Inventars zur Lehrveranstaltungsevaluation (HILVE-II; Kapitel 2.11) verwendet und an die Rahmenbedingungen von PROFiL-TT adaptiert. Die Untersuchung fand mit diesem Inventar einerseits in der Form einer *Ex-Ante*-Erhebung – in dessen Rahmen die Teilnehmer für die Inhalte der Veranstaltung sensibilisiert, sowie die Erwartungen und Einschätzungen an PROFiL-TT erhoben wurden²⁴⁷ –, andererseits wurde *ex-post* evaluiert – direkt nach der Veranstaltung und retrospektiv, acht Wochen nach Veranstaltungsende.

²⁴⁴ Vgl. Kapitel 2.4

²⁴⁵ „Where a summative evaluation is done [...], the aim is [...] to report *on* it, not to report *to* it.“ (Scriven 1991, 340)

²⁴⁶ Vgl. Kapitel 2.9.17

²⁴⁷ Vor diesem Hintergrund wurde das Heidelberger Inventar zur Lehrveranstaltungsevaluation bezogen auf eine Prospektion – Sensibilisierung, Erwartung bzw. Einschätzung der Veranstaltungsteilnehmer – synkretistisch angepasst.

Diese Dreiteilung der Datenerhebung wurde nach den wesentlichen Merkmalen von Ex-Post-Facto-Designs konzipiert (vgl. Kapitel 2.5.3).

Der Schwerpunkt bei der Evaluation von PROFiL-TT liegt auf der wiederholten Durchführung einer Untersuchung, die einen Vergleich innerhalb eines zeitlichen Verlaufs ermöglicht (*Längsschnittstudie*). Darüber hinaus fanden die Datenerhebungen wie in klassischen *Pretest- Posttest- Designs* vor der geplanten und nach der durchgeführten Intervention statt. Die Zielgruppe veränderte sich über diese Zeitspanne nicht. Es wurden somit für jeden Teilnehmer der Evaluation jeweils eine Erhebung vor, nach und retrospektiv durchgeführt (siehe Kapitel 2.5.3). Die Konstruktion der Erhebungsinstrumente wurde auf diese Vergleichbarkeit der drei Erhebungszeiträume ausgerichtet. Auf diese Weise und aufgrund der Tatsache, dass die Teilnehmer der Evaluation in allen drei Erhebungsstufen identisch sind, können die Ergebnisse der Pre-, Post- und Retrospektionserhebung miteinander verglichen werden.

Eine klare Zuweisung zu einer *internen* oder *externen* Evaluation ist bei dieser Untersuchung nicht eindeutig möglich. Einerseits fand die Evaluation intern statt – die Dissertationsarbeit wird von Institut für Psycholinguistik betreut, an dem ich auch studiert habe und somit Einblicke in die internen Prozesse und Abläufe bekommen konnten –, andererseits war ich zur Zeit der Evaluation nicht am Institut beschäftigt; es bestand kein Anstellungsverhältnis zwischen mir und den involvierten Personengruppen. Der Distanztyp dieser Evaluation lässt sich somit als „intern- externe“ Evaluation beschreiben – auch in diesem Fall eine Mischform.

Bilanzierend handelt es sich bei der vorliegenden Evaluation von PROFiL-TT um eine „intern- externe“ Evaluation, die im Rahmen eines Ex-Post-Facto-Designs mit einerseits formativen, andererseits summativem Hintergrund, basierend auf einem synkretistisch angepassten Inventar zur Lehrveranstaltungsevaluation unter dem Einfluss verschiedener Evaluationsmodelle deskriptiv, mit der Prämisse, die Zielerreichung, Effektivität und Effizienz des Projekts PROFiL-TT einzuschätzen, realisiert wurde. Die Vorgehensweise setzt eine klare Zielorientierung voraus.

3.2.1 Inhaltliche Vorgehensweise bei der Evaluation

Die inhaltliche Vorgehensweise bei der Evaluation des Projekts PROFiL-TT kann im Wesentlichen in drei Aspekten zusammengefasst werden.

- 1) Für die Konzeption und Entwicklung der Erhebungsinstrumente wurde das Heidelberger Inventar zur Lehrveranstaltungsevaluation als Basis verwendet und auf unterschiedliche Weise modifiziert.
- 2) In den Dimensionen und Items der Erhebungsinstrumente werden Einschätzungen und Werturteile der Teilnehmer abgefragt. Es besteht dabei die Möglichkeit, die abgefragten Bedeutungsinhalte subjektiv anzunehmen (zu akzeptieren) oder abzulehnen (z.B. durch Antwortskalen von „trifft nicht zu“ bis „trifft völlig zu“).²⁴⁸

²⁴⁸ Nach Lucke (1995) ist Akzeptanz „die Chance, für bestimmte Meinungen, Maßnahmen, Vorschläge und Entscheidungen bei einer identifizierbaren Personengruppe ausdrückliche oder stillschweigende Zustimmung zu finden und unter angebbaren Bedingungen aussichtsreich auf deren Einverständnis rechnen zu können“ (Lucke

- 3) Die Evaluation wurde als Längsschnittuntersuchung auf drei Ebenen konzipiert, um eine Tendenz der Teilnehmerakzeptanz vor der Maßnahme, direkt danach und acht Wochen nachher, feststellen zu können. Die Vergleichbarkeit der Daten wurde durch die Itemkonstruktion gesichert. Die Prospektion spielt dabei in erster Linie die Rolle, die Teilnehmer auf die Inhalte der Veranstaltung vorzubereiten und sie auf die Kursinhalte zu sensibilisieren – vorab zu informieren. Die Erwartung und Einschätzung der Veranstaltungsinhalte spielt bei der prospektiven Untersuchung vor dem Hintergrund der Vergleichbarkeit eine Rolle und weist Tendenzen bezüglich der Konzeption, bzw. Vorab- Kommunikation und Erwartungshaltung der Teilnehmer auf. Die Post- Erhebung hingegen zielt hauptsächlich auf eine Spontaneinschätzung und Beurteilung der Veranstaltung. Im Rahmen der Retrospektionserhebung besteht für die Teilnehmer die Möglichkeit, den Kurs unabhängig von unmittelbaren Seminareffekten (vgl. Kirkpatrick 2006 Level I: Evaluate Reaction)²⁴⁹ zu betrachten, andererseits allerdings auch, die Veranstaltung noch einmal – nachdem sich die Inhalte festigen konnten – zu reflektieren und dadurch einen Impuls zu geben, die Inhalte Revue passieren zu lassen.

3.2.2 Technische Vorgehensweise – Datenerhebung und -auswertung

Im folgenden Kapitel wird dargestellt, welche technischen Methoden bei der Umsetzung, Implementierung und Datenauswertung der Evaluation von PROFiL-TT angewendet wurden.

Paper & Pencil- vs. Online- Befragung

Vor Beginn der Evaluation stellte sich die Frage, auf welche Weise die Datenerhebung für die drei Evaluationsstufen von PROFiL-TT umgesetzt werden kann und welche Umsetzungsvariante die pragmatischste und effizienteste ist. Grundsätzlich stehen für eine derartige Datenerhebung zwei alternative Verfahren zur Wahl. Es besteht die Möglichkeit, als Instrument paper & pencil oder Online- Befragungen zur Datenerhebung zu verwenden. Die unterschiedlichen Aspekte beider Verfahren wurden vor der Evaluation von PROFiL-TT untersucht und gegenübergestellt. Aufgrund der Konzeption des Projekts, der Beschaffenheit der Zielgruppe und der logistischen Rahmenbedingungen habe ich mich bei der Umsetzung der Evaluation von PROFiL-TT für den Einsatz der Online-Variante entschieden. Diese Entscheidung beruht auf pragmatischen und effizienzorientierten Kriterien.

Die Evaluation von PROFiL-TT fand in drei Phasen statt, wobei sich die Teilnehmergruppe über den gesamten bayerischen Raum verteilt. Im Rahmen einer paper & pencil Befragung hätte jedem Teilnehmer vor Seminarbeginn, direkt nach Seminarendende und retrospektiv acht Wochen nach der Veranstaltung ein mehrseitiger Evaluationsfragebogen zugesendet und von den Teilnehmern auch wieder zurück geschickt werden müssen. Diese Vorgehensweise hätte sowohl die Akzeptanz bei den Teilnehmern, als auch den finanziellen Rahmen des Projekts stark beansprucht. Dieser logistische und finanzielle Aufwand konnte mittels der Online- Variante minimiert

D (1995). 104). Diese Definition umfasst auch das Gegenteil – mangelnde Akzeptanz –; wenn diese Zustimmung nicht vorhanden ist / etwas abgelehnt wird. Vgl. auch Kapitel 2.2

²⁴⁹ Vgl. Kirkpatrick DL / Kirkpatrick JD (2006). 27 – 41

werden, da die Teilnehmer die Evaluationsfragebögen ortsunabhängig – auch direkt nach dem Seminar vor Ort am eigenen Laptop, oder in den CIP- Pools der Universität – ausfüllen konnten. Das Problem, dass ich beim Ausfüllen der Fragebögen für eventuell auftretende Fragen nicht unmittelbar zur Verfügung stand, konnte durch eine telefonische Erreichbarkeit gelöst werden. Die Teilnehmer hatten zudem die Möglichkeit, Fragen oder Unklarheiten zur Erhebung an eine Projektmailadresse zu stellen.

Der ausschlaggebende Vorteil der Online- Umsetzung besteht in der Unmittelbarkeit der Datensammlung. Anders als bei paper & pencil Befragungen fließen die Daten direkt nachdem die Teilnehmer die Fragebögen ausgefüllt haben, in eine Datenbank und stehen dort sofort zur Verfügung. Die Fehleranfälligkeit, die z.B. durch eine händische Dateneingabe oder den Datenimport mittels einer OCR- Software²⁵⁰ in die Datenbank auftritt, konnte durch den Wegfall dieses sehr aufwändigen Zwischenschritts minimiert werden.

Die Projektverantwortlichen hatten darüber hinaus – durch einen eigenen Datenbank- Zugang – zeitnah die Möglichkeit, auf die Rohdaten der Erhebung zuzugreifen und dadurch den Projektfortschritt der Evaluation zu verfolgen und sich ein kontinuierliches Bild über die Ergebnisse zu verschaffen.

Die Teilnehmer erhielten für jede Stufe der Evaluation eine Einladungsemail mit einem Weblink, um den für sie bestimmten Fragebogen im Internet aufrufen und online ausfüllen zu können. Die Einladung für die prospektive Erhebung wurde eine Woche vor Veranstaltungsbeginn an die Teilnehmer verschickt.²⁵¹ Ebenso erhielten die Teilnehmer direkt nach der Veranstaltung eine Email mit einem Weblink für den Nachher- Fragebogen, bzw. acht Wochen nach Seminarende für den retrospektiven Fragebogen.

Die Teilnehmer wurden im Rahmen dieser Email- Kommunikation über Ziele, Abläufe und Hintergründe der Evaluation informiert. Der Evaluationsprozess konnte den Teilnehmern dadurch transparent dargelegt und offen nachvollziehbar gemacht werden.

Umsetzungsvarianten von Onlinefragebögen

Für die Umsetzung des Datenerhebungsinstruments standen wiederum zwei Alternativen zur Auswahl: der Kauf einer vorgefertigten Lösung, oder die Entwicklung einer Individuallösung.

Das Angebot an standardisierten Online- Erhebungstools am Markt ist relativ vielfältig.²⁵² Verschiedene Anbieter verkaufen Lizenzen für Online- Befragungs- Tools, die eine Datenerhebung und -auswertung angenehm und komfortabel gestalten können. Viele dieser Applikationen bieten die Möglichkeit, Umfragen an die individuellen Anforderungen der Nutzer anzupassen. Diese Anpassungen beziehen sich auf Layout, Gestaltung von Antwortskalen und Fragearten oder unterschiedliche Möglichkeiten bei der Datenauswertung. Häufig werden dabei stan-

²⁵⁰ Optical Character Recognition – elektronische Texterkennung; OCR- Software kann zur Erkennung von gescannten Fragebögen eingesetzt werden. Die Software erkennt z.B. welche Kästchen in einem Fragebogen angekreuzt wurden und ordnet diese Ergebnisse Werten in einer Datenbank zu. (vgl. Schmidt H (2002))

²⁵¹ Der Zeitraum von einer Woche wurde deswegen gewählt, weil eine Woche vor Seminarbeginn alle Teilnehmer feststanden.

²⁵² Als Beispiel seien hier Inquiry (<http://www.inworks.de>), FormGen (<http://www.formgen.de>) oder 2ask (<http://www.2ask.de>) erwähnt.

standardisierte Vorlagen zur Verfügung gestellt, die entweder unverändert, oder angepasst eingesetzt werden können. Die Lizenzen für diese Applikationen sind allerdings meist zeitlich befristet und mit – größtenteils erheblichen – finanziellen Belastungen verbunden.

Die Entscheidung für eine selbst entwickelte Lösung fiel aufgrund der zu hohen finanziellen Belastung für das Projekt PROFiL-TT durch eine Kauf- Lizenz, andererseits durch das Interesse und den Anspruch, die Datenerhebung unabhängig von diesen Kauf- Lizenzen individualisiert zu gestalten.

Für die Evaluation wurden drei individuelle, PHP- basierende Datenerhebungsinstrumente²⁵³ programmiert, die genau auf die Bedürfnisse der Befragung zurechtgeschnitten sind. Mit diesem Instrument – einer pragmatischen, und im Vergleich zu den oben genannten Alternativen finanziell günstigen Lösung – ist es möglich, Daten zu erheben und in eine Datenbank zu schreiben, von wo aus sie in weiterverarbeitende Programm exportiert werden können.

Datenauswertung – Datenhandling

Der Prozess, den die Daten von der Eingabe durch die Teilnehmer bis letztlich zur finalen Aufbereitung und Visualisierung durchlaufen haben, wird in mehrere Stufen unterteilt, die für die Daten jeder Erhebungsphase (Vorher-, Post- und Retrospektions- Erhebung) durchgeführt wurden. Bei der Datenauswertung wurden in erster Linie *deskriptive* Methoden verwendet.

In der ersten Stufe laufen die Daten, die in den Online- Fragebogen eingegeben wurden, in eine Oracle- Datenbank. Die Zuordnung der Werte erfolgte dabei analog zu der Aufstellung im Fragebogen. Für jede Frage im Fragebogen steht ein bestimmtes Feld in der Datenbank zur Verfügung, wohin die eingegebenen Werte gespeichert werden können. Der Auszug aus der Datenstruktur (Abb. 2) illustriert die Zuordnung der Einzelwerte in der Datenbank. Z.B. hat der Teilnehmer mit der Nutzer- ID 45 die Frage 9.1 in der Post- Erhebung („Das Kurstempo im Modul Rhetorik war: viel zu langsam=1/genau richtig=4/ viel zu schnell=7; ‚k.A.‘ bedeutet ‚keine Angabe‘“) mit 4 („genau richtig“) beantwortet.

²⁵³ PHP (Hypertext Preprocessor) bietet den Vorteil, mit relativ geringem Aufwand – Programmierkenntnisse vorausgesetzt – dynamische Webseiten mit Datenbankanbindung darstellen zu können; Abbildungen der Datenerhebungsinstrumente befinden sich in Anhang III.

ID	A_7_1	A_7_2	A_7_3	A_7_4	A_7_5	A_8_1	A_8_2	A_8_3	A_8_4	A_8_5	A_9_1	A_9_2	A_9_3	A_9_4	A_9_5	A_10_1	A_10_2	A_10_3	A_10_4	A_10_5
45	4	4	4	4	2	4	5	3	3	3	4	4	4	4	2	4	4	3	3	1
67	3	3	4	4	4	4	4	4	4	5	4	4	4	4	5	3	3	4	4	5
68	4	4	3	5	4	4	4	3	4	4	4	4	4	4	4	4	3	5	4	4
69	4	4	4	4	3	4	4	4	4	3	4	4	4	4	5	5	5	5	5	4
70	4	4	4	4	6	4	4	4	4	6	4	4	4	4	6	4	4	4	4	6
101	4	4	3	4	4	4	4	3	4	4	4	3	4	4	4	4	4	4	4	4
103	4	4	2	4	6	3	4	3	4	5	4	4	3	4	3	4	4	3	4	3
104	4	4	4	4	3	4	4	3	3	2	4	4	3	3	3	4	4	4	3	2
105	4	4	3	4	5	3	3	3	4	5	4	4	4	5	5	4	4	3	5	5
106	4	4	3	4	3	4	3	3	4	4	4	3	4	4	4	3	3	4	5	3
127	4	4	4	4	4	4	4	3	4	5	4	4	4	4	4	4	4	4	4	5
147	4	4	2	5	4	4	4	3	4	5	4	4	3	5	4	4	4	2	5	4
301	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
302	2	2	3	8	3	2	1	3	8	3	4	4	4	8	2	4	4	4	8	4
310	3	2	1	3	3	4	4	1	4	4	4	4	4	4	4	5	4	2	5	4
311	4	8	3	5	4	4	8	4	4	4	4	8	2	4	5	4	8	2	4	4
312	4	3	1	4	3	3	3	6	4	5	4	3	1	4	3	4	3	1	4	3
313	4	3	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4
314	4	4	3	4	4	1	2	5	4	5	7	7	4	5	5	5	5	5	5	5
320	4	4	3	4	3	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4
321	4	4	4	4	8	2	4	4	4	4	4	4	4	4	4	4	4	4	4	8
322	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

Abbildung 2: Auszug aus Datenstrukturmodell

Im nächsten Schritt wurden die Daten aus der Oracle- Datenbank in ein Excel- Template importiert, wo – basierend auf den Rohdaten – Median und Quartilabweichung (siehe Kapitel 2.7.2) für die einzelnen Items berechnet wurden (Abb. 3).

ID	A_7_1	A_7_2	A_7_3	A_7_4	A_7_5	A_8_1	A_8_2	A_8_3	A_8_4	A_8_5	A_9_1	A_9_2	A_9_3	A_9_4	A_9_5	A_10_1	A_10_2	A_10_3	A_10_4	A_10_5
45	4	4	4	4	2	4	5	3	3	3	4	4	4	4	2	4	4	3	3	1
46	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2
51	2	4	3	3	6	6	4	4	5	5	4	4	4	4	4	3	4	3	4	5
66	4	4	3	4	5	3	3	5	4	4	4	3	4	4	4	4	3	4	4	5
67	3	3	4	4	4	4	4	4	4	5	4	4	4	4	5	3	3	4	4	5
68	4	4	3	5	4	4	4	3	4	4	4	4	4	4	4	4	3	5	4	68
69	4	4	4	4	4	3	4	4	4	4	4	4	4	4	5	5	5	5	5	4
71	3	3	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
102	4	4	2	4	5	4	4	3	4	5	4	4	3	4	4	4	4	4	4	5
103	4	4	2	4	6	3	4	3	4	5	4	4	3	4	3	4	4	3	4	103
104	4	4	4	4	3	4	4	3	3	2	4	4	3	3	3	4	4	4	3	2
105	4	4	3	4	5	3	3	3	4	5	4	4	5	5	4	4	3	5	5	105
106	4	4	3	4	3	4	3	3	3	3	4	4	3	4	4	3	4	3	4	106
107	2	2	1	3	2	2	2	3	2	2	3	2	2	3	3	2	2	1	2	1
108	4	2	4	4	3	2	2	5	3	6	4	4	4	4	4	4	4	2	4	2
127	4	4	4	4	4	4	4	3	4	5	4	4	4	4	4	4	4	4	4	5
129	3	4	2	4	4	3	4	2	4	4	4	4	2	4	4	3	4	2	4	4
132	4	4	3	4	3	4	2	3	3	3	3	2	3	3	4	4	4	4	4	3
147	4	4	2	5	4	4	4	3	4	5	4	4	3	5	4	4	4	2	5	4
301	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
310	3	2	1	3	3	4	4	1	4	4	4	4	4	4	4	5	4	2	5	4
313	4	3	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4
314	4	4	3	4	4	1	2	5	4	5	7	7	4	5	5	5	5	5	5	5
315	4	4	1	4	3	3	1	4	3	4	4	1	4	3	4	4	4	1	4	5
316	4	5	5	4	4	4	5	3	4	5	4	6	2	3	2	4	4	4	4	3
319	4	4	1	4	4	3	1	7	3	5	4	4	1	4	4	4	4	1	4	4
320	4	4	3	4	3	4	4	3	4	4	4	4	4	4	4	4	4	4	4	320
Median	4,0	4,0	3,0	4,0	4,0	4,0	4,0	3,0	4,0	4,0	4,0	3,0	4,0	4,0	4,0	4,0	3,0	4,0	4,0	4,0
Quantil 0,25	3,0	3,5	2,0	4,0	3,0	3,0	3,0	3,0	3,0	3,0	4,0	4,0	2,0	4,0	3,0	3,5	4,0	2,0	4,0	3,0
Quantil 0,75	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0	5,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,5	5,0
Quartilabweichung	1,0	0,5	2,0	0,0	1,0	1,0	1,0	1,0	1,0	2,0	0,0	0,0	2,0	0,0	1,0	0,5	0,0	2,0	0,5	2,0

Abbildung 3: Auszug aus der Rohdatenübersicht in Excel mit Median / Quartilabweichung

Nach dem Import der Rohdaten wurden die Daten kumuliert und den einzelnen Antwortkategorien der Fragebögen zugeordnet. Die Daten liegen in dieser Verarbeitungsstufe immer noch in absoluter Form vor (in Abb. 4 hellblau hinterlegt) vor und werden im nächsten Verarbeitungsschritt zu relativen Werten umgewandelt (in Abb. 4 hellgrün hinterlegt).

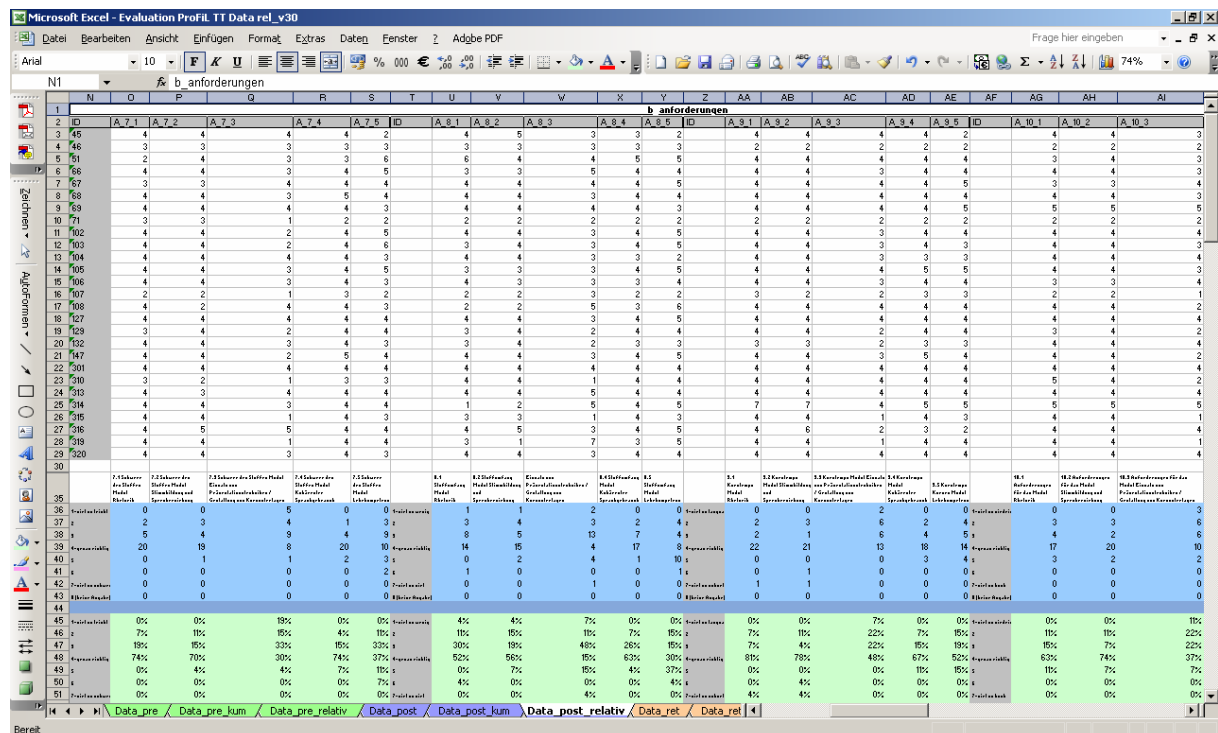


Abbildung 4: Auszug aus den kumulierten Daten / Relativ- Werten in Excel

Im letzten Schritt der Datenverarbeitung wurden die erhobenen Werte mittels Diagrammen visualisiert, wobei die drei Erhebungsphasen nebeneinander dargestellt werden, um die Daten übersichtlich vergleichen zu können (siehe Kapitel 3.2.4).

Kontrolle der Datenqualität

Die Basis für jedes verlässliche und valide Untersuchungsergebnis bildet ein konsistentes und nachvollziehbares Datenhandling. Fehler innerhalb der Datenverarbeitungskette führen in der Regel auch zu Ungenauigkeiten bei der Dateninterpretation und beim Untersuchungsergebnis.

Da es bei der Datenerhebung und -auswertung, sowie durch die Komplexität der Datenbankstruktur vorkommen kann, dass die Datenqualität nicht vollständig konsistent ist – z.B. durch Fehler bei der Kodierung, beim Data-handling oder der Vergabe von Variablen- und Identifikationscodes –, ist es generell ratsam und unverzichtbar, die Datenbasis fortlaufend auf derartige Fehler und Inkonsistenzen zu kontrollieren und diese zu bereinigen.²⁵⁴

Um solche Fehler generell zu vermeiden – oder zumindest die Häufigkeit zu minimieren – empfiehlt sich eine klare und übersichtliche Datenstruktur, sowie eine durchgängige Dokumentation der Datenverarbeitungsschritte, begonnen bei den erhobenen Rohdaten bis hin zu den fertig aufbereiteten Datensätzen. Zeichnen sich bei den

²⁵⁴ Vgl. Wottawa H, Thierau H (2003). 150 f.

Evaluationsergebnissen untypische, besonders herausragende oder unstimmgige Werte ab, empfiehlt es sich, diese Werte gesondert zu betrachten und den Datenverarbeitungsprozess bis zum Beginn der Verarbeitungskette zurückzuverfolgen und zu überprüfen.

Eine Hauptfehlerquelle bei der Datenauswertung besteht in der Verwechslung von Variablennamen bzw. Variablennummern. Häufig werden die Variablen durch fehlerhafte oder unterschiedliche im Umlauf befindliche Versionierungen vertauscht.

Durch eine regelmäßige Kontrolle der Daten und Konsistenzprüfungen kann das Risiko von Fehlern in der Datenauswertung minimiert, wenn auch nicht vollkommen ausgeschlossen werden.²⁵⁵

Im Rahmen der Evaluation von PROFiL-TT wurde die Datenbasis regelmäßig mit Hilfe von Excel- Makros auf Konsistenz überprüft. Zur Kontrolle der Datenqualität sind sowohl innerhalb jeder Verarbeitungsstufe, als auch bei jeder Datentransformation in Excel, vergleichende Sicherheitsmechanismen integriert worden, durch die – referenzierend auf die Rohdatenbasis – die Datenkonsistenz überprüft wurde.

3.2.3 Zeitmodell – Untersuchungsplan

Die hier vorgestellte Evaluation von PROFiL-TT hat sich insgesamt über den Zeitraum von Oktober 2004 bis Februar 2006 erstreckt.

Die Konzeptions- und Abstimmungsphase für den Aufbau und die Inhalte der Evaluation haben bis einschließlich der fünften Kalenderwoche 2005 gedauert. Im Anschluss daran fand die Evaluation in drei Stufen für jeden Veranstaltungstermin statt.

Die Evaluation fand für die beiden Zyklen Sprachkompetenz und Lehrkompetenz parallel statt.

Die Seminarteilnehmer hatten die Möglichkeit, die Teile Sprach- und Lehrkompetenz mit zeitlichem Versatz zu besuchen, diese Möglichkeit wurde allerdings – aus organisatorischen und inhaltlichen Gründen – von der Seminaradministration nicht empfohlen; zwei Drittel der untersuchten Population haben beide Schwerpunkte unmittelbar nacheinander besucht. Bei den restlichen Teilnehmern lag der Abstand zwischen den beiden Teilen zwischen zwei und fünf Wochen. Aus technischen Gründen war es allerdings in diesen Fällen nicht möglich, die Bewertung der Module zu separieren.

In der ersten Stufe (Diagramm 1; Erhebungszeitraum Vorher; hellgrün) wurden die Teilnehmer über den Ablauf der gesamten Evaluation informiert und gleichzeitig zum Ausfüllen der Vorabbeurteilung / Erwartungsabfrage / Sensibilisierung eingeladen. Der Link zum Beurteilungsbogen wurde eine Woche vor Seminarbeginn via Email verschickt. Die Teilnehmer wurden gebeten, die Pre- Evaluationsbogen bis *vor* Seminarbeginn auszufüllen, um Seminareffekte von vornherein ausschließen zu können.

Als nächste Stufe wurden die Teilnehmer zur Post- Evaluation direkt nach dem Seminar eingeladen (Diagramm 1; Erhebungszeitraum Nachher; dunkelblau). Im Rahmen dieser Evaluationsstufe ist der Seminareffekt (Kirk-

²⁵⁵ Vgl. *ibidem*; 153 f.

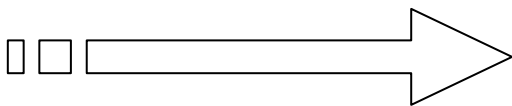
patrick 2006; Kapitel 3.2.1) häufig sehr stark ausgeprägt. Die Ergebnisse, die aus dieser Evaluationsstufe gewonnen wurden, werden vor diesem Hintergrund ausgewertet.

Die Einladung für die dritte – und letzte – Evaluationsstufe (Diagramm 1; Erhebungszeitraum Retrospektiv, gelb) erhielten die Teilnehmer acht Wochen nach Seminarende. Mittels dieser retrospektiven Seminarbetrachtung hatten die Teilnehmer die Möglichkeit, die Seminarinhalte und alle seminarbedingten Einflüsse aus einem Abstand von acht Wochen zu reflektieren. Nach dieser Zeit konnten die Seminarteilnehmer das Seminar, die Dozenten und die Rahmenbedingungen distanziert beurteilen.

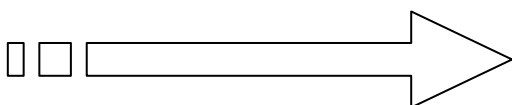
Diese drei Stufen wurden für jeden Zyklus des Projekts durchgeführt.

Die Datenauswertung fand jeweils parallel zur Datenerhebung statt. Reports, Zwischenberichte und Informationen an die Seminarorganisation wurden nach Bedarf erstellt.

Kalenderwochen 2005	bis KW05	KW06	KW07	KW08	KW09	KW10	KW11	KW12
Tätigkeiten								
Konzeptions- und Abstimmungsphase								
Modul Lehrkompetenz			14.-15.02.05		28.02.-01.03.05	07.-08.03.05		
Modul Sprachkompetenz			15.-18.02.05		01.-04.03.05	08.-11.03.05		
Erhebungszeitraum Vorher		für KW07		für KW09	für KW10			
Erhebungszeitraum Nachher				für KW07		für KW09	für KW10	
Erhebungszeitraum Retrospektiv								
Reports / Zwischenberichte								
Datenauswertung								



KW13	KW14	KW15	KW16	KW17	KW18	KW19 - KW21	KW22	KW28	KW29
									18.-19.07.05
	05.-08.04.05								19.-22.07.05
								für KW29	für KW30
		für KW14							
		für KW07		für KW09	für KW10		für KW14		



KW30	KW31	KW32 - KW36	KW37	KW38	KW39	KW40	KW41 - KW45	KW46	KW47
25.-26.07.05				19.-20.09.05					
26.-29.07.05				20.-23.09.05	27.-30.09.05				
			für KW38						
für KW29	für KW30				für KW38	für KW39			
			für KW29	für KW30				für KW38	für KW39

Diagramm 1: zeitliche Ablaufplanung der Evaluation PROFiL-TT

3.2.4 Ergebnisse der Evaluation von PROFiL-TT

Im folgenden Kapitel werden die Ergebnisse der Evaluation von PROFiL-TT deskriptiv dargestellt und erläutert. Grundlegend für die Darstellung werden die einzelnen Dimensionen der Evaluation und die darauf bezogenen Items angeführt.

Die einzelnen Diagramme enthalten die Daten der Vorher-, Nachher- und Retrospektionsuntersuchung. Zusätzlich werden für alle itembezogenen Diagramme Streuungsmaß (Quartilabweichung) und als Mittelwert der Median angegeben.

Die Darstellungsform der Diagramme ist prozentuell und veranschaulicht den Prozentsatz der Untersuchungsteilnehmer, die die unterschiedlichen Dimensionen in den verschiedenen Ausprägungen bewertet haben. Da die Grunddaten ordinalskaliert sind, werden Median und Quartilabweichung absolut auf die Antwortdimensionen bezogen.

Im Rahmen der Vorheruntersuchung hatten die Teilnehmer die Möglichkeit, ihre veranstaltungsbezogenen Erwartungen und Einschätzungen zu beurteilen.

Wesentlich bei der Vorheruntersuchung war die Sensibilisierung auf einerseits die Seminarinhalte – als Preview –, andererseits sollte durch die Vorherbefragung der Rahmen der Evaluation abgesteckt werden, als Orientierung der Teilnehmer über den Ablauf und die Inhalte der Evaluation. Diese Orientierung wurde bereits im Rahmen der Einladung zur Evaluation formuliert und explizit gemacht.

Aus pragmatischer Sicht dienen die Ergebnisse der Vorheruntersuchung als Referenzpunkt für die folgende Nachher- und Retrospektionsuntersuchung.

Enthaltungen bei Fragen („keine Angabe“) fließen zwar mit in die deskriptive Beschreibung, nicht allerdings in die Berechnung von Median oder Quartilabweichung mit ein.

Der in den Untersuchungsergebnissen hinterlegte Farbverlauf (rot – grün) soll die Akzeptanz und Ausprägung der abgefragten Items unterstreichen. Die Farbabstufung verläuft von rot (ungünstigste Ausprägung) bis grün (günstigste Ausprägung) des jeweiligen Items. Items, bei denen die Skala als günstigsten Wert den zentralen Wert zeigt, haben den Farbverlauf rot – grün – rot.

Die einzelnen Diagramme sind grundsätzlich folgendermaßen farblich codiert:

- Die hellgrüne Codierung bezieht sich immer auf die Ergebnisse der prospektiven Daten;
- Die blaue Codierung bezieht sich auf die Daten der Posterhebung;
- Die hellgelbe Codierung bezieht sich auf die Daten der retrospektiven Erhebung.

3.2.4.1 Eigenschaften der Population

Die untersuchte Population setzt sich aus den Teilnehmern von PROFiL-TT zusammen, die beide Teile – Sprachkompetenz und Lehrkompetenz – der Veranstaltung besucht und darüber hinaus an allen drei Stufen der Evaluation teilgenommen haben.

Die Geschlechtsverteilung innerhalb der Population liegt bei 56% Teilnehmer und 44% Teilnehmerinnen. Die Populationsgröße beträgt insgesamt 27 Teilnehmer.

Im Rahmen der demographischen Daten wurden die Teilnehmer gebeten, Aussagen über die Häufigkeit ihrer aktuellen und zukünftig geplanten Trainertätigkeit zu machen (*Diagramm 2*), sowie darüber, auf welchen Wegen sie von PROFiL-TT erfahren haben (*Diagramm 3*).

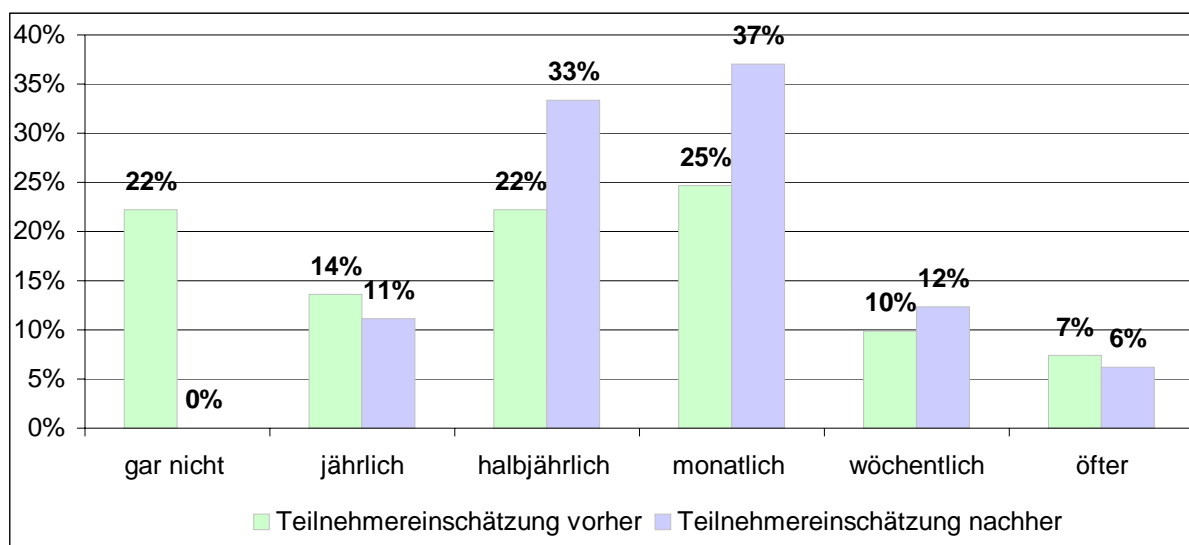


Diagramm 2: Aktuelle Trainingshäufigkeit im Vergleich zur zukünftigen

Diagramm 2 zeigt, wie häufig die Teilnehmer bis zum Zeitpunkt des Seminars, Trainings oder generell Weiterbildungen gehalten haben (grüner Balken). Dem gegenübergestellt wird die Einschätzung der Teilnehmer, wie oft sie nach der Veranstaltung Trainings oder Lehrveranstaltungen planen, Weiterbildungsveranstaltungen zu halten oder durchzuführen (blauer Balken).

Aus der Aufstellung wird sichtbar, dass ca. drei Viertel der Teilnehmer bereits vor Seminarbesuch – mehr oder weniger häufig – Trainings durchgeführt haben. Der Trend in dieser Verteilung geht grundsätzlich dahin, dass die meisten Teilnehmer die Veranstaltung vor dem Hintergrund einer zukünftig zunehmenden Trainertätigkeit besucht haben. Der Schwerpunkt liegt zwischen monatlicher und halbjährlicher Trainertätigkeit; alle Teilnehmer

haben die Perspektive, nach der Veranstaltung Seminare zu halten – auch die, die bis vor PROFiL-TT noch keine Trainings oder Kurse gegeben haben.

3.2.4.2 Publizitätsmaßnahmen

Im *Diagramm 3* wird dargestellt, mittels welcher Medien oder auf welchen Wegen die Kursteilnehmer auf das Projekt PROFiL-TT aufmerksam (gemacht) wurden. Gesteuerte Public Relation für PROFiL-TT bezog sich auf die beiden Schwerpunkte, Flyer und online über die Website (www.profil-tt.lmu.de). Dennoch hat die Hälfte der Teilnehmer von PROFiL-TT über Kollegen oder Mitarbeiter erfahren („Mundpropaganda“ als probates Werbemittel); jeweils ca. ein Viertel wurde auf PROFiL-TT über Flyer oder das Internet aufmerksam.

Die persönliche, zwischenmenschliche Kommunikation stellt sich dadurch als effektivste Werbemaßnahme für das Projekt PROFiL-TT heraus.

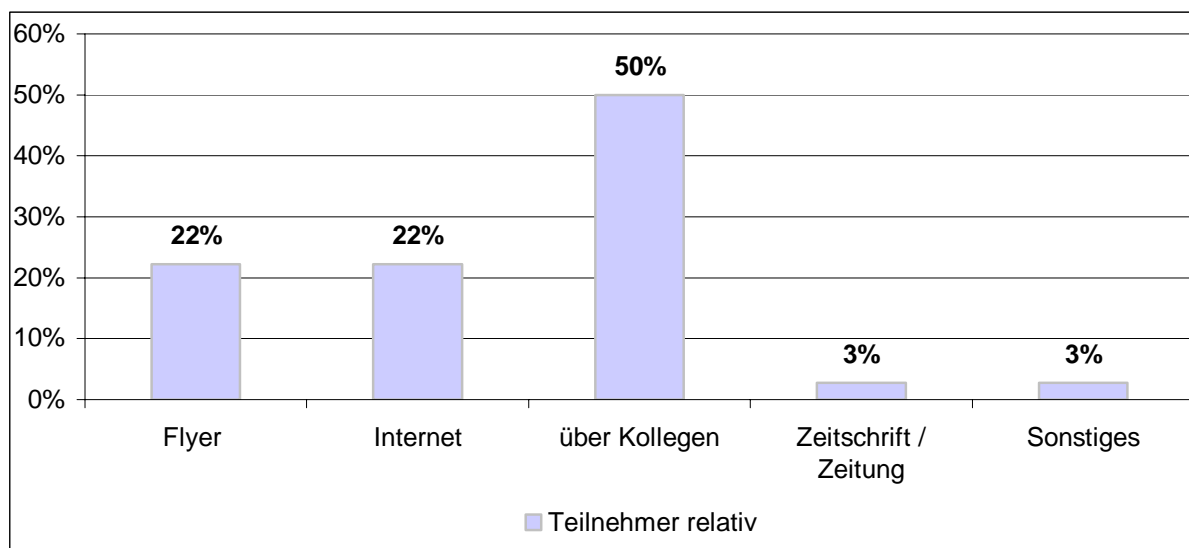


Diagramm 3: Publizitätsmaßnahmen von PROFiL-TT

3.2.4.3 Struktur

Die Dimension Struktur beschreibt den Aufbau und die Organisation der Veranstaltung. In diesem Rahmen wird bewertet, ob eine Veranstaltung für die Teilnehmer logisch und nachvollziehbar gestaltet wurde und ob der Teilnehmer mit der Durchführung und der Umsetzung des inhaltlichen Konzepts zufrieden war. In der Vorheruntersuchung sollen die Teilnehmer ihre Erwartungen, bezogen auf Aufbau und Struktur sowie Organisation der Veranstaltung einschätzen.

Aufbau

Untersuchungsergebnisse Aufbau

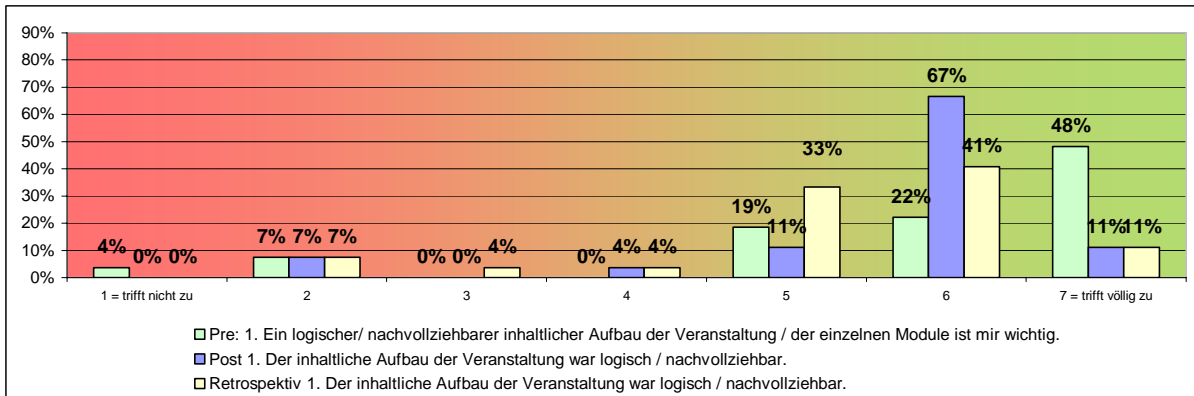


Diagramm 4: Untersuchungsergebnisse Evaluation PROFiL-TT Aufbau

Analyse / Ergebnisinterpretation Aufbau (siehe Diagramm 4)

Betrachtet man die Einschätzung und Erwartungen der Teilnehmer vor der Veranstaltung, kann man feststellen, dass ein logischer und klar nachvollziehbarer Aufbau von einem hohen Prozentsatz der Teilnehmer erwartet und gewünscht wird. Der Median liegt dementsprechend hoch (6 „trifft fast völlig zu“). Die Quartilabweichung bei der Vorabbefragung mit dem Wert 2 zeigt allerdings eine geringfügige Quartilabweichung im Bereich einer etwas stärker ausgeprägten Zustimmung (etwas größere Zustimmung als „trifft teilweise zu“) bis hin zur vollkommenen Zustimmung („trifft völlig zu“). Der relativ hohe Median zieht sich mit dem Wert 6 auch durch die Post- und Retrospektionserhebung. Bei der Post- Erhebung beträgt die Quartilabweichung 0. Die Werte konzentrieren sich hier auf dem Niveau einer sehr hohen Zustimmung („trifft fast völlig zu“). Bei der Retrospektionsbefragung ist die Quartilabweichung (1) wieder etwas größer, was vermutlich mit der konsolidierten Wahrnehmung nach acht Wochen zusammenhängt.

Organisation

Untersuchungsergebnisse Organisation

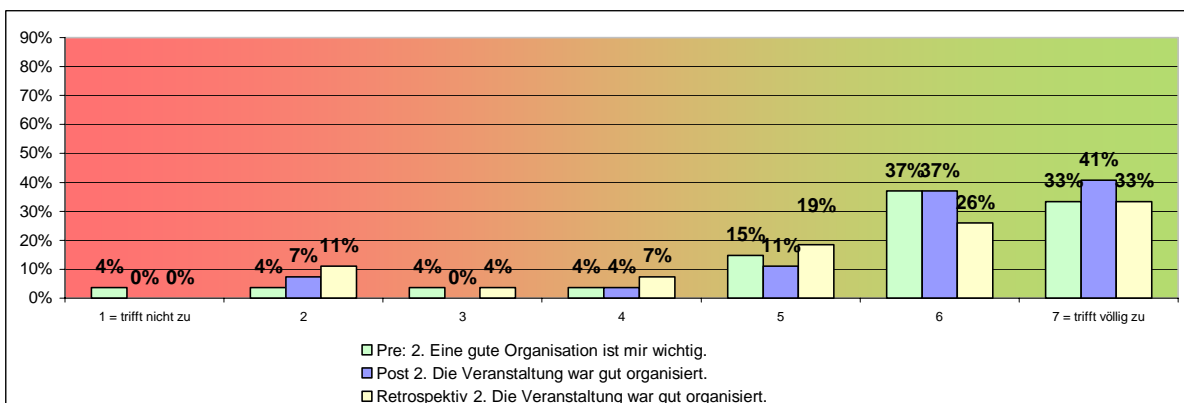


Diagramm 5: Untersuchungsergebnisse Evaluation PROFiL-TT Organisation

Analyse / Ergebnisinterpretation Organisation (siehe Diagramm 5)

Die Frage nach der Organisation der Veranstaltung (vorher: Einschätzung, nachher und retrospektiv: Empfindung) liefert ebenfalls ein eindeutiges Ergebnis. Bei einem Median von 6 und einer Quartilabweichung von 2 in der Vorherbefragung kann eine für die Teilnehmer hohe Relevanz einer guten Organisation festgestellt werden. Die Erwartung wurde von der Organisation sowohl in der Nachherbetrachtung als auch in der Retrospektion erfüllt (Median in der Post- und Retrospektions- Befragung: 6 bei einer Quartilabweichung von 1 und 2).

3.2.4.4 Auseinandersetzung

Diese Dimension nimmt Bezug auf die Art und Weise, wie im Rahmen der Veranstaltung die Lerninhalte vermittelt werden. Der Praxisbezug sowie die Anschaulichkeit der verwendeten Beispiele werden in den Items abgefragt. Bei der Vorher- Erhebung wird untersucht, welche Präferenzen die Teilnehmer bezüglich der Aufteilung von Theorie und Praxis, bzw., welche Erwartungen sie gegenüber der Praxisrelevanz von anschaulichen Beispielen in der Veranstaltung haben.

Veranschaulichung der Veranstaltungsinhalte anhand von Beispielen

Untersuchungsergebnisse Veranschaulichung

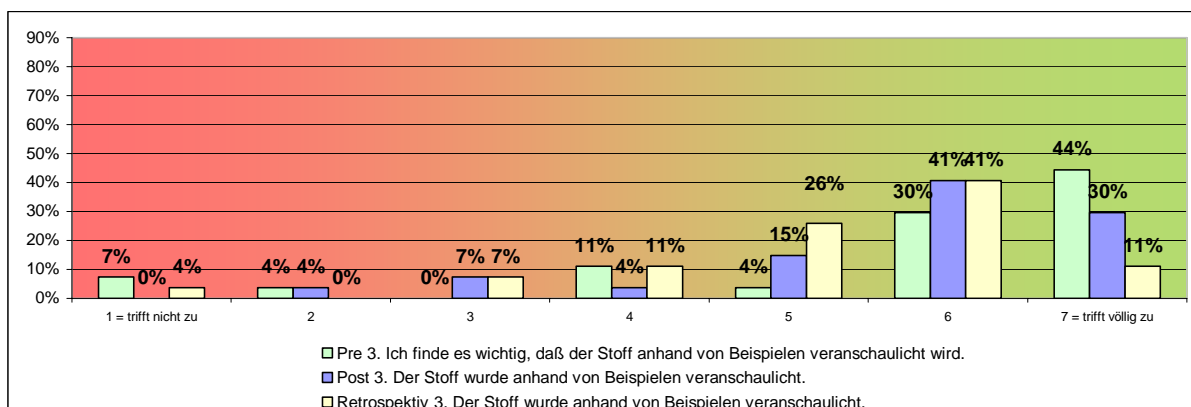


Diagramm 6: Untersuchungsergebnisse Evaluation PROFIL-TT Veranschaulichung

Analyse / Ergebnisinterpretation Veranschaulichung (siehe Diagramm 6)

Die Erwartung, die die Kursteilnehmer vorab an die Veranschaulichung des Lehrstoffes (anhand von Praxisbeispielen) stellen, ist hoch. Bei einem Median von 6 mit einer Quartilabweichung von 1,5 liegen dreiviertel der Teilnehmer- Erwartungen in einem Bereich stärkerer Zustimmung (Wert größer als 4). Dieses Bild kann auch auf die Post- und Retrospektionserhebung übertragen werden (jeweils ein Median von 6), allerdings ist die Quartilabweichung in der Nachherbefragung größer (Wert 2), was vermutlich daran liegt, dass hier der Prozentsatz derjenigen Teilnehmer die einer beispielhaften Stoffveranschaulichung voll und ganz zustimmen mit 30% schwer ins Gewicht fällt.

Bezug zwischen Theorie und Praxis

Untersuchungsergebnisse Theorie- Praxis- Bezug

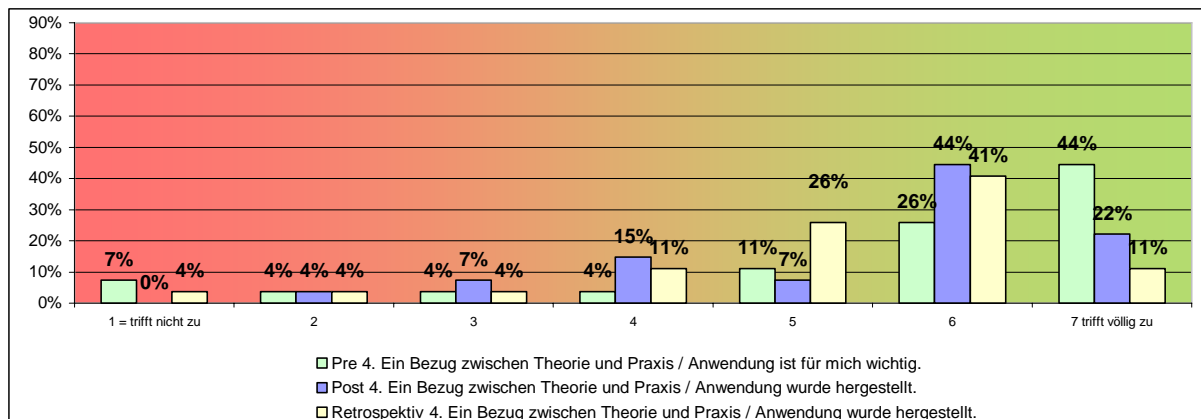


Diagramm 7: Untersuchungsergebnisse Evaluation PROFIL-TT Theorie-Praxis-Bezug

Analyse / Ergebnisinterpretation Theorie- Praxis- Bezug (siehe Diagramm 7)

Der Bezug zwischen Theorie und Praxis wurden von einem Großteil der Teilnehmer (der Median in der Vorherbefragung liegt bei 6, mit einer Quartilabweichung von 2) als wichtig eingeschätzt. Diese Einschätzung taucht auch in der Beurteilung nach der Veranstaltung und in der Retrospektion wieder auf. Die Frage, ob ein Theorie-Praxis-Bezug in der Veranstaltung stattgefunden hat, wurde nach der Veranstaltung von den meisten Teilnehmern als zutreffend beantwortet. (Median post / retrospektiv: 6 bei einer Quartilabweichung von 1,5 / 1). Die Verknüpfung von Theorie und Praxis wurde von über 70% (Posterhebung), bzw. über 80% (Retrospektion) als vorhanden eingeschätzt.

3.2.4.5 Redundanz

Redundanz kann sich auf den Kursverlauf und die Mitarbeit und Motivation der Veranstaltungsteilnehmer unterschiedlich auswirken. Die Teilnehmer sollten in diesem Rahmen einschätzen, in welchem Maße ihre themenbezogenen Vorkenntnisse für die Veranstaltung angemessen waren.

Vorwissen Modul Rhetorik

Untersuchungsergebnisse Vorwissen Modul Rhetorik

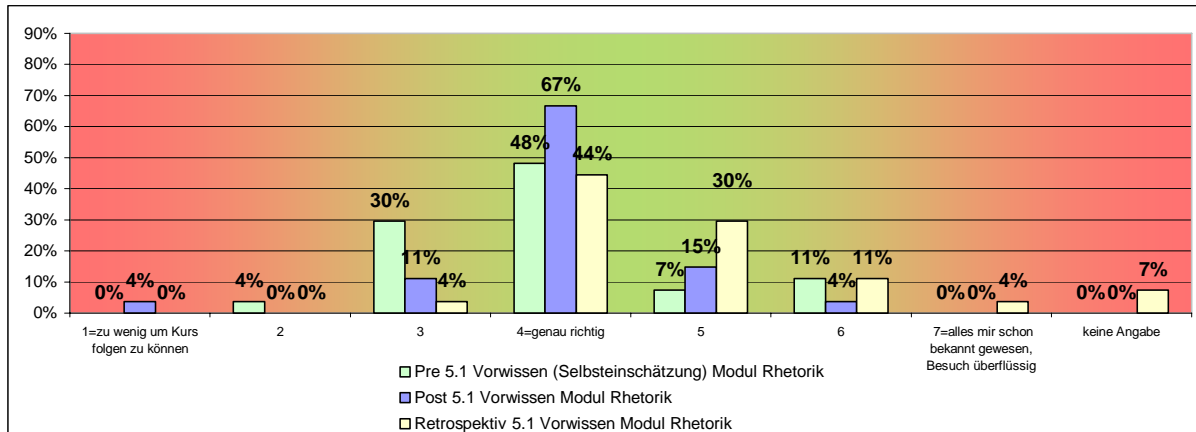


Diagramm 8: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Rhetorik

Analyse / Ergebnisinterpretation Vorwissen Modul Rhetorik (siehe Diagramm 8)

Die Skala in der Dimension „Akzeptanz“ hat ihre maximale Ausprägung bei 4. Die Einschätzung, bzw. Bewertung der Güte des eigenen Vorwissens im Modul Rhetorik zentriert sich auf diesen Maximalwert. Der Median in der Vorher-, Nachher- und Retrospektionsbefragung liegt jeweils bei 4. Auffällig in der Nachher- Erhebung ist die Quartilabweichung von 0. Fast alle Teilnehmer beurteilen ihre Vorkenntnisse im Modul Rhetorik als „genau richtig“. In der Vorab- und Retrospektionsbefragung misst die Quartilabweichung jeweils 1. Der Unterschied zwischen diesen beiden Erhebungen besteht allerdings in einer leichten Verlagerung der Ergebnisse vor der Veranstaltung und in der retrospektiven Betrachtung. Vor dem Seminar waren 30% der Teilnehmer der Ansicht, dass ihre Vorkenntnisse etwas zu gering sind, um den Inhalten im Modul Rhetorik folgen zu können, in der retrospektiven Betrachtung wird diese Einschätzung gespiegelt, womit 30% der Teilnehmer eher in die Richtung tendieren, dass ihre Vorkenntnisse stärker ausgeprägt waren. Die Ergebnisse der Vorher- Befragung gehen vermutlich mit der Erwartungshaltung der Teilnehmer an die Kursinhalte einher.

Vorwissen Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Vorwissen Modul Stimmbildung und Sprecherziehung

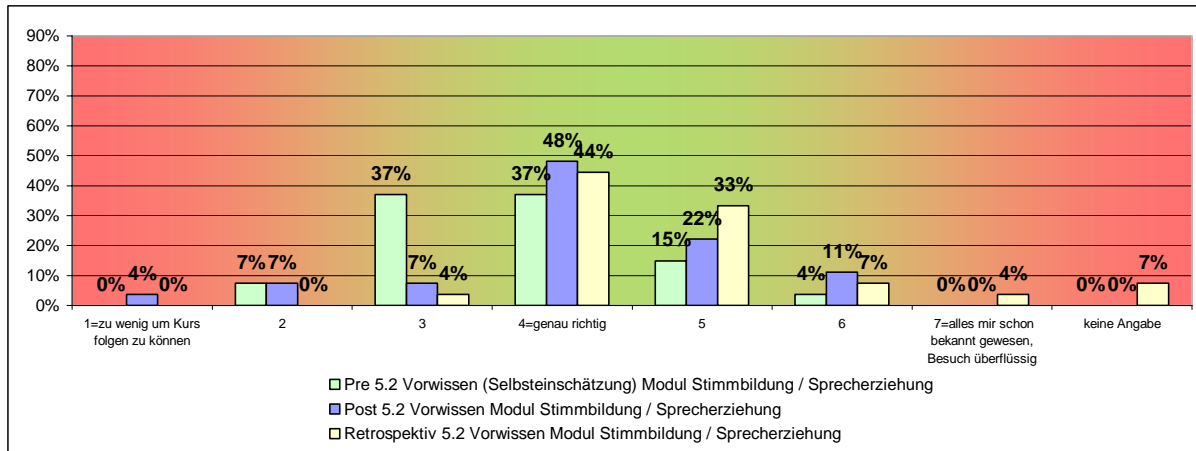


Diagramm 9: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Stimmbildung / Sprecherziehung

Analyse / Ergebnisinterpretation Vorwissen Modul Stimmbildung und Sprecherziehung (siehe Diagramm 9)

Das Vorwissen im Modul Stimmbildung und Sprecherziehung ist – betrachtet man Median und Quartilabweichung – über alle Erhebungszeitpunkte tendenziell gleich (jeweils Median 4 und Quartilabweichung 1). Anhand der Datenbasis kann allerdings analog zum Modul Rhetorik für dieses Modul eine Verlagerung der Selbsteinschätzung (bezogen auf Pre- und Retrospektionserhebung) beobachtet werden. Die Verteilung ist in der Vorab-Untersuchung leicht links- in der Retrospektionsuntersuchung leicht rechtslastig.

Vorwissen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Vorwissen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

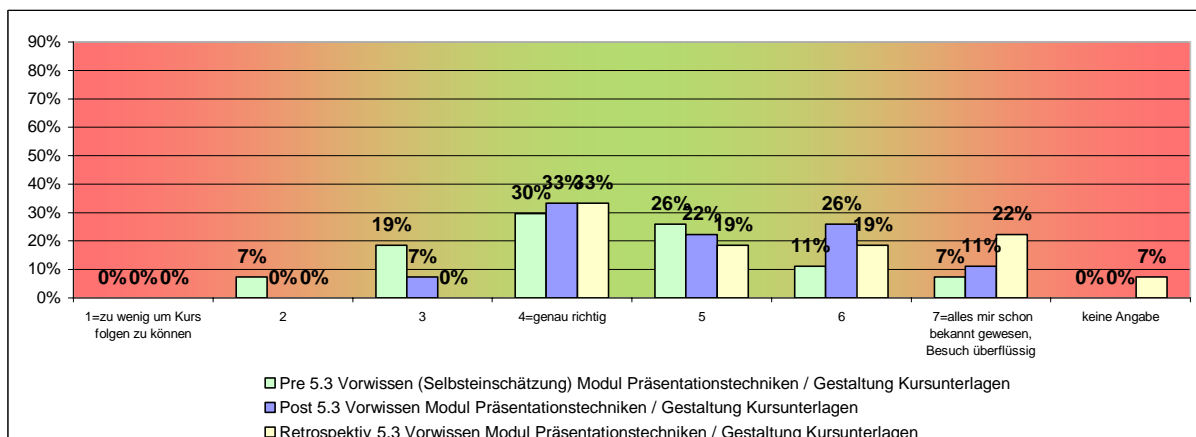


Diagramm 10: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Einsatz von Präsentationstechnik / Kursunterlagen

Analyse / Ergebnisinterpretation Vorwissen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 10)

Die Selbsteinschätzung der Teilnehmer, bezogen auf das Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen, unterscheidet sich geringfügig von der Einschätzung der beiden vorhergehenden Module. In der Vorher- Befragung liegt der Median bei einer Quartilabweichung von 1,5 bei 4, während in den beiden Einschätzungen nach der Veranstaltung der Median bei 5 mit einer Quartilabweichung von 2 liegen. In der retrospektiven Betrachtung bewerten die Teilnehmer ihr Vorwissen tendenziell als stärker ausgeprägt.

Vorwissen Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Vorwissen kohärenter Sprachgebrauch

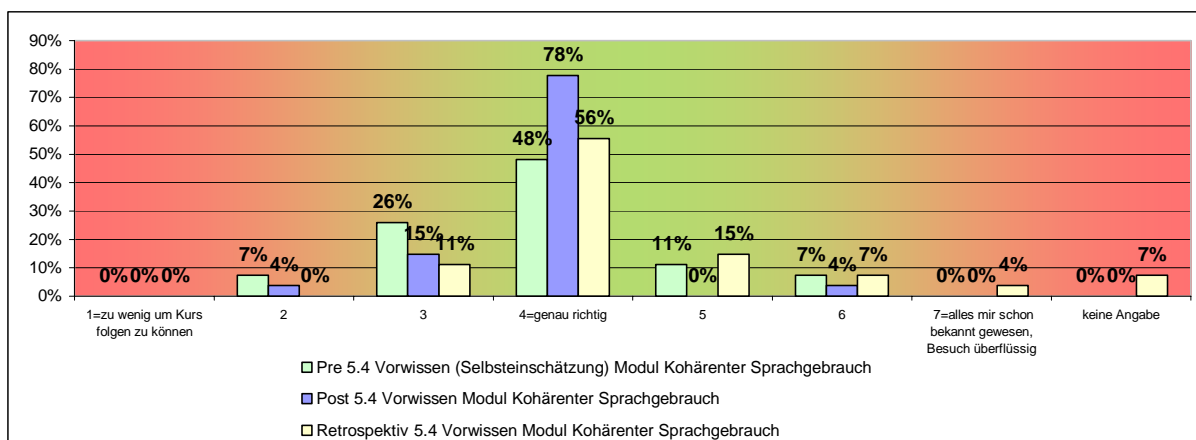


Diagramm 11: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Vorwissen kohärenter Sprachgebrauch (siehe Diagramm 11)

Die Teilnehmer schätzen ihr Vorwissen im Modul kohärenter Sprachgebrauch ziemlich eindeutig ein. Tendenziell besteht derselbe Unterschied zwischen Vorabereinschätzung und retrospektiver Betrachtung wie in den Modulen Rhetorik und Stimmbildung / Sprecherziehung. Diese Tendenz ist allerdings nur sehr schwach ausgeprägt. Interessant ist die Beurteilung des Vorwissens direkt nach dem Seminar. 78% der Teilnehmer bewerten in diesem Rahmen ihr Vorwissen als „genau richtig“, um den Kursinhalten folgen zu können (Median 4 – Quartilabweichung 0). Dies lässt auf eine genaue Passung der Teilnehmerbedürfnisse mit den Modulinhalt schließen.

Vorwissen Modul Lehrkompetenz

Untersuchungsergebnisse Vorwissen Lehrkompetenz

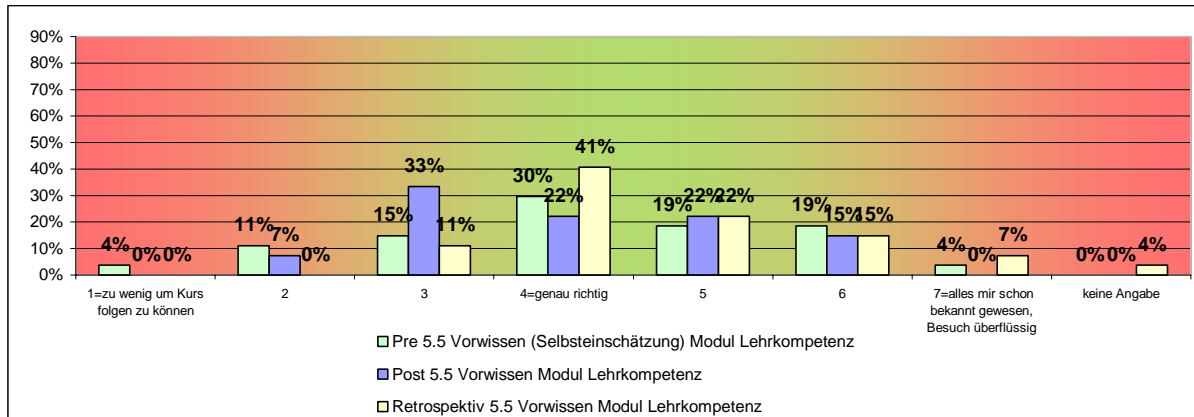


Diagramm 12: Untersuchungsergebnisse Evaluation PROFiL-TT Vorwissen Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Vorwissen Lehrkompetenz (siehe Diagramm 12)

Die Einschätzungen des Vorwissens im Modul Lehrkompetenz sind durchwegs heterogen. Der Median beträgt zwar in allen drei Erhebungsstufen 4, allerdings ist die Quartilabweichung vorher und nachher 2, in der Retrospektion 1. Ein Drittel der Teilnehmer schätzen ihr Vorwissen direkt nach der Veranstaltung als tendenziell etwas zu niedrig ein, in der retrospektiven Betrachtung liegt der Schwerpunkt (41%) allerdings auf dem Maximalwert. Dennoch kann aus dieser Verteilung keine eindeutige Tendenz gelesen werden.

Inhaltliche Überschneidungen zwischen den einzelnen Modulen

Untersuchungsergebnisse inhaltliche Überschneidungen

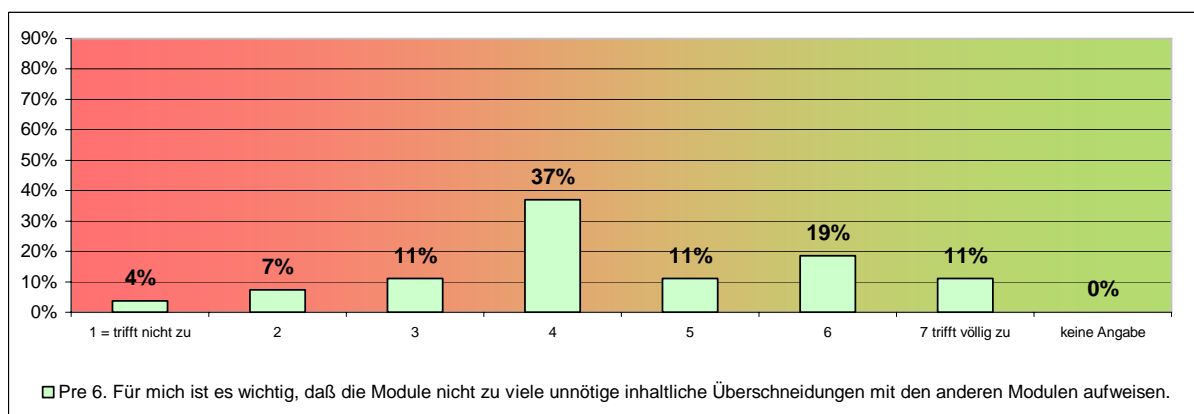


Diagramm 13a: Untersuchungsergebnisse Evaluation PROFiL-TT inhaltliche Überschneidungen vorher

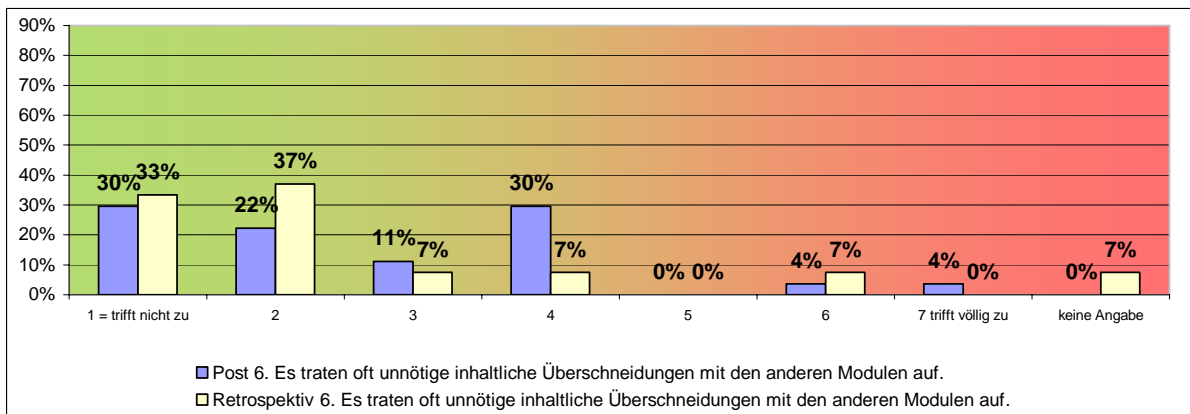


Diagramm 13b: Untersuchungsergebnisse Evaluation PROFiL-TT inhaltliche Überschneidungen nachher / retrospektiv

Analyse / Ergebnisinterpretation inhaltliche Überschneidungen (siehe Diagramm 13a/b)

Die Frage nach den inhaltlichen Redundanzen ist (bezogen auf die Skalierung in Vorheruntersuchung und Nachher- und Retrospektionsuntersuchung) aufgeteilt. Die Teilnehmereinschätzung weist vorab in eine neutrale Richtung. Inhaltliche Überschneidungen zwischen den Modulen zeigen eine mittlere Akzeptanz. Der Median ist demzufolge in der Vorabbefragung 4 (die Quartilabweichung 2). In der Erhebungen direkt nach der Veranstaltung liegt der Schwerpunkt der Einschätzungen – ob unnötige inhaltliche Überschneidungen zwischen den einzelnen Modulen stattgefunden haben – tendenziell auf „trifft nicht zu“, allerdings zeigt die Quartilabweichung für diese Erhebung (Wert 3) eine deutliche Verteilung der Daten auf – hier kann wiederum eine Analogie zur Vorheruntersuchung gezogen werden. In der Retrospektion ist diese Abweichung wesentlich geringer ausgeprägt (Wert 1). Das Gros der Teilnehmer ist in dieser Stufe der Untersuchung der Ansicht, dass unnötige Redundanzen im gesamten Kursverlauf selten auftraten.

Die Ergebnisse zu dieser Frage weisen im zeitlichen Verlauf der Untersuchung eine merkliche Diskrepanz auf. Unnötige inhaltliche – intermodulare – Überschneidungen werden vorab neutral (Schwerpunkt auf „teils/teils“) beurteilt, in der Nachheruntersuchung tendieren die Ergebnisse dazu, dass diese Redundanzen zwar gelegentlich aufgetreten sind, wobei der Median allerdings darauf hinweist, dass unnötige inhaltliche Überschneidungen (mit einer relativ hohen Streuung) kaum auftraten. In der letzten Stufe der Erhebung werden diese Überschneidungen ebenso als kaum vorhanden eingeschätzt; die Streuung ist dabei allerdings (mit dem Wert 1) wesentlich geringer als in der Post- Untersuchung.

3.2.4.6 Anforderungen

Überforderung oder Anforderungen beziehen sich auf die vier Schwerpunkte Stoffschwere, Stoffumfang, Kurs-tempo und Schwierigkeitsstufe der Anforderungen allgemein. Die Teilnehmer haben mit diesen Items die Möglichkeit zu bewerten, ob der Lerninhalt für sie auf dem richtigen Niveau und ob die Geschwindigkeit, in der dieser Inhalt dargeboten wurde, für sie angemessen war oder ob die Stoffmenge noch von ihnen verarbeitet werden konnte. Der Optimalwert ist bei diesen Items zentral.

3.2.4.6.1 Stoffschwere

Stoffschwere Modul Rhetorik

Untersuchungsergebnisse Stoffschwere Modul Rhetorik

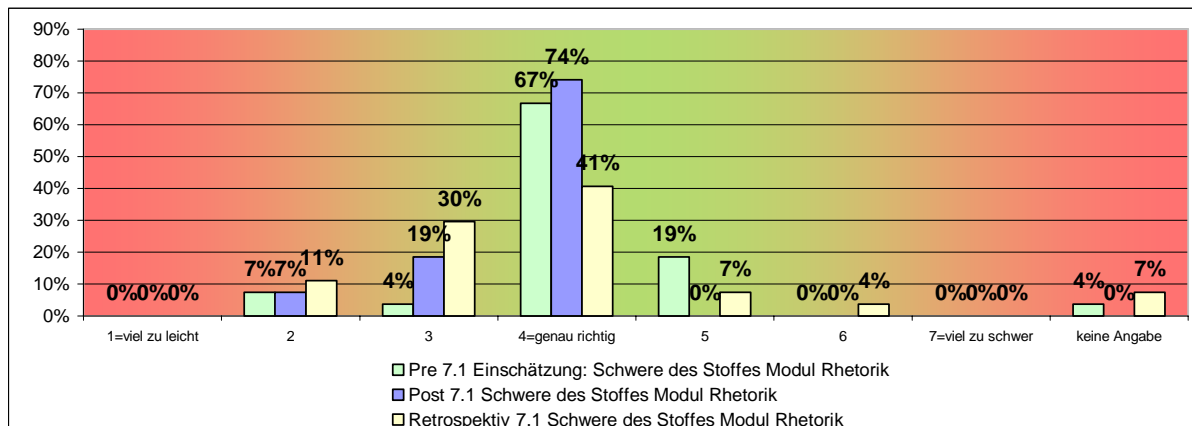


Diagramm 14: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Rhetorik

Analyse / Ergebnisinterpretation Stoffschwere Modul Rhetorik (siehe Diagramm 14)

Die Bewertung und Einschätzung der Stoffschwere im Modul Rhetorik, ist in allen drei Stufen der Evaluation sehr ähnlich. Vorab, nach dem Seminar und retrospektiv beträgt der Median 4. Das Niveau des Moduls Rhetorik wird somit wesentlich den Teilnehmeranforderungen entsprechend bewertet. Die Streuung in der retrospektiven Bewertung (Quartilabweichung 1) zeigt die geringfügige Tendenz dahingehend auf, dass den Teilnehmern die Stoffschwere im Modul Rhetorik etwas zu leicht erschien.

Stoffschwere Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Stoffschwere Modul Stimmbildung und Sprecherziehung

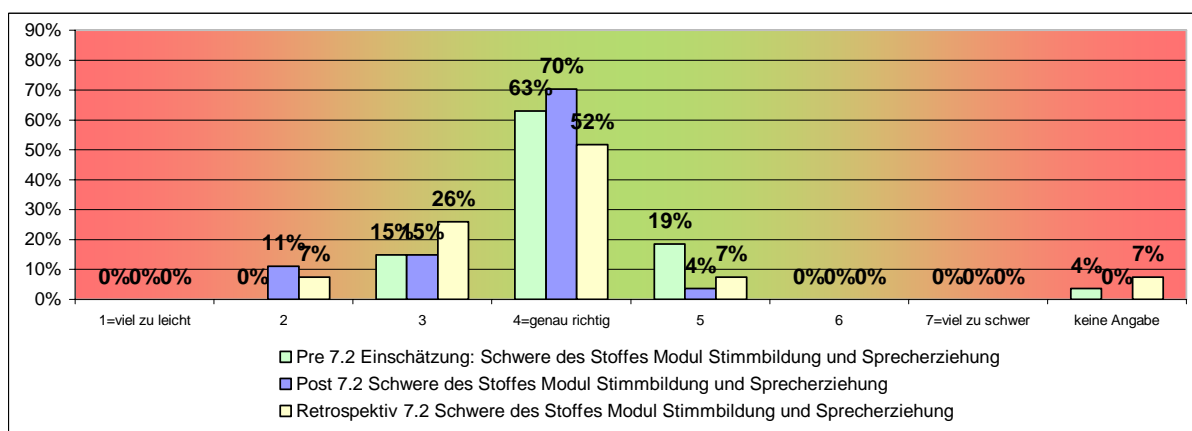


Diagramm 15: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Stoffschwere Modul Stimmbildung und Sprecherziehung (siehe Diagramm 15)

Die Tendenz in der Verteilung zur Stoffschwere im Modul Stimmbildung und Sprecherziehung ähnelt der Verteilung im Modul Rhetorik stark (Median und Quartilabweichung sind identisch $4 - 4 - 4 / 0 - 0,5 - 1$). Auch für dieses Modul konzentrieren sich die Werte um den Maximalwert, mit einer sehr geringen Streuung, die allerdings im Verlauf der Erhebung leicht zunimmt.

Stoffschwere Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Stoffschwere Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

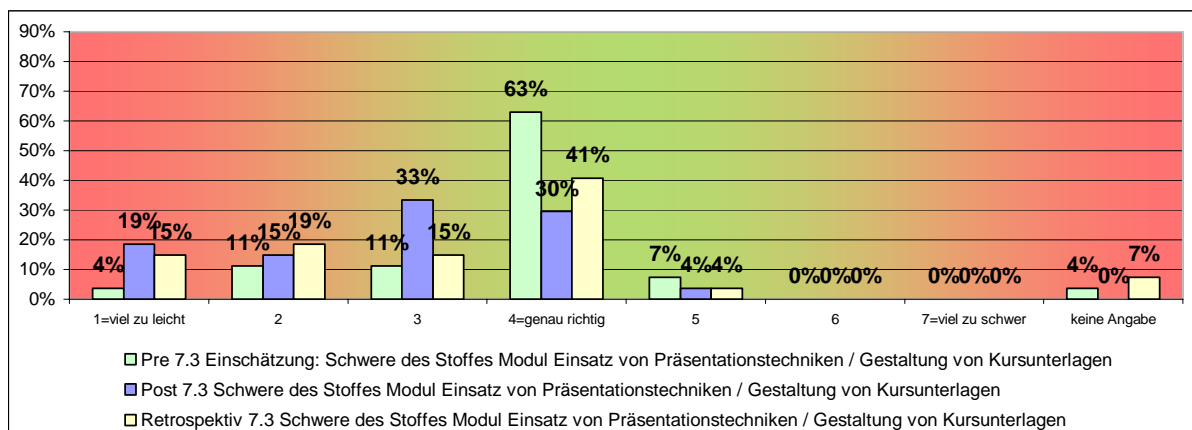


Diagramm 16: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Stoffschwere Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 16)

Die Einschätzung und Beurteilung der Stoffschwere im Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen tendiert – abgesehen von der Voraberechnung – in der Skala nach links (eher zu leicht für die Anforderungen der Teilnehmer). Der Median in der Voruntersuchung liegt bei 4 mit einer relativ geringen Quartilabweichung (0,8). Die Mediane in der Nachher- und Retrospektionsbefragung hingegen liegen bei 3 (einergehend mit einer Quartilabweichung von jeweils 2). Die Ausprägung der Verteilungen – vor allem der Beurteilungen nach der Veranstaltung – ist tendenziell linkslastig und neigt dazu, dass die Stoffschwere von den Teilnehmern als zu leicht empfunden wurde.

Stoffschwere Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Stoffschwere Modul kohärenter Sprachgebrauch

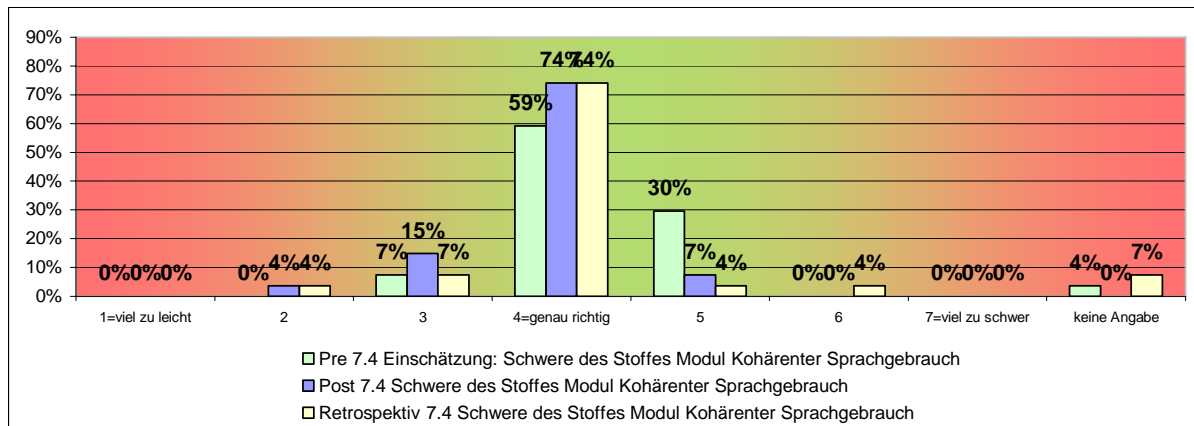


Diagramm 17: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Stoffschwere Modul kohärenter Sprachgebrauch (siehe Diagramm 17)

Die Stoffschwere im Modul kohärenter Sprachgebrauch wurde über alle drei Evaluationsstufen sehr ähnlich eingeschätzt. Der Median in allen Ebenen beträgt 4, die Quartilabweichung in der Nachher- und Retrospektions-Untersuchung 0. In der Vorabbefragung kann man eine leichte Tendenz der Teilnehmer erkennen, die die Stoffschwere in diesem Modul als „eher hoch“ einschätzen (diese Tendenz spiegelt sich in einer Quartilabweichung 1 in der Vorabbefragung wieder).

Stoffschwere Modul Lehrkompetenz

Untersuchungsergebnisse Stoffschwere Modul Lehrkompetenz

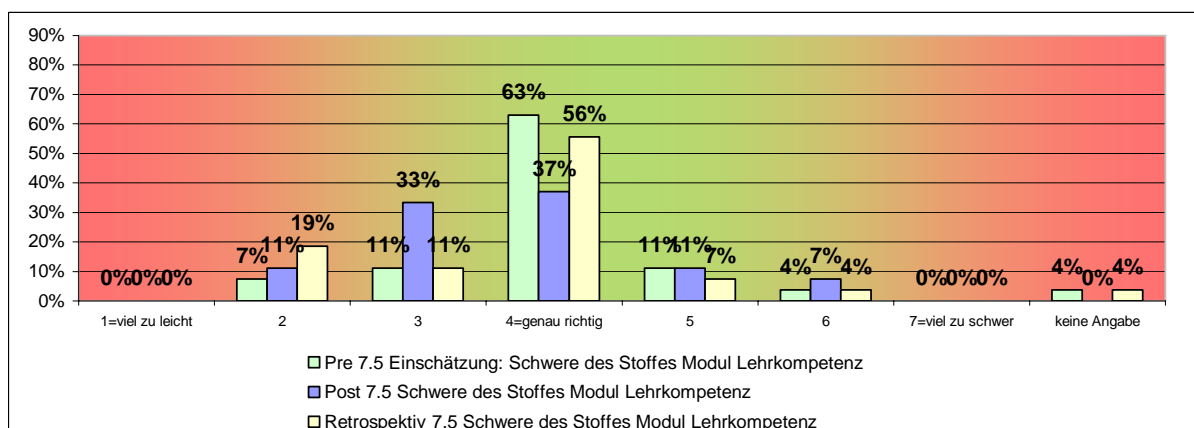


Diagramm 18: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffschwere Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Stoffschwere Modul Lehrkompetenz (siehe Diagramm 18)

Im Modul Lehrkompetenz wird die Stoffschwere von den meisten Teilnehmern vorab als „genau richtig“ eingeschätzt (Median 4 – Quartilabweichung 0). Diese Tendenz kann man auch in der Retrospektionsuntersuchung beobachten (Median 4 – Quartilabweichung 1). Die Werteverteilung direkt nach dem Seminar ist zwar mit den retrospektiven Daten identisch, betrachtet man allerdings die Datenbasis, kann man bei der Nachherbefragung einen Trend in Richtung „zu leicht“ feststellen (Vergleich post Wert 3: 33% - retrospektiver Wert 3: 11%; post Wert 4: 37% - retrospektiver Wert 3: 56%). Dieser Punkt spiegelt sich auch in der qualitativen Datenauswertung wider (siehe Kapitel 3.2.5.3).

3.2.4.6.2 Stoffumfang

Stoffumfang Modul Rhetorik

Untersuchungsergebnisse Stoffumfang Modul Rhetorik

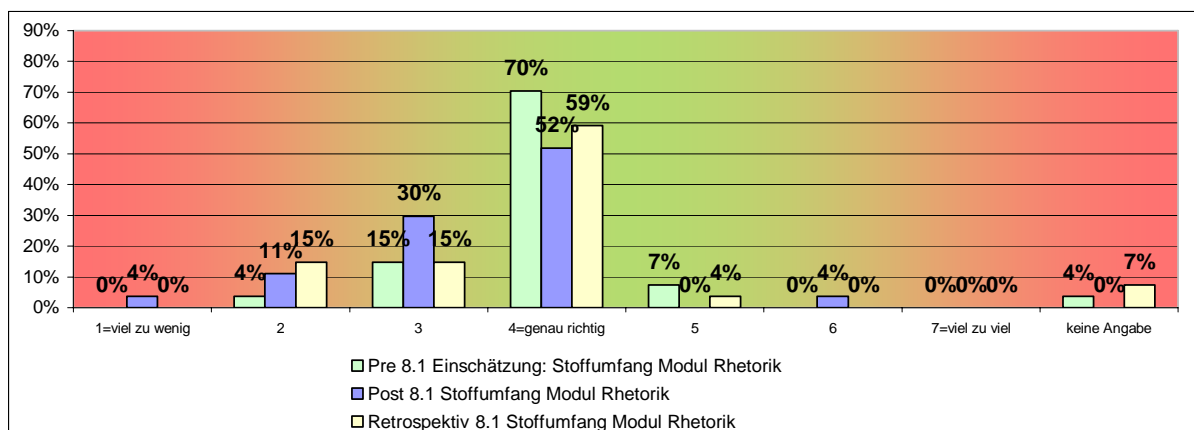


Diagramm 19: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Rhetorik

Analyse / Ergebnisinterpretation Stoffumfang Modul Rhetorik (siehe Diagramm 19)

Der Stoffumfang im Modul Rhetorik wird überwiegend mit 4 (der Stoffumfang ist „genau richtig“) beurteilt. In allen drei Untersuchungsstadien ergibt die Datenauswertung den Median 4. Betrachtet man die Werte der Nachher- und Retrospektionsuntersuchung, kann man feststellen, dass der Stoffumfang in diesem Modul – verglichen mit den Werten der Vorherbefragung – tendenziell zu wenig ist. Die Quartilabweichung in diesen beiden Untersuchungen ist jeweils 1.

Stoffumfang Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Stoffumfang Modul Stimmbildung und Sprecherziehung

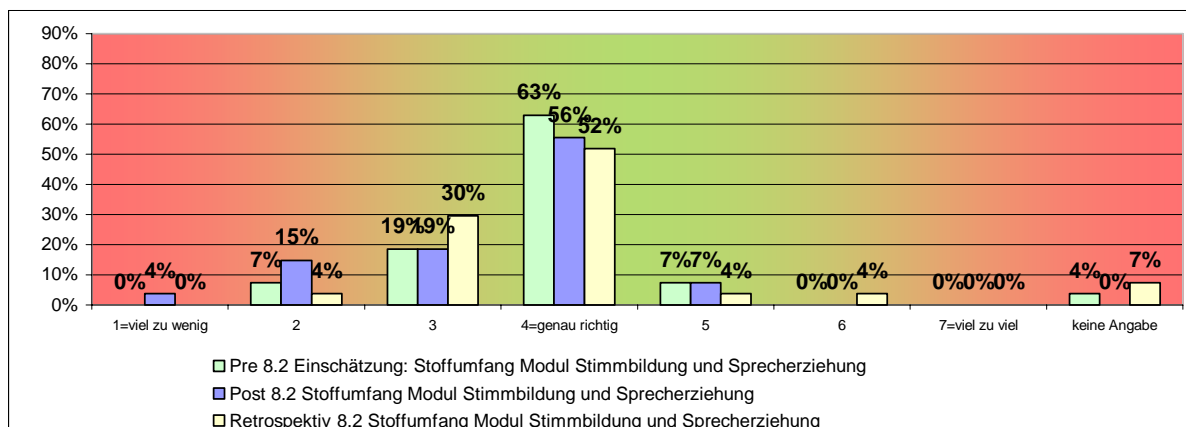


Diagramm 20: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Stoffumfang Modul Stimmbildung und Sprecherziehung (siehe Diagramm 20)

Median und Quartilabweichung bewegen sich in allen drei Stufen der Erhebung in etwa gleich (Median Wert 4 – Quartilabweichung vorher 0,8, ansonsten 1). Der Stoffumfang im Modul Stimmbildung und Sprecherziehung wurde – wie im Modul Rhetorik auch – ansatzweise als etwas zu wenig beurteilt. Auch für dieses Modul ist die Verteilung der Werte leicht linkslastig, wobei der Schwerpunkt der Werte sich auch hier um die Einschätzung „genau richtig“ bewegt.

Stoffumfang Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Stoffumfang Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

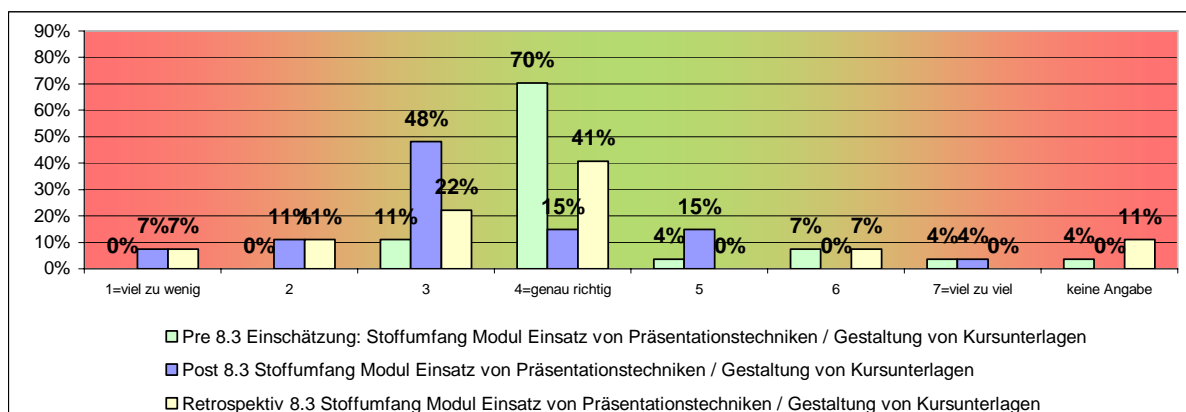


Diagramm 21: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Einsatz von Präsentationstechnik / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Stoffumfang Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 21)

Die Vorab einschätzung für den Stoffumfang im Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen hat ihren Schwerpunkt im Bereich „genau richtig“ (70%, Median 4, Quartilabweichung 0). In der Erhebung direkt nach der Veranstaltung verschiebt sich diese Verteilung etwas nach links („Stoffumfang war ansatzweise zu wenig“; Median 3, Quartilabweichung 1). Eine Verlagerung zwischen Einschätzung vor dem Seminar und nach dem Seminar kann hier deutlich abgelesen werden. In der retrospektiven Erhebung (Median 4, Quartilabweichung 1) verlagert sich diese Verschiebung wieder in den zentralen Bereich der Skala, wobei eine leichte Tendenz („Stoffumfang zu wenig“) immer noch erkennbar bleibt.

Stoffumfang Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Stoffumfang Modul kohärenter Sprachgebrauch

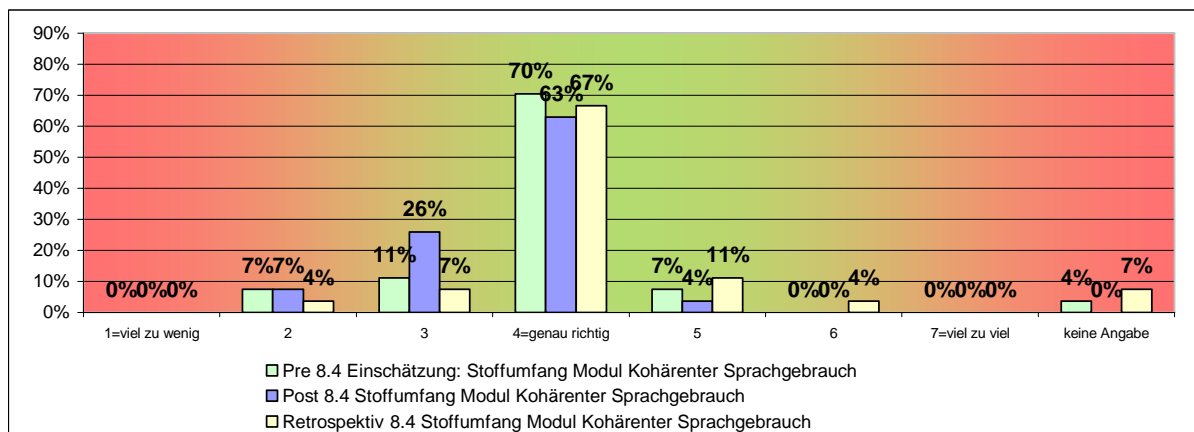


Diagramm 22: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Stoffumfang Modul kohärenter Sprachgebrauch (siehe Diagramm 22)

Die Wertverteilung bei der Frage nach dem Stoffumfang im Modul kohärenter Sprachgebrauch ist in allen drei Erhebungsstufen sehr eindeutig. Der Median vorher, nachher und retrospektiv beträgt 4, wobei lediglich bei der Post- Erhebung eine Quartilabweichung mit dem Wert 1 gemessen werden kann. Die Einschätzung, dass der Stoffumfang in diesem Modul „genau richtig“ ist / war teilen zwischen 60% und 70% der Teilnehmer. Der Quartilabstand in der Post- Erhebung resultiert aus einer Tendenz in Richtung „etwas zu wenig“ Stoffumfang, bei der retrospektiven Betrachtung egalisiert sich diese Tendenz in beide Richtungen.

Stoffumfang Modul Lehrkompetenz

Untersuchungsergebnisse Stoffumfang Modul Lehrkompetenz

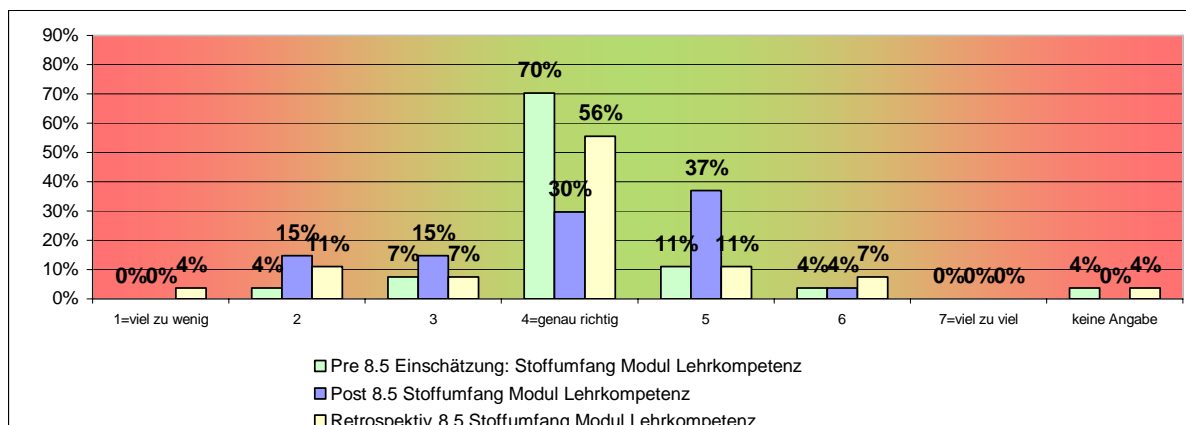


Diagramm 23: Untersuchungsergebnisse Evaluation PROFiL-TT Stoffumfang Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Stoffumfang Modul Lehrkompetenz (siehe Diagramm 23)

Bei der Einschätzung des Stoffumfangs im Modul Lehrkompetenz ist der Wert der Vorab- und der Retrospektions-Erhebung relativ eindeutig. Beide bewegen sich um den Maximalwert (Median 4) mit einer Quartilabweichung von 0. Bei der Post-Erhebung ist allerdings die Tendenz erkennbar, dass der Stoffumfang innerhalb dieses Moduls zu umfassend war. Der Median liegt zwar auch bei 4, allerdings beträgt die Quartilabweichung 2, was daran liegt, dass 37% der Untersuchungsteilnehmer den Stoffumfang dieses Moduls mit 5 („etwas zu viel“) bewertet haben. Andererseits haben 30% der Teilnehmer den Stoffumfang innerhalb dieses Moduls als „zu wenig“ (schwach bis mittelmäßig ausgeprägt) beurteilt. Hintergründe könnten hierfür die stark wissenschaftliche Orientierung dieses Moduls sein (siehe Kapitel 3.2.5.3).

3.2.4.6.3 Kurstempo

Kurstempo Modul Rhetorik

Untersuchungsergebnisse Kurstempo Modul Rhetorik

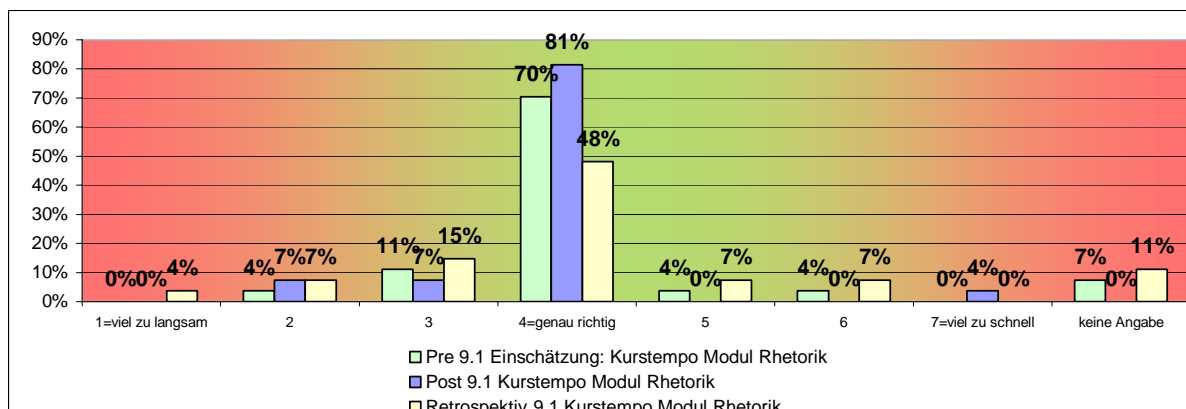


Diagramm 24: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Rhetorik

Analyse / Ergebnisinterpretation Kurstempo Modul Rhetorik (siehe Diagramm 24)

Die Einschätzung des Kurstempos für das Modul Rhetorik bewegt sich in der Vorab- und Nachher- Befragung vorwiegend um den zentralen Wert (Median 4 mit einer Quartilabweichung 0). Die äußerst hohe Zustimmung (81%) in der Post- Befragung ragt hierbei heraus und bleibt – allerdings in abgeschwächter Form – auch in der retrospektiven Betrachtung erhalten (Median 4 – Quartilabweichung 1). Tendenziell zeigt sich in der retrospektiven Betrachtung eine leichte Verlagerung der Verteilung in Richtung „Kurstempo etwas zu langsam“. Diese Tendenz ist allerdings nur sehr schwach ausgeprägt.

Kurstempo Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Kurstempo Modul Stimmbildung und Sprecherziehung

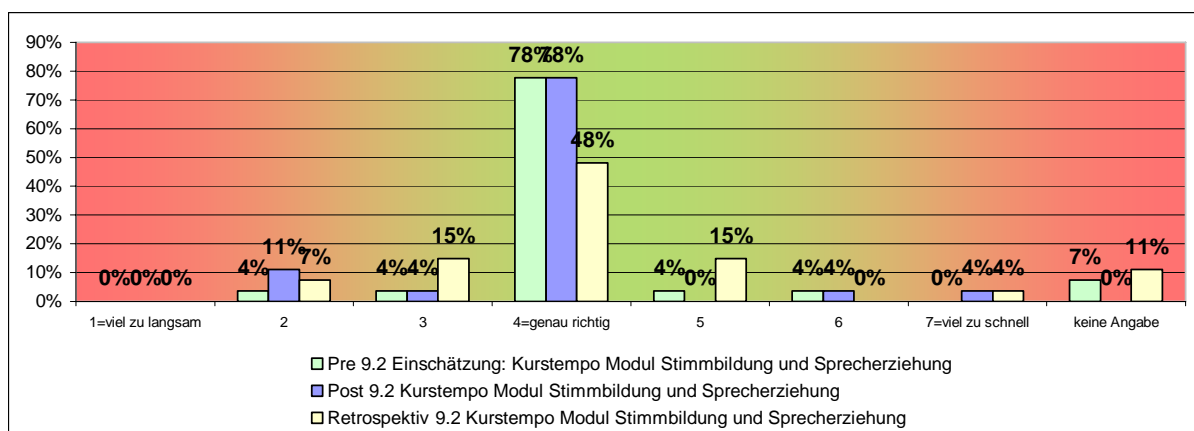


Diagramm 25: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Kurstempo Modul Stimmbildung und Sprecherziehung (siehe Diagramm 25)

Das Kurstempo im Modul Stimmbildung und Sprecherziehung wird ähnlich beurteilt, wie im Modul Rhetorik. Der Median liegt in allen drei Erhebungsstufen bei 4, die Quartilabweichung dabei ist sehr gering (pre und post 0, retrospektiv 0,3). Auch in dieser Verteilung ist der Unterschied zwischen der Post- und der retrospektiven Erhebung interessant. Während in der Befragung direkt nach der Veranstaltung noch 78% der Teilnehmer das Kurstempo als „genau richtig“ einschätzen, erhält man bei der Retrospektionsbetrachtung einen Wert von 48%. Die Verteilung um den Zentralwert ist für dieses Modul allerdings relativ ausgewogen – 22% der Teilnehmer tendieren dazu, das Kurstempo als zu schnell einzuschätzen, 19% als eher zu langsam.

Kurstempo Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Kurstempo Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

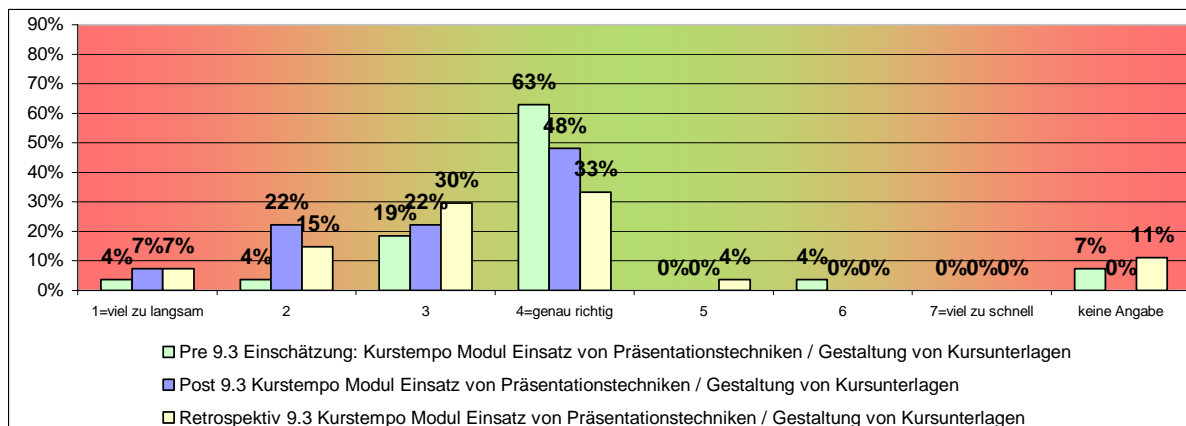


Diagramm 26: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Einsatz von Präsentationstechnik / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Kurstempo Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 26)

Auf den ersten Blick hat die Verteilung für das Kurstempo im Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen ihren Schwerpunkt in der Mitte. Eine Tendenz in die Richtung „Kurstempo ist zu langsam“ kann allerdings eindeutig festgestellt werden. Diese Einschätzung wird von Median und Quartilabweichung (vor allem in der Post- und Retrospektionsbefragung) gestärkt. In der Vorabbefragung liegt der Median bei 4 (Quartilabweichung 1), auch hier zeigt die Einschätzung der Teilnehmer bereits in Richtung „etwas zu langsam“. Der Median in der Post- und Retrospektions- Erhebung zeigt mit dem Wert von jeweils 3 (allerdings mit der Quartilabweichung von 2 (post) und 1,3 (retrospektiv)) wiederum in die Richtung „Kurstempo etwas zu langsam“.

Kurstempo Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Kurstempo Modul kohärenter Sprachgebrauch

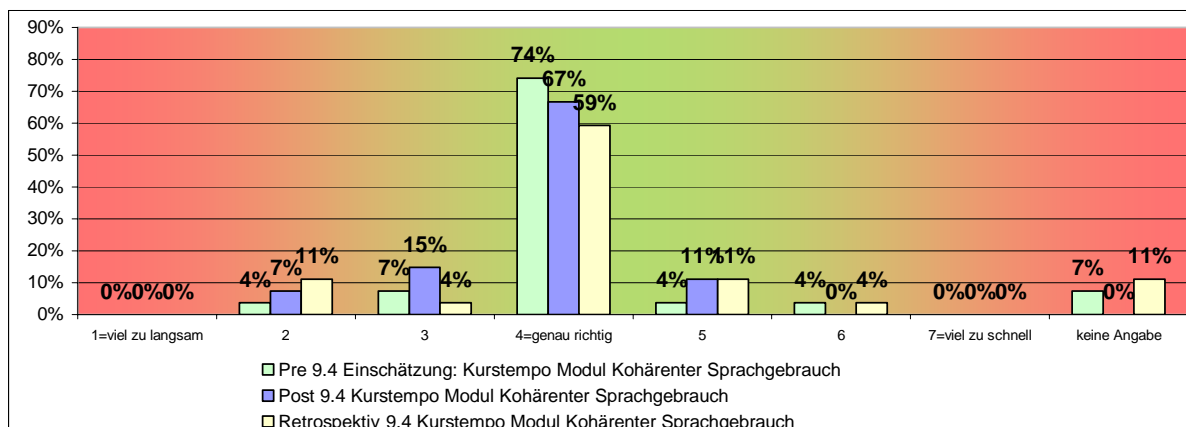


Diagramm 27: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Kurstempo Modul kohärenter Sprachgebrauch (siehe Diagramm 27)

Die Verteilung der Werte – Kurstempo im Modul kohärenter Sprachgebrauch – ist eindeutig zentral gelagert (Median jeweils Wert 4, Quartilabweichung jeweils 0). Die Tendenz „Kurstempo genau richtig“ hält sich über alle drei Einschätzungsstufen, wobei sie im Verlauf von prospektiver Erhebung bis zur retrospektiven Betrachtung tendenziell etwas abgeschwächt wird. Die Relationen der einzelnen Verteilungen bleiben allerdings – genau wie die Quantilswerte – konstant.

Kurstempo Modul Lehrkompetenz

Untersuchungsergebnisse Kurstempo Modul Lehrkompetenz

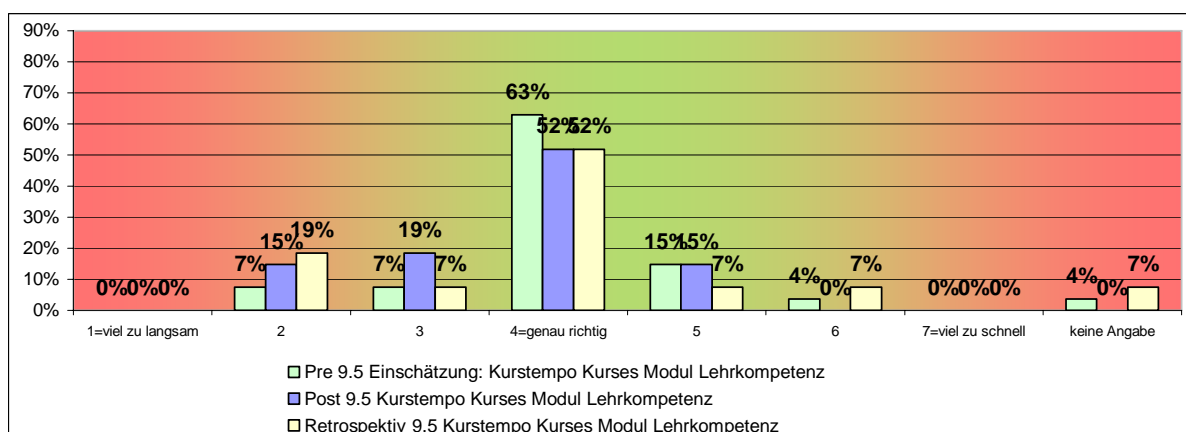


Diagramm 28: Untersuchungsergebnisse Evaluation PROFiL-TT Kurstempo Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Kurstempo Modul Lehrkompetenz (siehe Diagramm 28)

Die Einschätzungen der Teilnehmer bezogen auf das Kurstempo im Modul Lehrkompetenz ist konstant. Die Streuung im Rahmen der Vorabbefragung beträgt 0 und zentriert sich auf den Wert „Kurstempo genau richtig“ (Median 4). Diese Tendenz setzt sich auch in der Nachher- und Retrospektions- Befragung fort, wobei diese Erhebungen eine etwas größere Streuung (Quartilabweichung 1) aufweisen.

3.2.4.6.4 Anforderungen

Anforderungen Modul Rhetorik

Untersuchungsergebnisse Anforderungen Modul Rhetorik

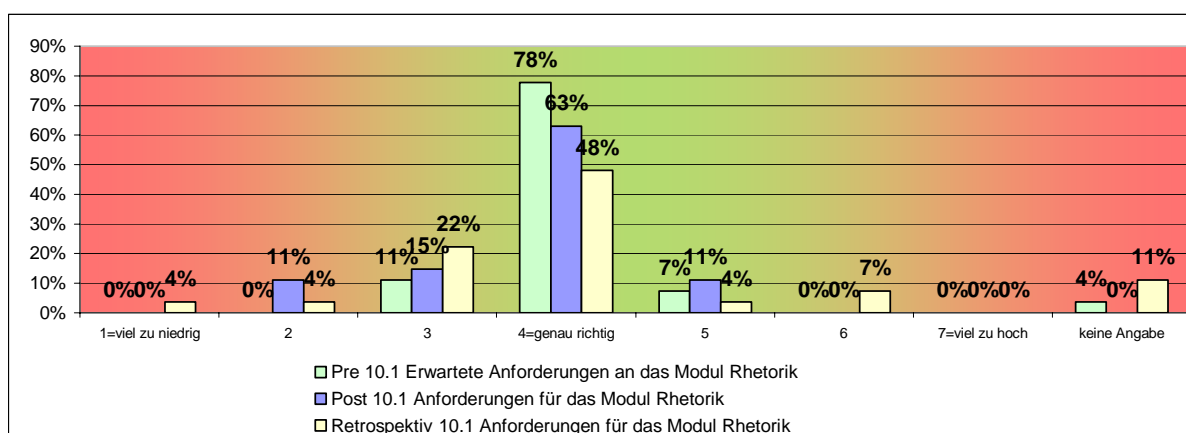


Diagramm 29: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul Rhetorik

Analyse / Ergebnisinterpretation Anforderungen Modul Rhetorik (siehe Diagramm 29)

Die Einschätzung der Anforderungen im Modul Rhetorik nimmt im Verlauf Vorher-, Nachher- und retrospektive Betrachtung etwas ab. Der Median in allen drei Untersuchungsstufen hat den Wert 4, wobei die Quartilabweichung geringfügig zunimmt (0 – 0,5 – 1). Die Werte der Nachherbefragung tendieren in die Richtung „Anforderung etwas zu niedrig“. Diese Tendenz kann auch in den Werten der retrospektiven Betrachtung wieder gefunden werden.

Anforderungen Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Anforderungen Modul Stimmbildung und Sprecherziehung

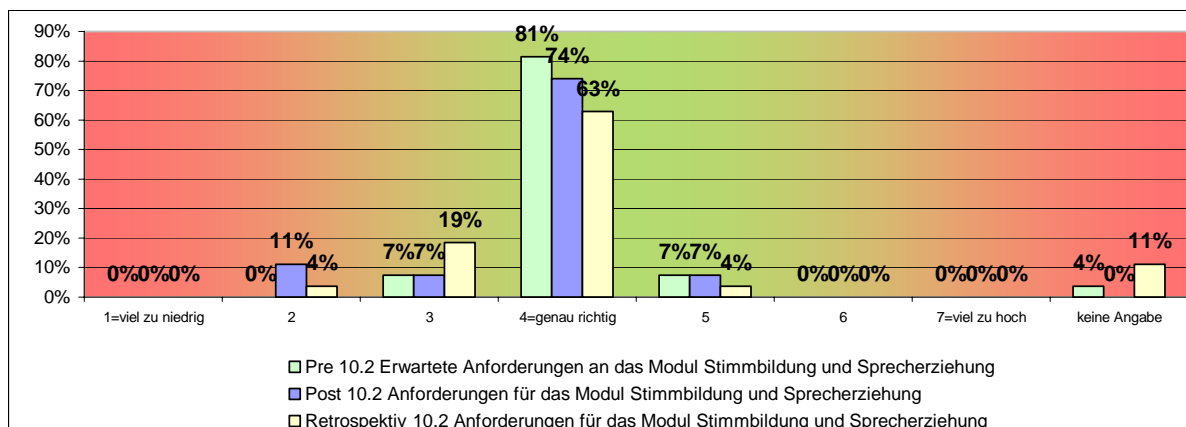


Diagramm 30: Untersuchungsergebnisse Evaluation PROFIL-TT Anforderungen Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Anforderungen Modul Stimmbildung und Sprecherziehung (siehe Diagramm 30)

Die Anforderungen an die Teilnehmer im Modul Stimmbildung und Sprecherziehungen werden von vornherein als „genau richtig“ eingeschätzt. Lediglich ein geringer Prozentsatz in der Post- und Retrospektionserhebung ist der Ansicht, dass die Anforderungen etwas zu niedrig waren (Anforderungen „zu niedrig“ / „etwas zu niedrig“ in der Post- Erhebung 18%, in der Retrospektionserhebung 23%). Dementsprechend hat der Median über alle drei Untersuchungsstufen den Wert 4 bei einer Quartilabweichung von 0, bzw. 0,3 in der Retrospektion.

Anforderungen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Anforderungen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

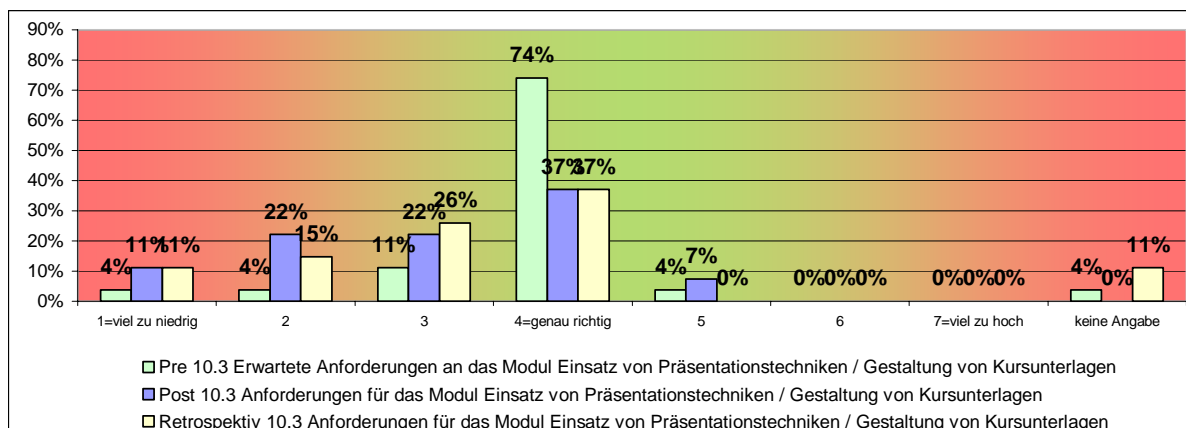


Diagramm 31: Untersuchungsergebnisse Evaluation PROFIL-TT Anforderungen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Anforderungen Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 31)

Für das Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen werden die Anforderungen von den Teilnehmern heterogen beurteilt. In der Vorab-Untersuchung schätzen fast drei Viertel der Teilnehmer die Anforderungen in diesem Modul als weder zu hoch noch zu niedrig ein (Median 4, Quartilabweichung 0). In den Werten der Posterhebung verlagert sich diese Einschätzung allerdings insofern, dass 55% der Teilnehmer die Anforderungen als „etwas zu niedrig“ bis hin zu „viel zu niedrig“ beurteilen (Median 3 – Quartilabweichung 2). Diese Tendenz setzt sich auch in den Werten der retrospektiven Erhebung fort.

Anforderungen Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Anforderungen Modul kohärenter Sprachgebrauch

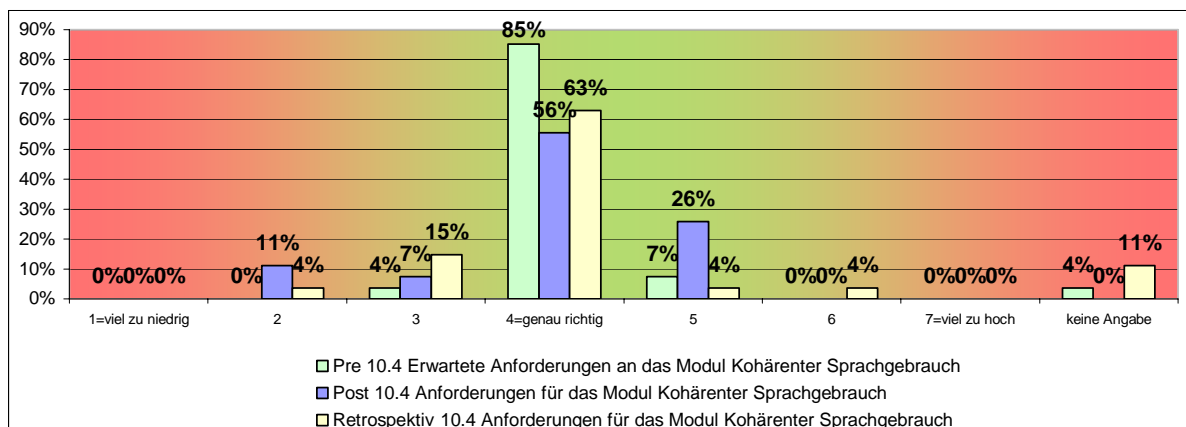


Diagramm 32: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Anforderungen Modul kohärenter Sprachgebrauch (siehe Diagramm 32)

Die Einschätzungen der Teilnehmer in der Vorher- Erhebung bezogen auf die Anforderungen im Modul kohärenter Sprachgebrauch liegen mit 85% genau im zentralen Wert der Skala („genau richtig“).

Betrachtet man lediglich die Mittel- und Streuungswerte der drei Erhebungsstufen, kommt man zu dem Ergebnis, dass die Werte aller drei Untersuchungsabschnitte sich im Rahmen des zentralen Skalenwerts befinden (Median jeweils 4). Die Quartilabweichung der drei Erhebungsabschnitte bekräftigen diesen Eindruck (0 – 0,5 – 0). In der Post- Erhebung lässt sich eine geringe Tendenz in Richtung *Anforderungen etwas zu hoch* feststellen.

Anforderungen Modul Lehrkompetenz

Untersuchungsergebnisse Anforderungen Modul Lehrkompetenz

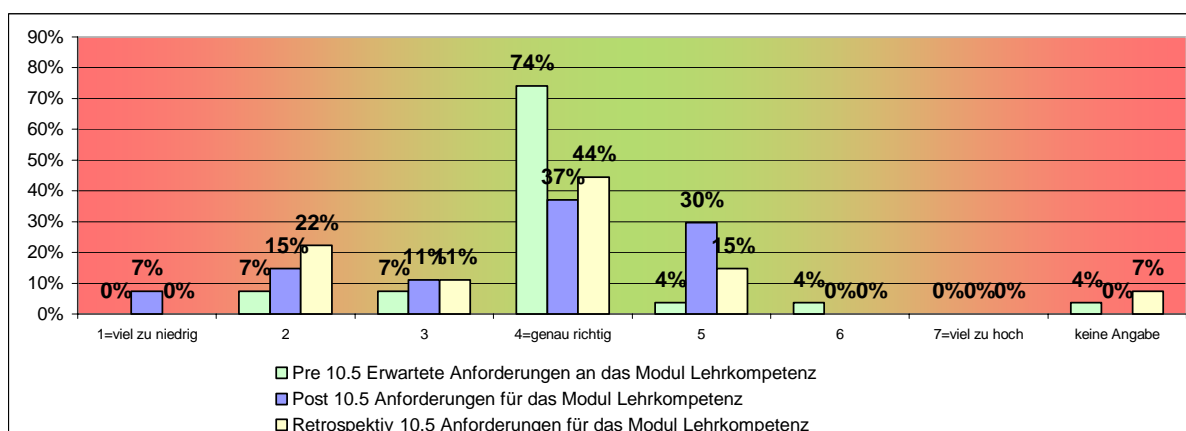


Diagramm 33: Untersuchungsergebnisse Evaluation PROFiL-TT Anforderungen Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Anforderungen Modul Lehrkompetenz (siehe Diagramm 33)

Das Modul Lehrkompetenz weist – bezogen auf die Anforderungen – in allen drei Erhebungssequenzen einen Median von 4 auf. Während die Quartilabweichung 0 vorher die Tendenz der vorherigen Module fortsetzt, kann man in der Quartilabweichung von 2 in der Erhebung direkt nach der Veranstaltung eine Zweiteilung der Einschätzungen erkennen. Einerseits beurteilen ein Drittel der Teilnehmer die empfundenen Anforderungen des Moduls Lehrkompetenz als zu niedrig (in unterschiedlich starker Ausprägung), andererseits tendieren ebenso fast ein Drittel der Teilnehmer dazu, die Anforderungen in der Post- Erhebung als etwas zu hoch einzustufen. In der retrospektiven Betrachtung bleibt allerdings nur eine Seite dieser Einschätzung konstant – 33% der Teilnehmer beurteilen in diesem Rahmen die Anforderungen als etwas zu niedrig. Diese Aussage wird in der qualitativen Auswertung bekräftigt (siehe Kapitel 3.2.5.3).

3.2.4.7 Lernen – quantitativ

Grundlegend für die Beurteilung der Quantität des Lernens ist eine Selbsteinschätzung der Teilnehmer, bezogen auf den Lernumfang, der in der Veranstaltung vermittelt wird. Die Bewertung bezieht sich auf den Vergleich des Wissenstandes *vor der Maßnahme* mit dem Wissenstand *nach der Maßnahme*. Generell soll in dieser Itemgruppe die Sinnhaftigkeit und Effizienz der Lehre beurteilt werden.

Im prospektiven Teil der Evaluation wurden die Teilnehmer dazu aufgefordert, ihre Erwartungen und Einschätzung bezüglich der Lern- Quantität innerhalb der einzelnen Module wiederzugeben.

3.2.4.7.1 Lern- Quantität

Lern- Quantität Modul Rhetorik

Untersuchungsergebnisse Lern-Quantität Modul Rhetorik

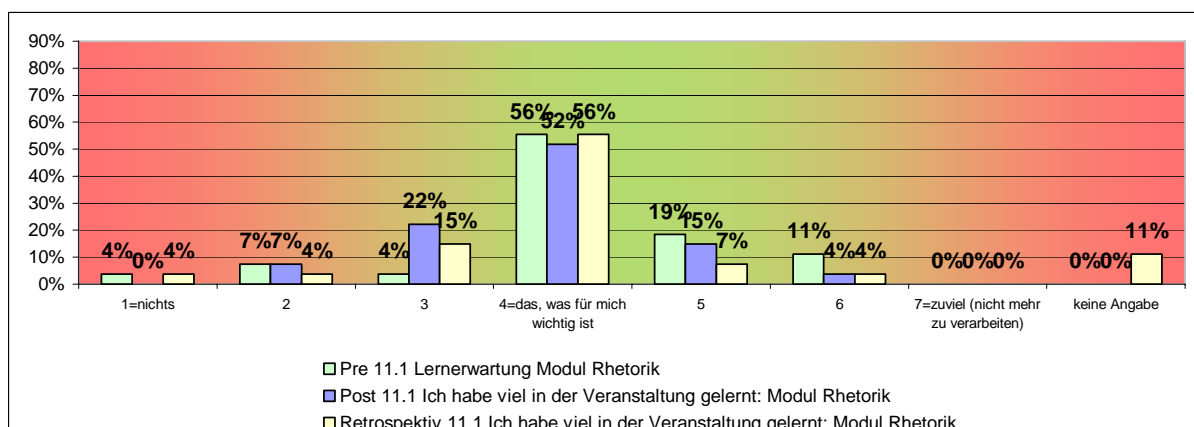


Diagramm 34: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Rhetorik

Analyse / Ergebnisinterpretation Lern-Quantität Modul Rhetorik (siehe Diagramm 34)

Die Einschätzungen der Lern-Quantität im Modul Rhetorik sind über alle drei Erhebungsstufen ähnlich. Der Median hat in allen drei Phasen den Wert 4, die Quartilabweichung vorher und nachher beträgt 1, in der Retrospektion ist die Quartilabweichung 0,3. In den beiden Erhebungen nach der Veranstaltung ist eine Tendenz dahingehend vorhanden, dass die Lernqualität in diesem Modul etwas zu gering war. Diese Tendenz ist allerdings nur sehr gering.

Lern- Quantität Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Lern-Quantität Modul Stimmbildung und Sprecherziehung

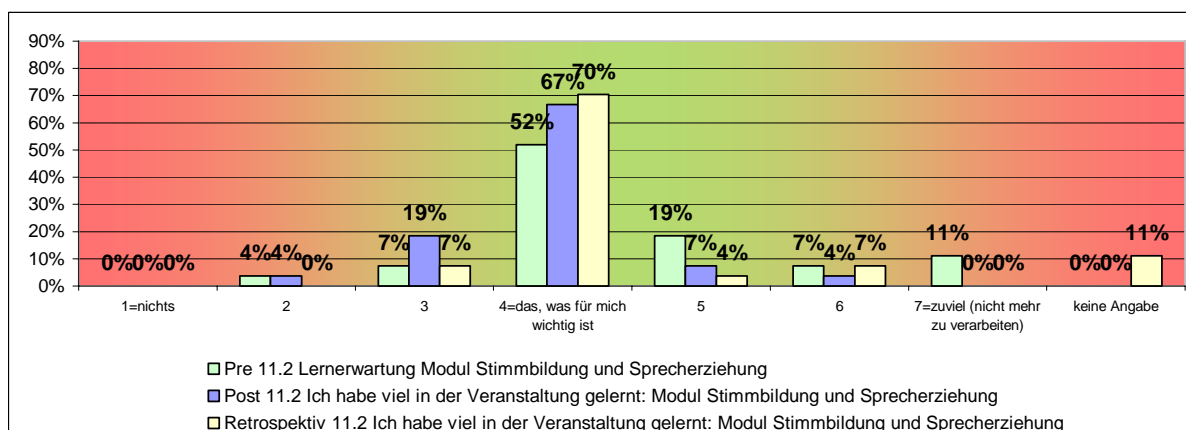


Diagramm 35: Untersuchungsergebnisse Evaluation PROFIL-TT Lernquantität Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Lern-Quantität Modul Stimmbildung und Sprecherziehung (siehe Diagramm 35)

Interessant bei der Verteilung der Teilnehmereinschätzung über die Lern- Quantität im Modul Stimmbildung und Sprecherziehung ist der Unterschied zwischen der Vorab- und Post- bzw. Retrospektions- Erhebung. Etwas mehr als die Hälfte der Teilnehmer bewerten ihre Lernerwartung vorab als genau das, was für sie wichtig ist. In den Erhebungen nach der Veranstaltung steigt die Prozentzahl dieser Teilnehmer auf 67% bzw. 70% an – in beiden Fällen liegt der Median bei 4 und die Quartilabweichung bei 0. Vor allem bezogen auf die Ergebnisse der retrospektiven Betrachtung kann für dieses Modul eine stark ausgeprägte Nachhaltigkeit festgestellt werden.

Lern- Quantität Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Lern-Quantität Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

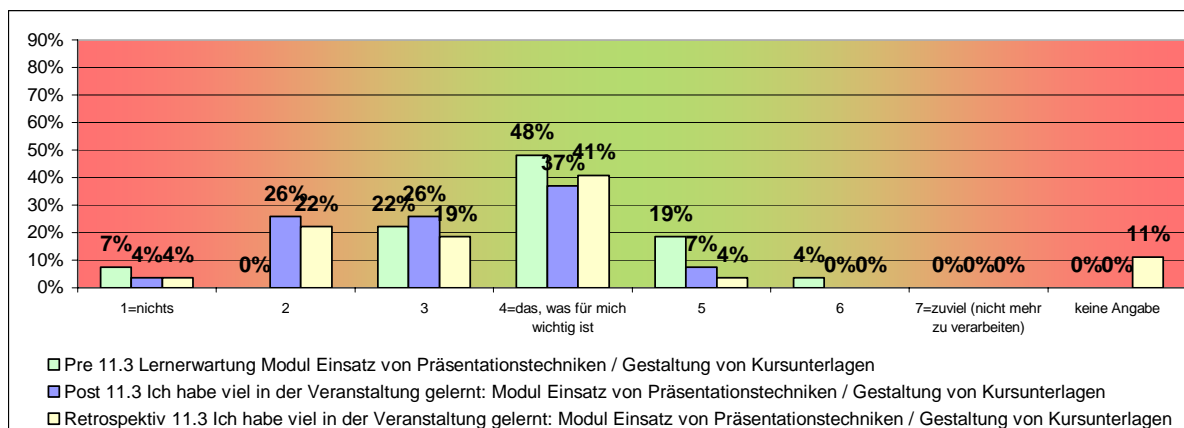


Diagramm 36: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Lern-Quantität Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 36)

Die Verteilung der Einschätzung zur Lern-Quantität im Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen zeigt in den beiden Erhebungen nach der Veranstaltung einen Median von 3 bzw. 3,5 bei einer Quartilabweichung von jeweils 2 – im Vergleich zum Median der Voreinschätzung 4 bei einer Quartilabweichung von 1. Über 50% der Teilnehmer beurteilten die Lern-Quantität dieses Moduls nach der Veranstaltung als zu gering (in unterschiedlich starken Ausprägungen). Die Gesamteinschätzung tendiert ebenfalls in diese Richtung.

Lern- Quantität Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Lern-Quantität Modul kohärenter Sprachgebrauch

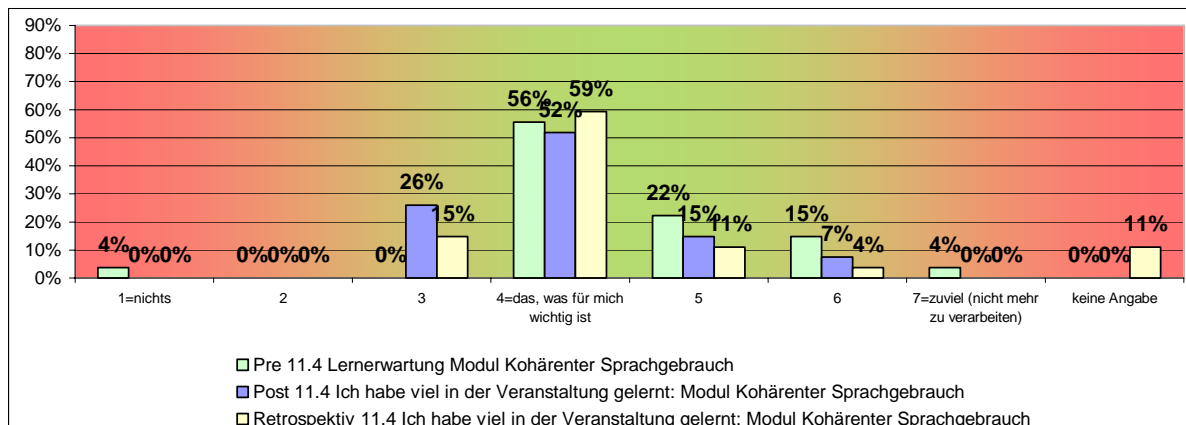


Diagramm 37: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Lern-Quantität Modul kohärenter Sprachgebrauch (siehe Diagramm 37)

Bei der Einschätzung der Lern-Quantität für das Modul kohärenter Sprachgebrauch liegt der Schwerpunkt in allen drei Erhebungsphasen in der Mitte der Skala (Median jeweils 4). Die Quartilabweichung wird mit jeder Stufe der Evaluation geringer (1 – 0,5 – 0). Die Werte aller drei Erhebungsstufen sind durchwegs ausgeglichen, wobei auch in diesem Modul der Schwerpunkt der retrospektiven Bewertung auffällt. Ähnlich wie beim Modul Stimmbildung und Sprecherziehung kann der relativ hohe Wert für das Skalenzentrum (59%) vor dem Hintergrund einer ausgeprägten Nachhaltigkeit interpretiert werden. Die Werte der Vorab- Erhebung tendieren geringfügig in Richtung einer quantitativen Überbeanspruchung.

Lern- Quantität Modul Lehrkompetenz

Untersuchungsergebnisse Lern-Quantität Modul Lehrkompetenz

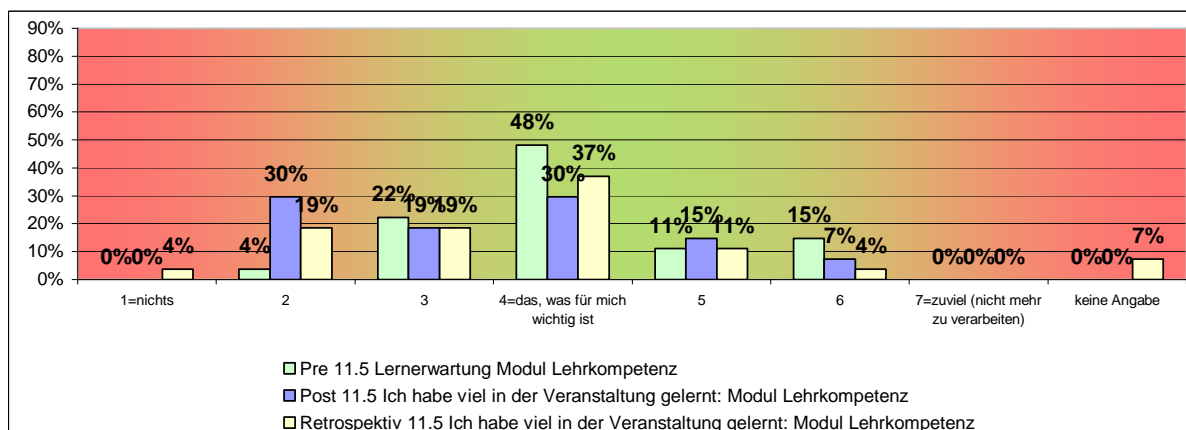


Diagramm 38: Untersuchungsergebnisse Evaluation PROFiL-TT Lernquantität Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Lern-Quantität Modul Lehrkompetenz (siehe Diagramm 38)

Bei der Untersuchung der Bewertungsergebnisse für die Lern-Quantität im Modul Lehrkompetenz fallen in erster Linie die Werte der Post- Erhebung auf. Bei einem Median von 4 tritt hierbei eine Quartilabweichung von 2 auf. Diese Abweichung zielt vorwiegend in die Richtung, dass die Lern-Quantität von einigen Teilnehmern als zu wenig umfangreich beurteilt worden ist. Der Wert für „ich habe in diesem Modul relativ wenig gelernt“ (2) ist dabei mit 30% relativ prominent. Diese Tendenz ist in der Vorab- Erhebung nicht erkennbar, in der retrospektiven Betrachtung abgeschwächt.

3.2.4.7.2 Subjektiver Wissenszuwachs

Subjektiver Wissenszuwachs Modul Rhetorik

Die Einschätzung des Wissenszuwachses sollte in allen fünf Modulen von PROFiL-TT sehr differenziert betrachtet werden, da in diesem Bereich bereits ein erster Blick auf die Deskription ein stark heterogenes Bild zeigt, das sich von den Verteilungen der vorhergehenden Werte unterscheidet.²⁵⁶

Die generell recht niedrigen Werte in der Retrospektions- Untersuchung können – bezogen auf die Vergessenskurve von Ebbinghaus – daherrühren, dass der subjektive Wissenszuwachs der Teilnehmer in einem Abstand von acht Wochen trotz der implizierten Reflektion eher gering eingeschätzt wird.²⁵⁷

Untersuchungsergebnisse Wissenszuwachs Modul Rhetorik

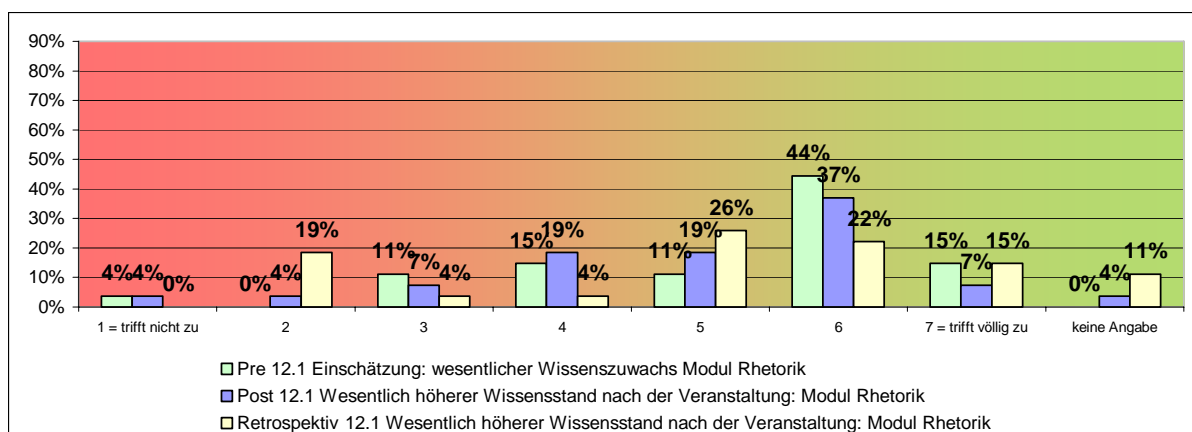


Diagramm 39: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Rhetorik

²⁵⁶ Die Datengrundlage und die darauf basierende weitere Datenverarbeitung wurden aufgrund dieser Unterscheidung zu den restlichen Werten mehrfach kontrolliert und überprüft, wobei keine Inkonsistenzen festgestellt wurden.

²⁵⁷ Vgl. Ebbinghaus H (1885)

Analyse / Ergebnisinterpretation Wissenszuwachs Modul Rhetorik (siehe Diagramm 39)

Betrachtet man die Vorabanschätzung des Wissenszuwachses für das Modul Rhetorik, kann man eine klare Tendenz erkennen. So wird im Vorherein für dieses Modul von mehr als der Hälfte der Teilnehmer ein relativ hoher Wissenszuwachs erwartet (Median 6 – Quartilabweichung 2). Diese Tendenz ist – allerdings in etwas abgeschwächter Form – auch an den Werten der Post- Erhebung abzulesen (Median 5 – Quartilabweichung 2). In der retrospektiven Erhebung verlagert sich der Schwerpunkt der Verteilung noch weiter nach links, wobei der Median immer noch 5 beträgt die Quartilabweichung allerdings auf den Wert 2,3 anwächst.

Subjektiver Wissenszuwachs Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Wissenszuwachs Modul Stimmbildung und Sprecherziehung

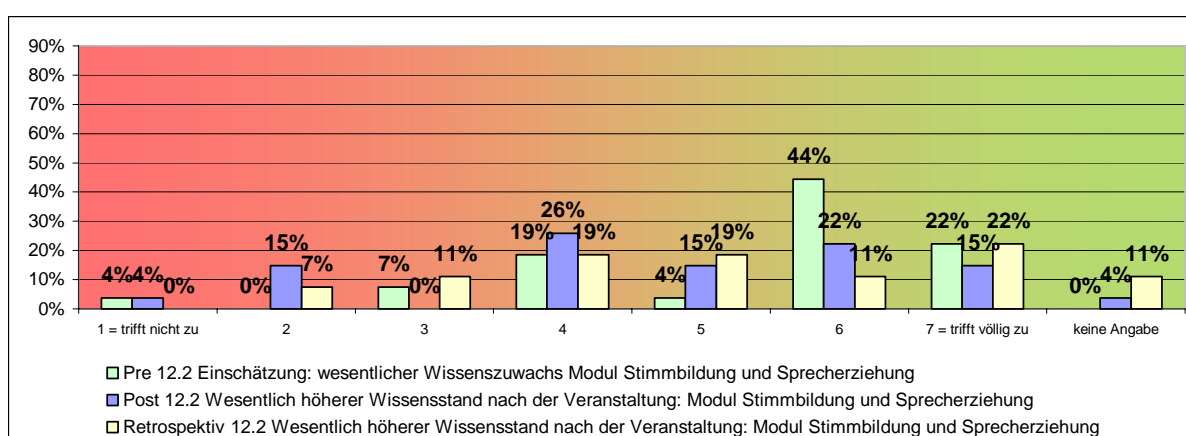


Diagramm 40: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Wissenszuwachs Modul Stimmbildung und Sprecherziehung (siehe Diagramm 40)

Die Verteilung der Werte im Modul Stimmbildung und Sprecherziehung ist – bezogen auf Median und Quartilabweichung – mit den Werten des Moduls Rhetorik identisch. (Median 6 – 5 – 5 / Quartilabweichung 2 – 2 – 2,3). Eine Differenz fällt allerdings bei der Betrachtung der Werte der Post- Erhebung auf. Die Wertverteilung über nahezu das gesamte Skalenspektrum (mit Ausnahme der Werte 1 und 3) scheint etwas ausgewogener zu sein. Auch für diese Einschätzungen fehlt die klare Eindeutigkeit der Untersuchungsdaten, die bis zur Frage 11 beobachtet werden kann.

Subjektiver Wissenszuwachs Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Wissenszuwachs Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

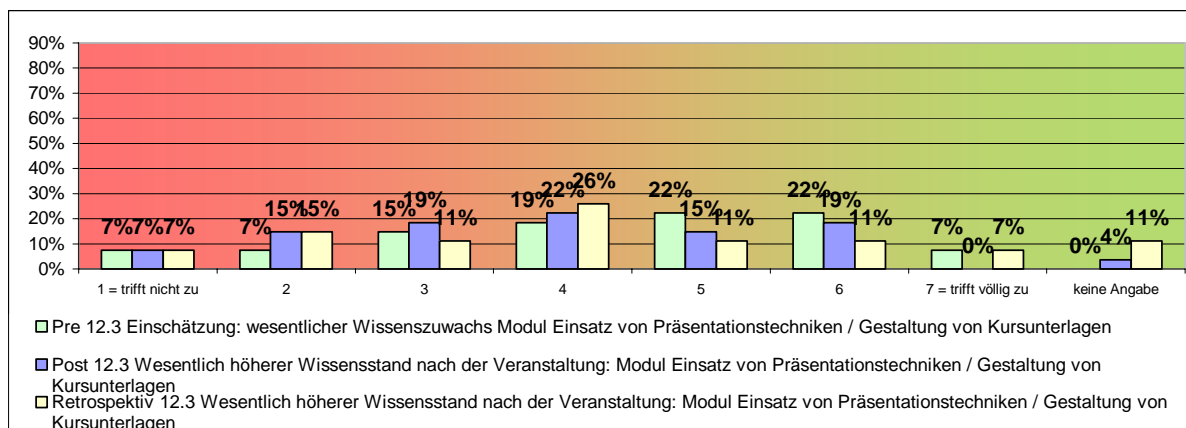


Diagramm 41: Untersuchungsergebnisse Evaluation PROFIL-TT Wissenszuwachs Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Wissenszuwachs Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 41)

Die Verteilung der Werte im Bezug auf den subjektiven Wissenszuwachs der Veranstaltungsteilnehmer im Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen ist, ähnlich wie die der beiden vorhergehenden Module, breit gefächert.

Die Vorabeeschätzung der Teilnehmer hat ihre beiden Gipfel bei den Werten 5 und 6, wobei der Median hierfür eindeutig auf dem Wert 5 liegt. Dennoch ist eine Tendenz in Richtung „geringer Erwartung auf Wissenszuwachs“ erkennbar (Quartilabweichung 3). Der Median der Post- Erhebung liegt bei 4 (mit einer Quartilabweichung von 2). Auch in diesem Bereich sind die Werte sehr heterogen verteilt. Die Ausprägung dieser Verteilung ist in den Werten der retrospektiven Erhebung noch deutlicher (Median 4 – Quartilabweichung 2,3).

Subjektiver Wissenszuwachs Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Wissenszuwachs Modul kohärenter Sprachgebrauch

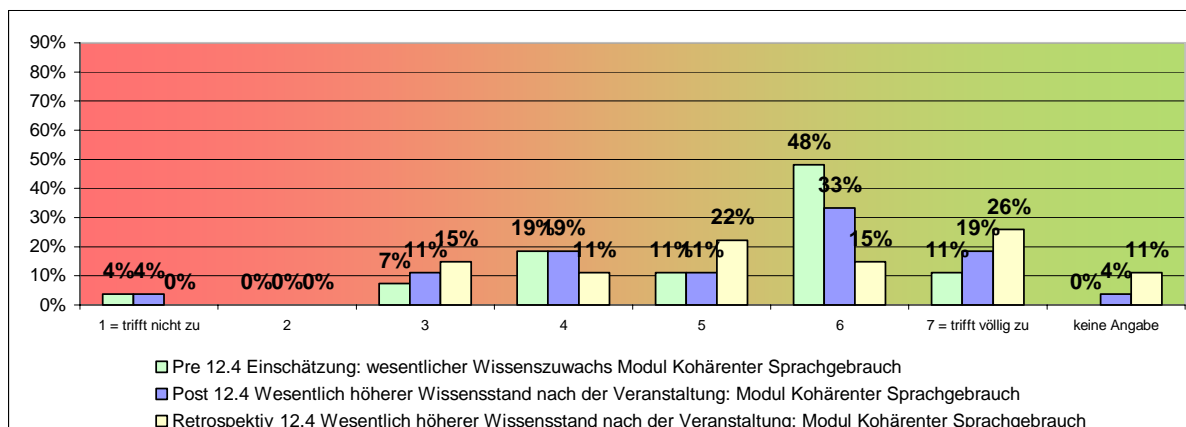


Diagramm 42: Untersuchungsergebnisse Evaluation PROFIL-TT Wissenszuwachs Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Wissenszuwachs Modul kohärenter Sprachgebrauch (siehe Diagramm 42)

Die Einschätzung der Teilnehmer (bezogen auf den Wissenszuwachs im Modul kohärenter Sprachgebrauch) sind ähnlich verteilt, wie in den vorhergehenden Modulen. Der Zentralwert liegt in der Vorher- und Nachher- Erhebung jeweils bei 6 mit einer Quartilabweichung von 2. Der Median für die retrospektive Betrachtung liegt mit dem Wert 5 etwas darunter, wobei diese Erhebung eine größere Quartilabweichung (Wert 3) aufweist.

Im Rahmen der Vorab- Untersuchung weisen die Werte – 48% der Teilnehmer haben für dieses Modul die Erwartung, eines relativ hohen, 11% eines sehr hohen Wissenszuwachses – eine deutliche Tendenz auf. In der Einschätzung nach dem Seminar ist diese Tendenz etwas abgeschwächt (analog zu den prospektiven Werten 33% - 19%). In der retrospektiven Sicht verteilen sich die Werte auf den Skalenbereich 3 bis 7.

Subjektiver Wissenszuwachs Modul Lehrkompetenz

Untersuchungsergebnisse Wissenszuwachs Modul Lehrkompetenz

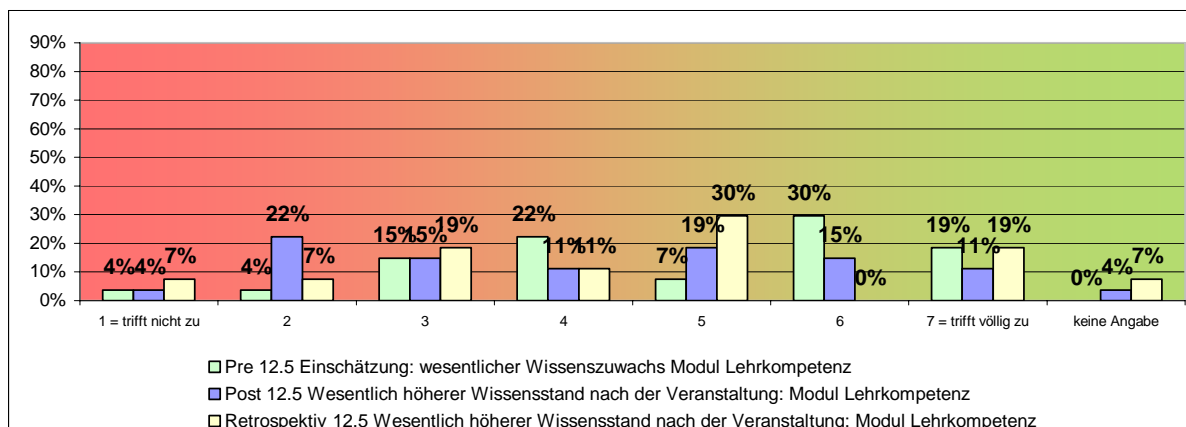


Diagramm 43: Untersuchungsergebnisse Evaluation PROFiL-TT Wissenszuwachs Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Wissenszuwachs Modul Lehrkompetenz (siehe Diagramm 43)

Für das Modul Lehrkompetenz ist die Erwartung bezogen auf Wissenszuwachs bereits in der Prospektion etwas niedriger (Median 5 – Quartilabweichung 2), wobei die Einschätzungsschwerpunkte auf den Skalenwerten 3 und 4 (15% und 22%), sowie 6 und 7 (30% und 19%) liegen. Die Beurteilungen des Wissenszuwachses direkt nach der Veranstaltung verteilen sich gleichmäßig auf die Skalenwerte 2 bis 7. Die Einschätzung des subjektiven Wissenszuwachses ist bei einem Median von 4 mit einer Quartilabweichung von 3,5 stark gestreut. Diese Verteilung verschiebt sich in der retrospektiven Betrachtung nach rechts (Median 5 - Quartilabweichung 2).

3.2.4.8 Verarbeitung

Die Dimension *Verarbeitung* bezieht sich auf die Veranstaltungsinhalte. Die Fragestellungen zielen hauptsächlich auf die Vorgehensweise bei der Vermittlung der Lehrinhalte. Es soll hinterfragt werden, ob und inwiefern Veranstaltungsteilnehmer aktiv in den Prozess der Inhaltsvermittlung eingebunden werden und diesen auch kritisch begleiten können.

In der Prospektion wurden diese Items dazu benutzt, die Relevanz von Interaktivität und den Stellenwert einer kritischen Betrachtung der Veranstaltungsinhalte abzufragen.

Anregung zum aktiven Mitdenken

Untersuchungsergebnisse Anregung zum aktiven Mitdenken

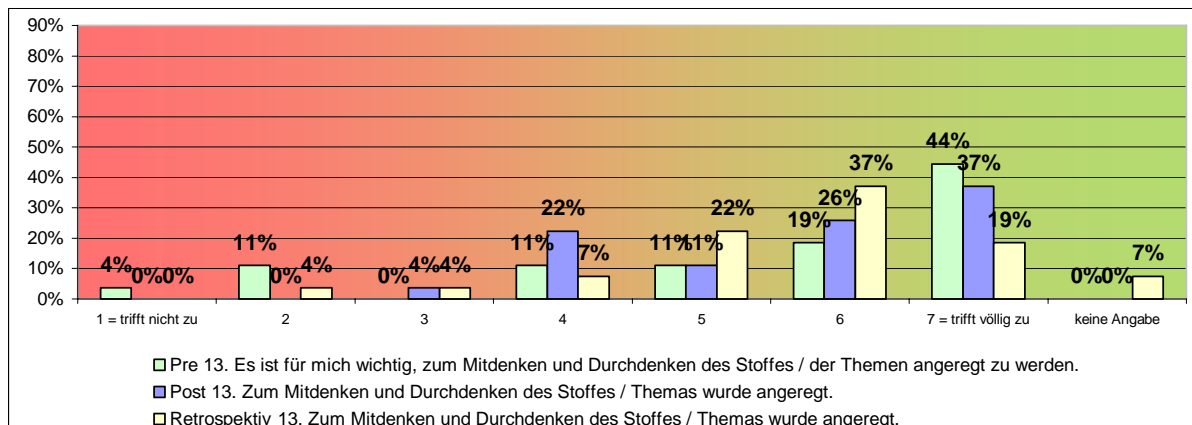


Diagramm 44: Untersuchungsergebnisse Evaluation PROFiL-TT aktives Mitdenken

Analyse / Ergebnisinterpretation Anregung zum aktiven Mitdenken (siehe Diagramm 44)

Der Reiz zur aktiven Mitarbeit wird von den meisten Teilnehmern vorab als sehr hoch eingeschätzt.²⁵⁸ Der Median liegt dementsprechend bei 6, die Quartilabweichung bei 2,5.

Eine ähnliche Verteilung lässt sich in der Erhebung nach der Veranstaltung wieder finden (Median und Quartilabweichung weisen dieselben Werte auf), wobei eine leichte Verschiebung der Werte von „sehr hoher Relevanz“ zu „hoher Relevanz“ beobachtet werden kann.²⁵⁹

Retrospektiv beurteilen die Teilnehmer die Anregung zur aktiven Mitarbeit in der Veranstaltung als relativ stark ausgeprägt. Dies zeigt ein Median von 6 mit einer Quartilabweichung von 1.

²⁵⁸ 44% der Teilnehmer weisen der Anregung zu aktiven Mitarbeit vorab eine sehr hohe, 19% eine hohe Relevanz zu.

²⁵⁹ Die Werte liegen hierbei bei 37% (Skalenwert 7) und 26% (Skalenwert 6) in der Beurteilung der Anregung zur Mitarbeit.

Kritische Auseinandersetzung mit den Themen

Untersuchungsergebnisse kritische Auseinandersetzung mit den Themen

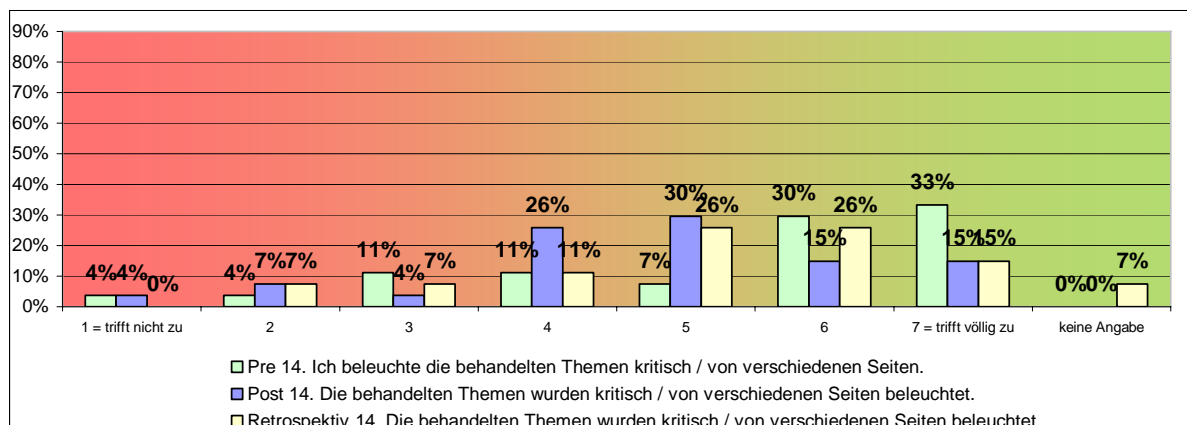


Diagramm 45: Untersuchungsergebnisse Evaluation PROFiL-TT kritische Auseinandersetzung

Analyse / Ergebnisinterpretation kritische Auseinandersetzung mit den Themen (siehe Diagramm 45)

Die Selbsteinschätzung der Teilnehmer bezogen auf ihre kritische Auseinandersetzung mit den Kursthemen vor der Veranstaltung hat einen relativ hohen Median (Wert 6), wobei die Quartilabweichung für diese Erhebung auch stark ausgeprägt ist (Wert 3). Der eigene, kritische Umgang mit Fachthemen wird von jeweils etwa einem Drittel der Teilnehmer als stark, bzw. sehr stark ausgeprägt eingeschätzt. Übertragen auf die kritische Behandlung von Themen innerhalb der Veranstaltung ist die Beurteilung etwas weniger stark ausgeprägt (Median 5 – Quartilabweichung 2). Dieser Trend hält auch in der retrospektiven Betrachtung an. Diese Ergebnisse spiegeln eine kritische Betrachtung der Inhalte im Rahmen der Veranstaltung wider.

3.2.4.9 Betreuung

Der Punkt *Betreuung* bezieht sich sowohl auf die Veranstaltung selbst, als auch auf das Umfeld. Im eigentlichen Rahmen der Veranstaltung soll gemessen werden, ob ein Dozent den Veranstaltungsteilnehmern Rückmeldung gibt, bzw. ob diese Rückmeldung für die Teilnehmer nützlich ist. Außerhalb der Veranstaltung bezieht sich die Betreuung auf die Präsenz des Dozenten gegenüber den Teilnehmern (zum Beispiel bei inhaltlichen und organisatorischen Fragen vor oder im Lauf der Veranstaltung).

3.2.4.9.1 Feedback

Untersuchungsergebnisse Feedback

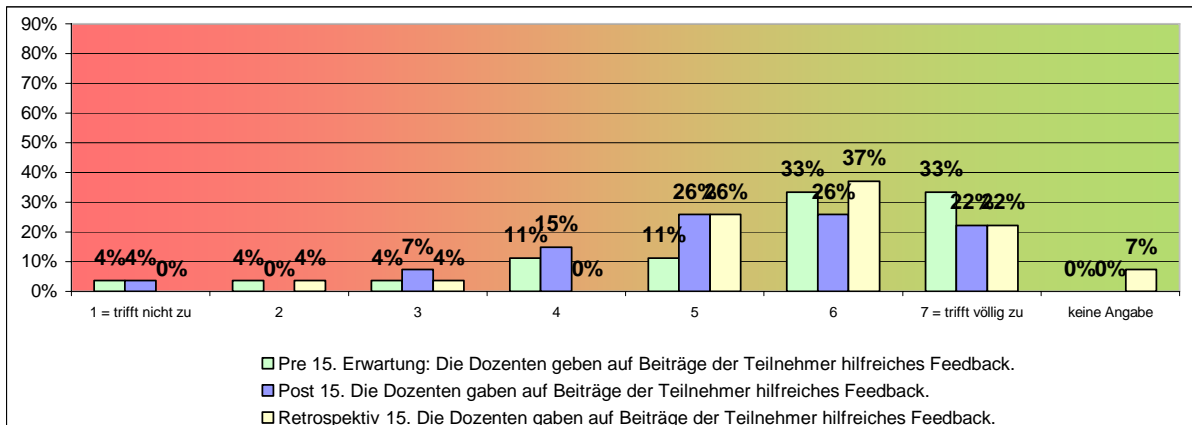


Diagramm 46: Untersuchungsergebnisse Evaluation PROFiL-TT Feedback

Analyse / Ergebnisinterpretation Feedback (siehe Diagramm 46)

Die Vorab einschätzung der Teilnehmer in Bezug auf das Feedback ist relativ stark ausgeprägt. Die Teilnehmer erwarten von den Dozenten der Veranstaltung hilfreiches Feedback auf eigene Beiträge. Der Median in der Vorher-Befragung beträgt dementsprechend den Wert 6 – bei einer Quartilabweichung von 2. Die Beurteilung des tatsächlich erhaltenen Feedbacks tendiert ebenso in die Richtung, dass Feedback gegeben wurde (Median der Werte 6 – Quartilabweichung 1,5). Diese Tendenz verstärkt sich noch in der Retrospektiv- Beurteilung. Der Median hier liegt ebenso bei 6, die Quartilabweichung beträgt dabei 1. Die Ergebnisse der Post- und Retrospektionsbetrachtung lassen darauf schließen, dass den Teilnehmern im Rahmen der Veranstaltung hilfreiches Feedback auf ihre Beiträge gegeben wurde.

3.2.4.9.2 Betreuung außerhalb der Veranstaltung

Untersuchungsergebnisse Betreuung außerhalb der Veranstaltung

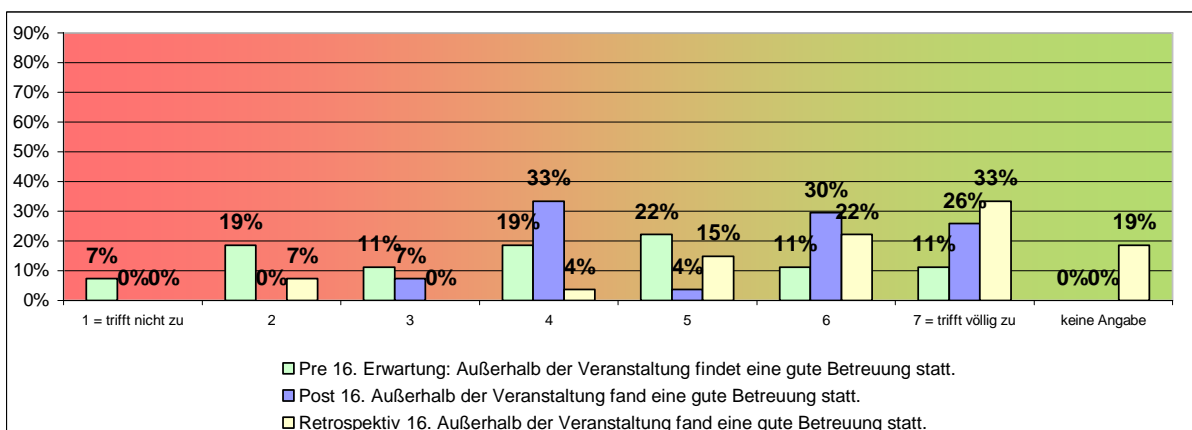


Diagramm 47: Untersuchungsergebnisse Evaluation PROFiL-TT Betreuung außerhalb der Veranstaltung

Analyse / Ergebnisinterpretation Betreuung außerhalb der Veranstaltung (siehe Diagramm 47)

Die Teilnehmer schätzen die Betreuung außerhalb der Veranstaltung vorab stark heterogen ein. Ein – relativ geringer – Teil der untersuchten Personengruppe erwartet eine Betreuung über den Rahmen der Veranstaltung hinaus, ein anderer Teil geht nicht davon aus, dass eine solche Betreuung im Rahmen von PROFiL-TT stattfindet. Der Median spiegelt diese Verteilung der Werte wider (Wert 4 bei einer Quartilabweichung von 2,5). Aus den Werten der Post- und Retrospektion- Erhebung wird allerdings ersichtlich, dass die Teilnehmer eine gute Betreuung außerhalb der Veranstaltung durchaus wahrgenommen haben (Median post und retrospektiv jeweils 6, Quartilabweichung 2,5, bzw. 2).

3.2.4.10 Beteiligung

Die Dimension *Beteiligung* erhebt das Engagement der Veranstaltungsteilnehmer mit dem Fokus auf die Veranstaltung selbst.

Untersuchungsergebnisse Beteiligung

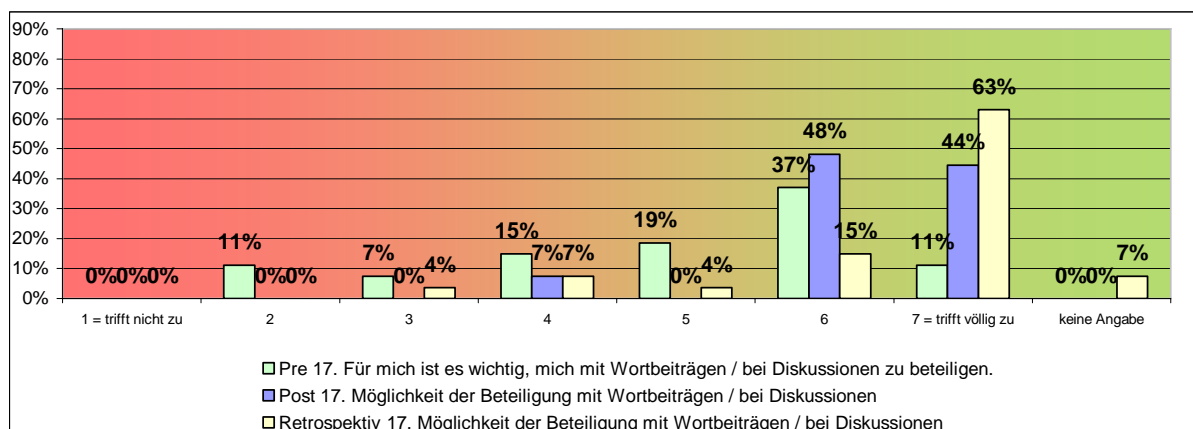


Diagramm 48: Untersuchungsergebnisse Evaluation PROFiL-TT Beteiligung

Analyse / Ergebnisinterpretation Beteiligung (siehe Diagramm 48)

Bei der Vorab einschätzung der Teilnehmer bezogen auf eine interaktive Beteiligung am Kursverlauf kann eine Verteilung der Werte über fast das komplette Skalenspektrum beobachtet werden.³³ Der Median liegt dementsprechend bei 5 (mit einer Quartilabweichung von 2), und somit etwas höher als eine teilweise Zustimmung. Aus diesen Werten kann geschlossen werden, dass eine relativ große Gruppe von Teilnehmern die Veranstaltung mit der Erwartung einer mehr monologisch geführten und eher weniger dialogisch geprägten Informationsvermittlung besucht haben.

Die Möglichkeit zur Beteiligung mit Wortbeiträgen und an Diskussionen wurde allerdings in den beiden Erhebungsstufen nach der Veranstaltung sehr eindeutig bestätigt (Median Post: 6 – Quartilabweichung 1; Median Retrospektiv: 7, Quartilabweichung 1). Die Teilnehmer schätzten ihre Möglichkeit einer dialogischen Beteiligung am Kursverlauf als sehr hoch ein. Diese Einschätzung wurde in retrospektiver Sicht noch verstärkt.

3.2.4.11 Kommunikative Unterrichtsformen

Kommunikative Unterrichtsformen werden vorwiegend in Seminaren oder kleineren Veranstaltungen verwendet.²⁶⁰ Mittels dieser Dimension soll das Vorhandensein und die Möglichkeit beurteilt werden, Diskussionen oder Gruppenarbeiten innerhalb einer Veranstaltung in ausreichender Form abzuhalten. Der Bewertungsfokus liegt vor allem auf der angemessenen Quantität von Diskussionen und dem ausreichenden Einsatz von kommunikativen Lehrformen.

3.2.4.11.1 Diskussionen – quantitativ

Untersuchungsergebnisse Diskussionen – quantitativ

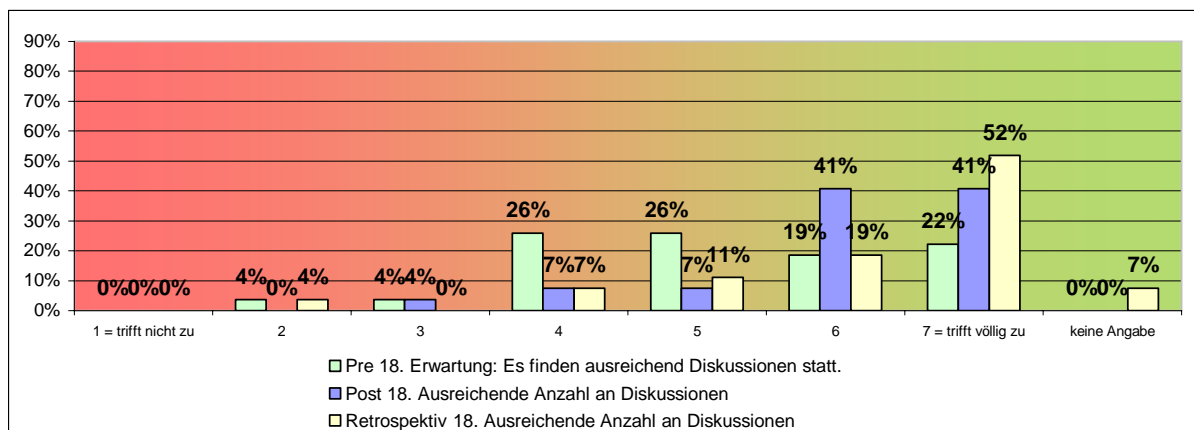


Diagramm 49: Untersuchungsergebnisse Evaluation PROFiL-TT Diskussionen - quantitativ

Analyse / Ergebnisinterpretation Diskussionen – quantitativ (siehe Diagramm 49)

Die Erwartung der Teilnehmer, dass im Rahmen der Veranstaltung ausreichend Diskussionen stattfinden, ist in der Vorab- Erhebung – ähnlich wie die Einschätzung zur Interaktivität (Frage 17) – etwas stärker als „durchschnittlich“ ausgeprägt (Median 5, Quartilabweichung 2). Die Tendenzen in der Post- und Retrospektions- Erhebung sind ebenso analog zu den Ergebnissen der „Beteiligung“ (Median 6/7, Quartilabweichung jeweils 1). Auch in diesem Fall kann festgestellt werden, dass die Teilnehmer vorab weniger Interaktion – in Form von Diskussionen – erwartet haben. Die Beurteilungen nach der Veranstaltung weisen eindeutig darauf hin, dass die Menge an Diskussionen im Rahmen von PROFiL-TT ausreichend bzw. für die Teilnehmer angenehm war. Diese subjektive Einschätzung wird mit zunehmendem zeitlichem Abstand abermals verstärkt.

²⁶⁰ PROFiL-TT ist als ebensolche *kommunikative Unterrichtsform* konzipiert worden.

3.2.4.11.2 Einsatz kommunikativer Unterrichtsformen

Untersuchungsergebnisse kommunikative Unterrichtsformen

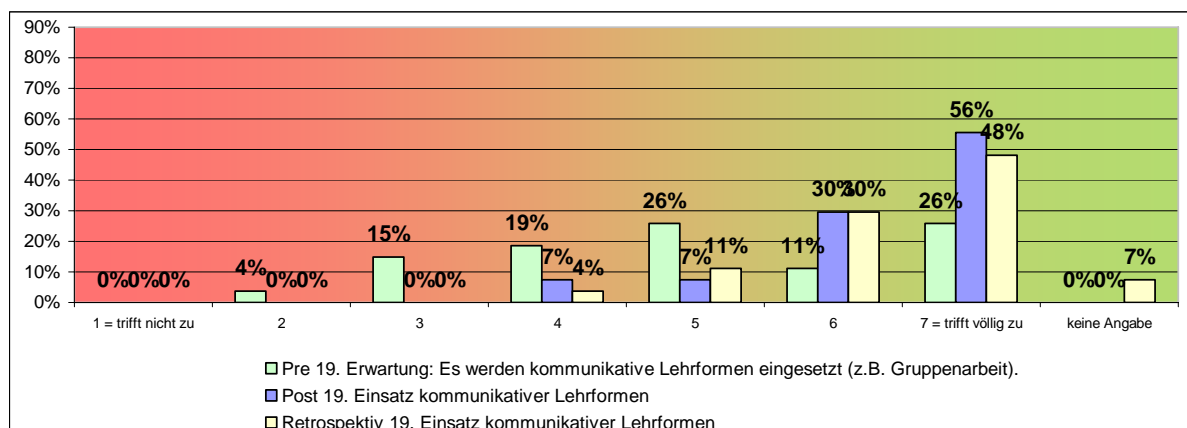


Diagramm 50: Untersuchungsergebnisse Evaluation PROFiL-TT kommunikative Unterrichtsformen

Analyse / Ergebnisinterpretation kommunikative Unterrichtsformen (siehe Diagramm 50)

Der Unterschied der subjektiven Einschätzungen bezüglich Einsatz kommunikativer Unterrichtsformen zwischen der Vorab- Erhebung und den Erhebungen nach der Veranstaltung, ist deutlich sichtbar. Die Teilnehmereinschätzung vorab ist stark heterogen (Median 5 bei einer Quartilabweichung von 2,5); die Verteilung dieser Erhebung ist ausgewogen über die Skala verteilt. Die Beurteilungen nach der Veranstaltung zeigen wiederum eindeutig, dass der Einsatz von kommunikativen Unterrichtsformen sehr intensiv wahrgenommen wurde. Dementsprechend liegen die Werte für den Median bei einer relativ niedrigen Quartilabweichung sehr hoch (Median bei beiden Erhebungsstufen 7, Quartilabweichung jeweils 1).

3.2.4.12 Dozentenbewertung

Im Rahmen der Dozentenbewertung haben die Teilnehmer die Möglichkeit, nicht die Person der einzelnen Dozenten zu bewerten, sondern den Dozent als *Instanz*, Ressource und somit als Bestandteil der Veranstaltung.

Der Hintergrund dieser relativ kurzen Abhandlung der Dozentenbewertung besteht in den Grundzügen der konzeptionellen Planung von PROFiL-TT. Durch die modulare Entwicklung der gesamten Veranstaltung sollte ein Konzept generiert werden, das dozentenunabhängig ist und aufgrund dieser personellen Unabhängigkeit beliebig „ervielfältigt“ werden kann.

Da allerdings der Dozent in der Umsetzung einer Weiterbildungsmaßnahme eine tragende Rolle spielt,²⁶¹ wurde auf den kompletten Entfall der Dozentenbewertung verzichtet, um zumindest eine Tendenz im aktuell vorliegenden Projekt PROFiL-TT bezüglich der Kompetenz des Lehrpersonals aufzeigen zu können. Die „Dozentenbeur-

²⁶¹ Siehe Kapitel 2.11 (HILVE)

teilung“ ist auf die einzelnen Module aufgeteilt, wobei auch hier – aufgrund von personellem Wechsel – vor allem tendenzielle Aussagen getroffen werden können. Die Möglichkeit, den Dozenten als *Inстанz* bewerten zu können, wurden von einigen Teilnehmer (im Rahmen der qualitativen Anmerkungen der der Post- und Retrospektions- Erhebungen) durchaus kritisch betrachtet, war aber von Seiten des projektleitenden Gremiums nachdrücklich gewünscht worden. Diesem Wunsch wurde in der Evaluation im Rahmen eines Zusatzmoduls entsprochen.

Die Einschätzung der Dozenten wird anhand der „Interessantheit“ der einzelnen Module erhoben. Es wird somit in erster Linie die Art und Weise eingeschätzt, in der ein Modul durchgeführt worden ist, bzw. in der Prospektion die Erwartung, wie interessant ein Modul für die einzelnen Teilnehmer werden wird. Die Prospektion dient dabei in erster Linie dazu, eine Vorab- Tendenz bezüglich der Darbietung der einzelnen Module abzufragen – die „Interessantheit“ des Themas spielt hierbei natürlich auch eine wesentliche Rolle.²⁶²

Modul Rhetorik

Untersuchungsergebnisse Dozenteneinschätzung / -beurteilung Modul Rhetorik

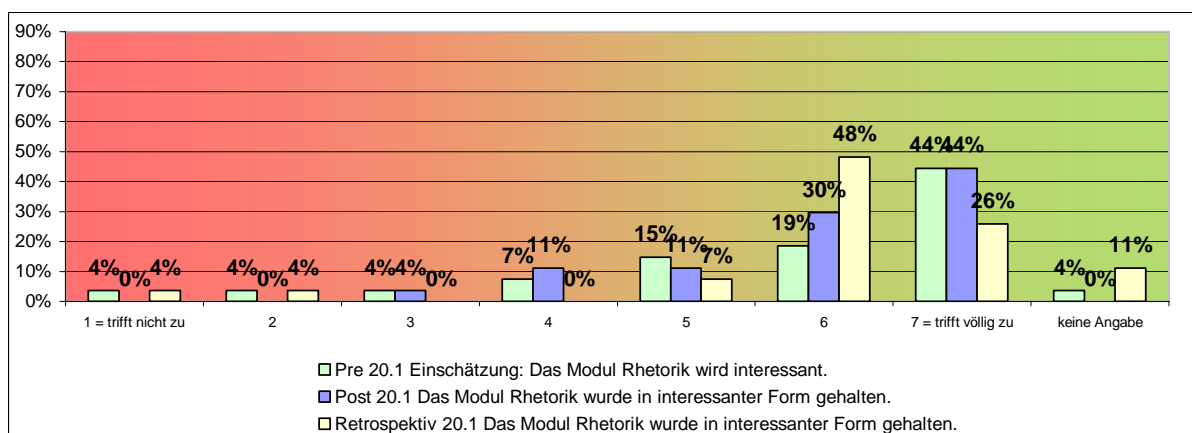


Diagramm 51: Untersuchungsergebnisse Evaluation PROFIL-TT Dozenteneinschätzung /-beurteilung Modul Rhetorik

Analyse / Ergebnisinterpretation Dozenteneinschätzung / -beurteilung Modul Rhetorik (siehe Diagramm 51)

Bei der Dozenteneinschätzung für das Modul Rhetorik kann man über alle drei Erhebungsstufen eine klare Tendenz feststellen, die ihr Zentrum bei Median 6 hat. Die Erwartung für dieses Modul ist hoch, die Teilnehmer messen dem Thema Rhetorik vor dem Hintergrund „Dozent“ eine hohe Relevanz zu. Diese Relevanz kristallisiert sich in den Werten der Post- Erhebung noch weiter heraus (die Quartilabweichung bei gleichem Medianwert fällt von 2 auf 1,5; betrachtet man die Werte der Verteilung, kann man einen deutlichen Anstieg von Skalenwert 3 bis 7 erkennen). In der retrospektiven Betrachtung konsolidiert sich die Verteilung weiter um den Median 6 (mit einer Quartilabweichung von 1).

²⁶² Siehe Kapitel 2.11

Modul Stimmbildung und Sprecherziehung

Untersuchungsergebnisse Dozenteneinschätzung / -beurteilung Modul Stimmbildung und Sprecherziehung

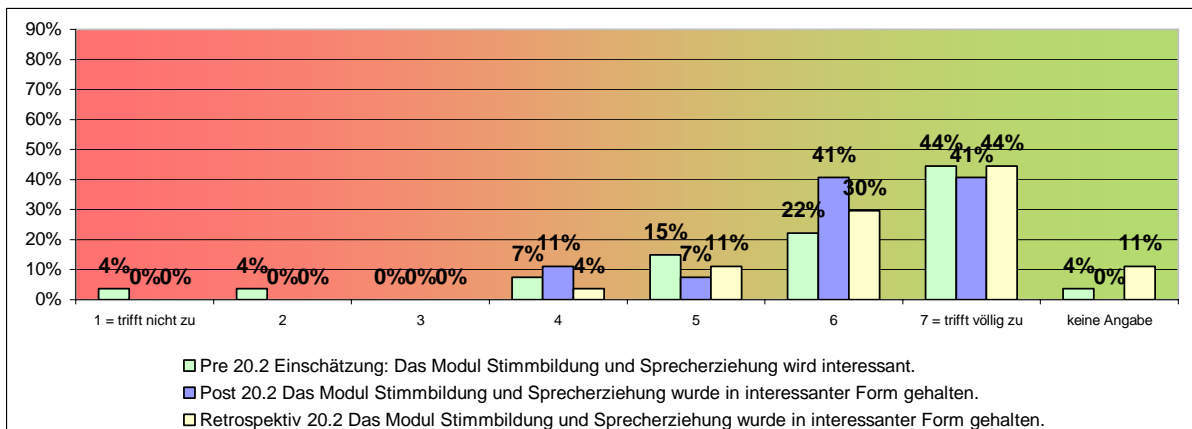


Diagramm 52: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /-beurteilung Modul Stimmbildung und Sprecherziehung

Analyse / Ergebnisinterpretation Dozenteneinschätzung / -beurteilung Modul Stimmbildung und Sprecherziehung (siehe Diagramm 52)

Die Verteilung der Werte für die Dozenteneinschätzung im Modul Stimmbildung und Sprecherziehung kann – abgesehen von geringfügigen Abweichungen der Werte im Modul Rhetorik – analog betrachtet werden. Die Relevanz ist auch hier klar erkennbar. Bei der Verteilung der Werte für dieses Modul fällt allerdings auf, dass die Beurteilung nach dem Seminar sich sehr stark auf den Wert 6 und 7 konzentriert (beide von 41% der Teilnehmer beurteilt)²⁶³. Bei den Werten der retrospektiven Erhebung fällt in erster Linie der hohe Median auf (6,5 bei einer Quartilabweichung von 1), der nachhaltig auf eine sehr interessante Darbietung der Modul Inhalte schließen lässt.

²⁶³ Bei einem Median mit Wert 6 und einer Quartilabweichung 1

Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Untersuchungsergebnisse Dozenteneinschätzung / -beurteilung Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

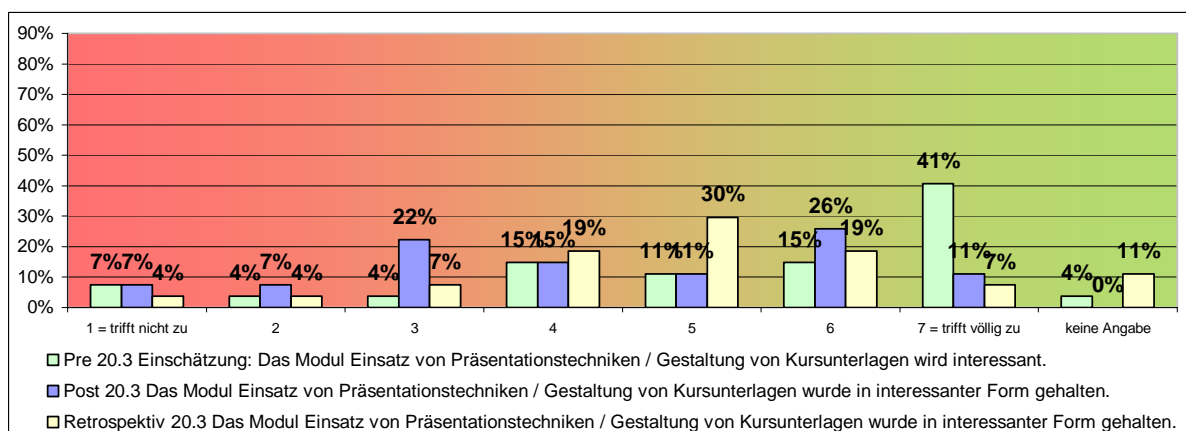


Diagramm 53: Untersuchungsergebnisse Evaluation PROFIL-TT Dozenteneinschätzung /-beurteilung Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

Analyse / Ergebnisinterpretation Dozenteneinschätzung / -beurteilung Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen (siehe Diagramm 53)

Die Werte der Vorher- Erhebung zur Dozenteneinschätzung für das Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen weisen einen ähnlichen Charakter auf, wie die der vorangehenden Module, allerdings ist die Quartilabweichung hier vorab relativ hoch und spiegelt eine leichte Heterogenität der Verteilung wider (Median 6 – Quartilabweichung 3). Aus diesen Werten kann geschlossen werden, dass der Interessenschwerpunkt der Teilnehmer für dieses Modul bereits vorab etwas verschoben war. Die Werte der Post- Erhebung liegen für die Dozenteneinschätzung dieses Moduls bei einem relativ niedrigen Median (Wert 4) vergleichsweise weit gestreut (Quartilabweichung 3) und bewegen sich im Rahmen der kompletten Einschätzungsskala.²⁶⁴ Im Rahmen der retrospektiven Erhebungen steigt der Median auf den Wert 5 (bei einer Quartilabweichung 2)

²⁶⁴ Hinweise auf den etwas niedrigen Median finden sich in der Auswertung der qualitativen Daten, wo vereinzelt darauf hingewiesen wird, dass der Vortragstil des Dozenten in diesem Modul etwas frontal und monoton wirkt (siehe Anhang IV: Verbatims).

Modul kohärenter Sprachgebrauch

Untersuchungsergebnisse Dozenteneinschätzung / -beurteilung Modul kohärenter Sprachgebrauch

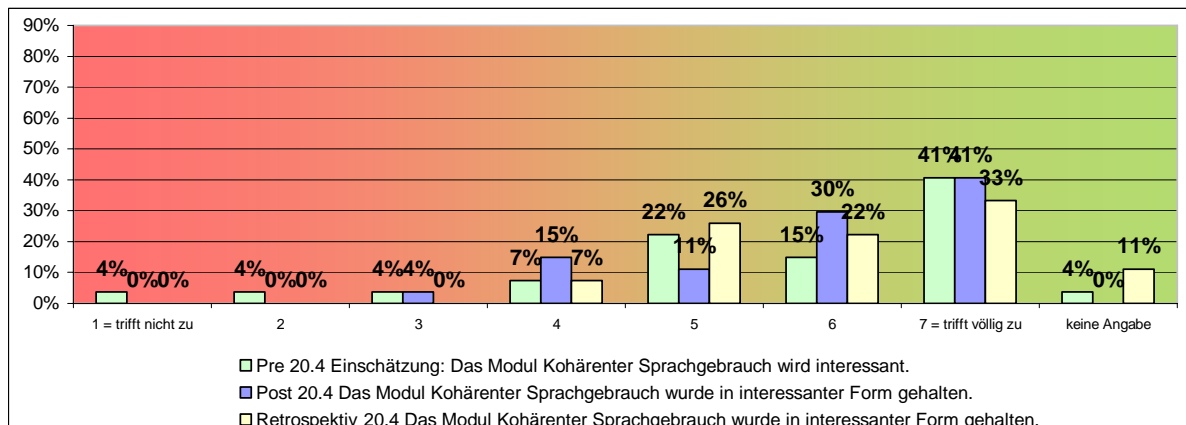


Diagramm 54: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /-beurteilung Modul kohärenter Sprachgebrauch

Analyse / Ergebnisinterpretation Dozenteneinschätzung / -beurteilung Modul kohärenter Sprachgebrauch (siehe Diagramm 54)

Die Dozenteneinschätzung für das Modul kohärenter Sprachgebrauch ist über alle drei Erhebungsstufen konsistent. Der Median liegt jeweils bei 6, die Quartilabweichung bei 2. Die Vorab-Einschätzung ist dabei analog zu den restlichen Modulen. Eine geringfügige Unterscheidung der beiden Erhebungen nach der Veranstaltung kann man in Bezug auf eine moderate Verlagerung der Werte in der Retrospektion auf die Skalenwerte 5 bis 7 feststellen; in der Post-Erhebung bezieht die Ausdehnungsspanne die Werte 4 bis 7 stärker mit ein. Diese Durchgängigkeit und Konstanz der Werte begründet sich vermutlich darin, dass der Dozent für dieses Modul nicht gewechselt wurde. Das Modul kohärenter Sprachgebrauch wurde komplett über den Projektverlauf vom selben Dozenten durchgeführt.

Modul Lehrkompetenz

Untersuchungsergebnisse Dozenteneinschätzung / -beurteilung Modul Lehrkompetenz

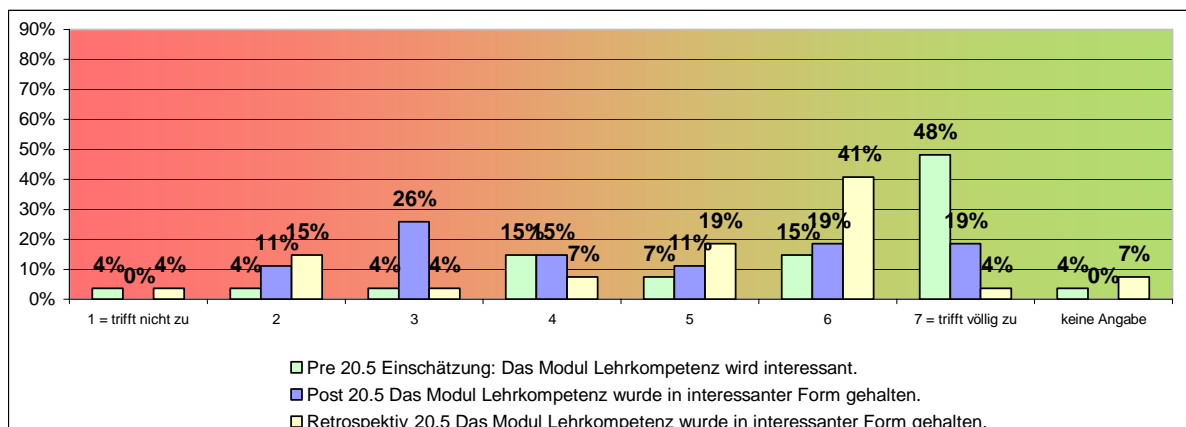


Diagramm 55: Untersuchungsergebnisse Evaluation PROFiL-TT Dozenteneinschätzung /-beurteilung Modul Lehrkompetenz

Analyse / Ergebnisinterpretation Dozenteneinschätzung / -beurteilung Modul Lehrkompetenz (siehe Diagramm 55)

Die Vorabeeschätzung der dozentenbedingten Interessantheit des Moduls Lehrkompetenz hat im Vergleich mit den restlichen Modulen den höchsten Median (Wert 6,5), allerdings auch eine verhältnismäßig hohe Streuung. Diese Werte können als hohe Erwartungshaltung der Teilnehmer an die Dozenten interpretiert werden. Im Vergleich zu den Werten der Vorher- Erhebung zentrieren sich die Erhebungsdaten der Post- Untersuchung um den Median 4 (mit einer gleichfalls hohen Quartilabweichung von 3). Die Verteilung ist dabei nahezu gleichmäßig über die Skalenwerte 2 bis 7 gestreut. Diese Diskrepanz ist verhältnismäßig auffällig.²⁶⁵ Die retrospektiven Werte bewegen sich bei der Dozenteneinschätzung dieses Moduls um den Zentralwert 5 (bei einer Quartilabweichung 2).

3.2.4.13 Allgemeineinschätzung

Als Fazit über die gesamte Veranstaltung wird mit der Dimension *Allgemeineinschätzung* eine Gesamtbeurteilung durchgeführt. Diese Kompletteneinschätzung ist inhaltlich betrachtet weit gestreut und zielt vor allem auf die Gesamtwirkung der Veranstaltung. Fokussiert wird dabei die Einschätzung, inwiefern die Teilnahme an der Maßnahme einen persönlichen Nutzen bewirken konnte. Dieser Nutzen wird mittels einer Gesamtnote bestimmt; es fließen Dozenten- und Teilnehmer- bezogene Faktoren sowie Rahmenbedingungen der Veranstaltung mit ein.

Die Erhebung der Allgemeineinschätzung wurde im vorliegenden Fall im Rahmen der Post- und Retrospektionsuntersuchung durchgeführt.

²⁶⁵ Vermeintliche Erklärungen und Hinweise über die Hintergründe dieser Verteilung finden sich in der qualitativen Auswertung.

Aus den Werten der Allgemeinschätzung kann – für die Post- und Retrospektions- Erhebung analog – eine eindeutige Einschätzung abgelesen werden. Der Schwerpunkt der Werte liegt bei 2,0, eine starke Tendenz zu 1,5 und 1,0 ist offensichtlich.

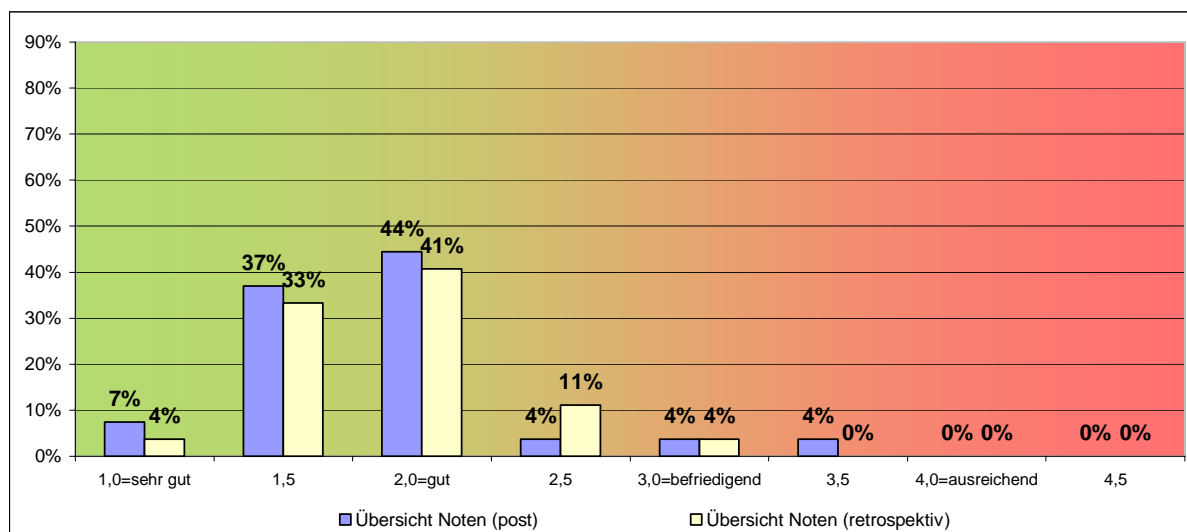


Diagramm 56: Untersuchungsergebnisse Evaluation PROFiL-TT Allgemeinschätzung

3.2.5 Qualitative Auswertung der Verbatims

In der qualitativen Auswertung der Teilnehmeranmerkungen und -anregungen tauchen mehrere Aspekte gehäuft auf. Im folgenden Teil werden diese Aspekte dargestellt und ergänzend auf die quantitative Auswertung der Erhebungen bezogen. Die ursprünglichen Anmerkungen und Kommentare der Post- und Retrospektions- Erhebungen befinden sich im Anhang (Anhang IV: Verbatims).

Die explizite Hervorhebung von einzelnen Dozenten wurde in der Aufstellung im Anhang aus Gründen des Schutzes der Privatsphäre anonymisiert.

Interessant beim Vergleich der Verbatims in der Post- Erhebung mit denen in der retrospektiven Erhebung, ist die Rekursivität einzelner Aspekte, die in der Wahrnehmung der Teilnehmer fest verankert sind.

3.2.5.1 Anmerkungen aus der Vorab- Untersuchung

Die Anmerkungen und Fragen aus der Pre- Erhebung beziehen sich größtenteils auf die Erhebung. Hierbei tauchen in erster Linie Anmerkungen zum Fragebogendesign auf. Die Dreistufigkeit – auf der die Vergleichbarkeit der einzelnen Erhebungen basiert – wird von einigen Teilnehmern als kritisch betrachtet. Generell wird von PROFiL-TT erwartet, dass die Dozenten im Rahmen dieser Weiterbildung Erfahrungen und Hilfen, sowie Anregungen für die einerseits außeruniversitäre Weiterbildung, allerdings auch – als Synergieeffekt – für ihre eigene inneruniversitäre Lehre ziehen können. Aus einigen Anmerkungen wird zudem deutlich, dass bereits eine Vorbildung im Bereich der Weiterbildung vorhanden ist. Für diese bereits vorhandenen Fähigkeiten und Kenntnisse taucht der Wunsch nach Konsolidierung, Festigung und Vertiefung auf. Die Themen Struktur (bezogen auf einen genauen Zeitplan) sowie Nachhaltigkeit wurden explizit angesprochen.

3.2.5.2 Zusammenfassung positiver Anmerkungen

Post- Erhebung

Die Anmerkungen der Teilnehmer nehmen größtenteils auf die Kategorisierung der standardisierten Teile der Erhebungsinstrumente Bezug.

Die Themen Konzeption und Zusammensetzung der Module,²⁶⁶ Praxisbezug, Feedback, Interaktivität und Dozentenengagement finden sich in den Anmerkungen häufig wieder. Ebenso häufig wird die Gruppengröße in den Veranstaltungen positiv hervorgehoben – teilweise wird dieser Aspekt auch mit Praxisbezug und Feedback in Verbindung gebracht, da eine kleinere Gruppe unter Umständen für den Praxisbezug und das „eigene Ausprobieren“ der Teilnehmer förderlich sein kann, andererseits besteht in kleineren Gruppen auch stärker die Möglichkeit für den Dozenten Teilnehmern detailliertes Feedback zu geben.

Die Interdisziplinarität der Teilnehmer – die Spannbreite geht hier von Naturwissenschaften über Wirtschaftswissenschaften bis hin zu unterschiedlichen Fachgebieten der Geisteswissenschaft – wurde von einigen Teilnehmern ebenso als angenehm und positiv eingestuft, wie die Atmosphäre und das Klima innerhalb der Gruppen.

In den Verbatims werden vorwiegend positive Anmerkungen bezogen auf die Module kohärenter Sprachgebrauch, Rhetorik und Stimmbildung und Sprecherziehung abgegeben. Die jeweiligen Dozenten, Inhalte, methodisches und didaktisches Vorgehen, sowie Einsatz und integrative Anwendung der Fachthemen werden dabei explizit hervorgehoben. Der Bezug zum eigentlichen Thema „Train-the-Trainer“ wurde hauptsächlich in den Modulen der Sprachkompetenz wieder gefunden, und in diesem Rahmen auch als wertvoll und positiv eingeschätzt.

Das Engagement einzelner Dozenten – vorwiegend der Module kohärenter Sprachgebrauch, Rhetorik und Stimmbildung / Sprecherziehung – wurde vermehrt ausführlich erwähnt.

Retrospektions- Erhebung

In der retrospektiven Erhebung tauchen ähnliche Aspekte wie in der Post- Erhebung auf. Hervorgehoben werden dabei wiederum die angenehme Gruppengröße und der daraus resultieren hohe Praxisbezug, die stimmige Kombination der Veranstaltungsinhalte und der einzelnen Module sowie das anspruchsvolle Kursniveau und das methodische Vorgehen in den einzelnen Modulen. Zudem werden die Interaktivität innerhalb der Veranstaltung und der interdisziplinäre Austausch von Kollegen vor demselben methodischen Hintergrund, allerdings unterschiedlichen Fachinhalten, als sehr nützlich betrachtet.

Die angenehme Gruppenatmosphäre wurde darüber hinaus als wertvoll und hilfreich eingeschätzt.

²⁶⁶ Fällt in die Kategorie Struktur und Organisation

Allgemein ist der Detaillierungsgrad in der retrospektiven Erhebung wesentlich geringer als in der Erhebung direkt nach der Veranstaltung – die Teilnehmer erinnern sich nach einer Zeitspanne von acht Wochen nur noch an die für sie besonders stark ausgeprägten Aspekte.²⁶⁷

3.2.5.3 Zusammenfassung konstruktiver Anmerkungen

In den konstruktiven Anmerkungen der beiden Erhebungen nach der Veranstaltung finden sich viele Aspekte wieder, die die quantitativen Daten bestätigen und unterstreichen. Generell wird das Modul Lehrkompetenz – Inhalte, Darbietung der Inhalte, Didaktik, Kursunterlagen sowie der eigentliche Themenbezug zu „Train-the-Trainer“ vor einem außeruniversitären Hintergrund –, sowie vereinzelt das Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen mit kritischen Anmerkungen versehen.

Post- Erhebung

Im Rahmen der Post- Erhebung geben die Teilnehmer teilweise sehr detaillierte konstruktive Inputs zu den Inhalten, der didaktischen Darbietung sowie den Lernunterlagen ab.

Bezogen auf das Modul Lehrkompetenz werden dabei folgende Aussagen hervorgehoben:

- Zu starke einseitig wissenschaftliche Ausprägung
- Zu intensiver Einsatz pädagogisch-psychologischer Fachterminologie
- Praxisbezug auf „Train-the-Trainer“ im außeruniversitären Bereich nicht immer vorhanden
- Etwas zu wenig Zeit für Gruppenarbeiten
- Zielsetzung der Gruppenarbeiten nicht immer klar
- Didaktische Umsetzung der Modulthemen teilweise im Widerspruch mit den eigentlichen Inhalten
- Orthographische und grammatikalische Fehler in den Kursunterlagen

Zum Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen wurde vereinzelt eine zu frontale und etwas monotone Kursgestaltung festgestellt.

²⁶⁷ Begründen kann man diese Beschränkung auf die wesentlichen Aspekte mit der Vergessenskurve nach Ebbinghaus. Zwar handelt es sich bei dieser Bewertung nicht um die Wiedergabe von Gelerntem, sondern von Einschätzungsaspekten, die Memorierleistung, wonach auf lange Sicht lediglich 15% des Gelernten, bzw. Erfahrenem im Gedächtnis verankert bleiben sollte allerdings auf die Einschätzung und Bewertung der Teilnehmer übertragen werden können. (vgl. Ebbinghaus H (1885))

Retrospektions- Erhebung

In der retrospektiven Betrachtung finden sich – ebenso wie bei der Zusammenfassung der positiven Aspekte – einige Themen aus der Einschätzung direkt nach der Veranstaltung wieder.

Im Wesentlichen werden drei Aspekte genannt:

- Zielsetzung PROFiL-TT: Ein zu starker Bezug auf universitäre Weiterbildung (statt außeruniversitär)
- Inkonsistenz in Bezug auf die Inhalte und Darbietung im Modul Einsatz von Präsentationstechniken und Gestaltung von Kursunterlagen
- Stärkerer Bezug auf aktuelle Forschungsergebnisse

Darüber hinaus wird zusätzlich eine Heterogenität der Erfahrungen und des Vorwissens der Teilnehmer angesprochen. Das Feedback diesbezüglich bezieht sich auf die Auswahl und Zusammensetzung der einzelnen Gruppen.

3.2.5.4 Kommentare und Anmerkungen zur Erhebung

Die Anmerkungen zur Vorgehensweise bei der Datenerhebung sowie zum Datenerhebungsinstrument selbst, beschränken sich auf einige zentrale Punkte, die – größtenteils – bereits vor der Erhebung bekannt waren, allerdings aus technischer, pragmatischer und konzeptioneller Sicht nicht anderweitig realisiert werden konnten.

Die dozentenunabhängige Bewertung der Veranstaltung wurde – trotz einer Erklärung dieser Vorgehensweise im Rahmen der Befragung – von einigen Teilnehmern als kritisch betrachtet. Darüber hinaus wurde die Funktion der ersten Stufe der Evaluation teilweise fehlinterpretiert – nicht als Sensibilisierung und Orientierung für die folgende Veranstaltung. Fragen und Unklarheiten zu diesem Punkt wurden jedoch bereits beim Ausfüllen des Vorab- Fragebogens am Telefon oder via Email geklärt.

3.2.6 Resümee

Im Resümee werden sowohl die Mediane, als auch die Quartilabweichungen der einzelnen Erhebungssitems für die Themenbereiche Organisation – Gesamtveranstaltung, modulübergreifende Themen, sowie alle modulbezogene Themen auf den drei Erhebungsstufen gegenübergestellt. Mittels einer Farbkodierung wird die Intensität der Quartilabweichung visualisiert – bei einer Quartilabweichung von 0 oder 1 ist der Farbcode grün, bei 2 gelb und ab einer Quartilabweichung von 3, rot). Bei der Leseart der folgenden Tabellen sollte generell beachtet werden, dass sich diese Werte auf die kumulierten Rohdaten – auf Ordinal- Niveau – beziehen und dass die Beschaffenheit der Frageskalen – ob die Skala ihren Wert maximaler Akzeptanz in der Mitte, auf dem rechten oder auf dem linken Extremwert hat – wechselt. Der Median bei Skalen mit Maximalwert in der Mitte – Wert 4 – wird im Farbcode in den Bereichen kleiner 2 und größer 6 rot, die Werte 2 und 6 gelb und der Bereich von 3 bis 5 grün

gekennzeichnet. Bezogen auf Skalen mit dem Maximalwert 7 wird der Bereich von 1 bis 3 rot, 3 bis 5 gelb und größer als 5 grün markiert.²⁶⁸

3.2.6.1 Organisation – Gesamtveranstaltung

Zu *Organisation und Gesamtveranstaltung* werden im Folgenden die Themen *Aufbau*, *Organisation* und *Betreuung außerhalb der Veranstaltung* zusammengefasst. Betrachtet man die Einschätzung und Bewertung des Aufbaus von PROFiL-TT, lässt sich ein hoher Grad der Zustimmung bezüglich Klarheit und Strukturierung der Gesamtveranstaltung erkennen. Die Deckung zwischen prospektiver und retrospektiver Sicht ist sehr deutlich zu erkennen. Dieses Resümee lässt sich auch für die Organisation ziehen. Die Teilnehmer hatten die Erwartung eine gut organisierte Veranstaltung zu besuchen, diese Erwartung konnte nach dem Kurs im Wesentlichen bestätigt werden. Die Betreuung außerhalb der Veranstaltung zeigt in allen drei Erhebungsphasen ein heterogenes Bild. Die Werte über die drei Erhebungszeiträume zeigen allerdings deutlich, dass die Erwartungshaltung der Teilnehmer bezogen auf eine Betreuung außerhalb der Veranstaltung in der prospektiven Erhebung eher gering war, sich allerdings in den Erhebungen nach der Veranstaltung dennoch bestätigt hat. Die Teilnehmer sind davon ausgegangen, außerhalb der Veranstaltung mehr oder minder auf sich selbst gestellt zu sein, schätzten die Betreuung nach der Veranstaltung allerdings zunehmend positiv ein.

	Pre: 1. Ein logischer/nachvollziehbarer inhaltlicher Aufbau der Veranstaltung / der einzelnen Module ist mir wichtig.	Pre: 2. Eine gute Organisation ist mir wichtig.	Pre 16. Erwartung: Außerhalb der Veranstaltung findet eine gute Betreuung statt.
Median Pre	6	6	4
Quartilabstand	2	2	3
	Post 1. Der inhaltliche Aufbau der Veranstaltung war logisch / nachvollziehbar.	Post 2. Die Veranstaltung war gut organisiert.	Post 16. Außerhalb der Veranstaltung fand eine gute Betreuung statt.
Median Post	6	6	6
Quartilabstand	0	1	3
	Retrospektiv 1. Der inhaltliche Aufbau der Veranstaltung war logisch / nachvollziehbar.	Retrospektiv 2. Die Veranstaltung war gut organisiert.	Retrospektiv 16. Außerhalb der Veranstaltung fand eine gute Betreuung statt.
Median Retrospektiv	6	6	7
Quartilabstand	1	2	2

Tabelle 2: Überblick Median und Quartilabstand Aufbau, Organisation und Betreuung außerhalb der Veranstaltung²⁶⁹

²⁶⁸ Alle Frageskalen der Kategorien Organisation – Gesamtveranstaltung und die modulübergreifenden Themen haben ihren Maximalwert bei 7. Die modulbezogenen Items haben ihren Wert mit maximaler Akzeptanz bei 4 (mit Ausnahme der Fragen 12.* und 20.*).

²⁶⁹ Die Skalierung für diese Items verläuft vom Minimalwert 1 (*trifft nicht zu*) bis 7 (*trifft völlig zu*).

3.2.6.2 Modulübergreifende Themen

Die *modulübergreifenden Themen* fassen all jene Items zusammen, die nicht unmittelbar auf die einzelnen Module oder die Veranstaltungsorganisation bezogen werden. Darunter fallen z.B. die Veranschaulichung von Inhalten, der generelle Praxisbezug in der Veranstaltung, Redundanzen, Interaktion und Kommunikation.

Die Veranschaulichung der Lerninhalte sowie der Theorie- Praxis- Bezug von PROFiL-TT zeigt in der Prospektion jeweils eine eindeutige Verteilung. Den Teilnehmer ist Praxisnähe und Anwendungsorientierung der Inhalte wichtig. Die Vorabeeschätzung wurde durch die Erhebungen nach der Veranstaltung größtenteils bestätigt.

Bezogen auf die inhaltlichen Überschneidungen zwischen den einzelnen Modulen weisen die prospektiven gegenüber den retrospektiven Werten eine Differenz auf. Aus den prospektiven Werten kann die Schlussfolgerung gezogen werden, dass sich einige Teilnehmer vorab nicht vollständig über die Beziehung und Verzahnung der einzelnen Module untereinander im Klaren waren. Im Nachhinein wurden die Überlagerungen zwischen den Modulen als angenehm beurteilt.²⁷⁰

Die Anregung zum aktiven Mitdenken weist in der Pre- und Post- Erhebung noch relativ hohe Quartilabweichungen auf, hat aber in allen drei Erhebungsphasen einen hohen Median. Die kritische Auseinandersetzung mit den Themen spiegelt eine ausgeprägte Sensibilisierung der Teilnehmer in den Werten wider. Das Dozenten-Feedback ist in der Nachher- Erhebung ebenso stark ausgeprägt und wird in der retrospektiven Betrachtung noch verstärkt. Aus den sehr eindeutigen Werten zu den Themen Beteiligung, kommunikative Unterrichtsformen und Einsatz kommunikativer Unterrichtsformen kann man generell schließen, dass die Veranstaltung ihre Zielsetzung – Kommunikation und Interaktion als wesentliche Handwerkzeuge von Weiterbildnern – erfüllt hat.

	Pre 3. Ich finde es wichtig, daß der Stoff anhand von Beispielen veranschaulicht wird.	Pre 4. Ein Bezug zwischen Theorie und Praxis / Anwendung ist für mich wichtig.	Pre 6. Für mich ist es wichtig, daß die Module nicht zu viele unnötige inhaltliche Überschneidungen mit den anderen Modulen aufweisen.	Pre 13. Es ist für mich wichtig, zum Mitdenken und Durchdenken des Stoffes / der Themen angeregt zu werden.	Pre 14. Ich beleuchte die behandelten Themen kritisch / von verschiedenen Seiten.
Median Pre	6	6	4	6	6
Quartilabstand	2	2	2	3	3
	Post 3. Der Stoff wurde anhand von Beispielen veranschaulicht.	Post 4. Ein Bezug zwischen Theorie und Praxis / Anwendung wurde hergestellt.	Post 6. Es traten oft unnötige inhaltliche Überschneidungen mit den anderen Modulen auf.	Post 13. Zum Mitdenken und Durchdenken des Stoffes / Themas wurde angeregt.	Post 14. Die behandelten Themen wurden kritisch / von verschiedenen Seiten beleuchtet.
Median Post	6	6	2	6	5
Quartilabstand	2	2	3	3	2
	Retrospektiv 3. Der Stoff wurde anhand von Beispielen veranschaulicht.	Retrospektiv 4. Ein Bezug zwischen Theorie und Praxis / Anwendung wurde hergestellt.	Retrospektiv 6. Es traten oft unnötige inhaltliche Überschneidungen mit den anderen Modulen auf.	Retrospektiv 13. Zum Mitdenken und Durchdenken des Stoffes / Themas wurde angeregt.	Retrospektiv 14. Die behandelten Themen wurden kritisch / von verschiedenen Seiten beleuchtet.
Median Retrospektiv	6	6	2	6	5
Quartilabstand	1	1	2	2	2

Tabelle 3a: Überblick Median und Quartilabstand modulübergreifende Themen

²⁷⁰ Bei dieser Frage ist die Antwortskala in den retrospektiven Erhebungsphasen umgepolt. Der Wert mit der höchsten Akzeptanz liegt in diesem Fall bei 1.

	Pre 15. Erwartung: Die Dozenten geben auf Beiträge der Teilnehmer hilfreiches Feedback.	Pre 17. Für mich ist es wichtig, mich mit Wortbeiträgen / bei Diskussionen zu beteiligen.	Pre 18. Erwartung: Es finden ausreichend Diskussionen statt.	Pre 19. Erwartung: Es werden kommunikative Lehrformen eingesetzt (z.B. Gruppenarbeit).
Median Pre	6	5	5	5
Quartilabstand	2	2	2	3
	Post 15. Die Dozenten gaben auf Beiträge der Teilnehmer hilfreiches Feedback.	Post 17. Möglichkeit der Beteiligung mit Wortbeiträgen / bei Diskussionen	Post 18. Ausreichende Anzahl an Diskussionen	Post 19. Einsatz kommunikativer Lehrformen
Median Post	5	6	6	7
Quartilabstand	2	1	1	1
	Retrospektiv 15. Die Dozenten gaben auf Beiträge der Teilnehmer hilfreiches Feedback.	Retrospektiv 17. Möglichkeit der Beteiligung mit Wortbeiträgen / bei Diskussionen	Retrospektiv 18. Ausreichende Anzahl an Diskussionen	Retrospektiv 19. Einsatz kommunikativer Lehrformen
Median Retrospektiv	6	7	7	7
Quartilabstand	2	1	1	1

Tabelle 3b: Überblick Median und Quartilabstand modulübergreifende Themen²⁷¹

3.2.6.3 Modulbezogene Themen

Im folgenden Abschnitt werden die modulbezogenen Themen zusammengefasst. Die Items Vorwissen, Stoffschwere, -umfang, Kurstempo, Anforderungen, Lernerwartung, subjektive Steigerung des Wissenszuwachses und die Einschätzung der dozentenbezogenen Interessantheit der Veranstaltung werden dabei für jedes Modul im Verlauf der drei Erhebungsphasen abgebildet. Die deskriptive Darstellung in den folgenden Unterkapiteln resümiert die Ergebnisse für jedes Modul.

Modul Rhetorik

	Pre 5.1 Vorwissen (Selbsteinschätzung) Modul Rhetorik	Pre 7.1 Einschätzung: Schwere des Stoffes Modul Rhetorik	Pre 8.1 Einschätzung: Stoffumfang Modul Rhetorik	Pre 9.1 Einschätzung: Kurstempo Modul Rhetorik
Median Pre	4	4	4	4
Quartilabstand	1	0	0	0
	Post 5.1 Vorwissen Modul Rhetorik	Post 7.1 Schwere des Stoffes Modul Rhetorik	Post 8.1 Stoffumfang Modul Rhetorik	Post 9.1 Kurstempo Modul Rhetorik
Median Post	4	4	4	4
Quartilabstand	0	1	1	0
	Retrospektiv 5.1 Vorwissen Modul Rhetorik	Retrospektiv 7.1 Schwere des Stoffes Modul Rhetorik	Retrospektiv 8.1 Stoffumfang Modul Rhetorik	Retrospektiv 9.1 Kurstempo Modul Rhetorik
Median Retrospektiv	5	4	4	4
Quartilabstand	1	1	1	1

Tabelle 4a: Überblick Median und Quartilabstand Modul Rhetorik²⁷²

²⁷¹ Die Skalierung für diese Items verläuft vom Minimalwert 1 (*trifft nicht zu*) bis 7 (*trifft völlig zu*).

²⁷² Skalierung: Frage 5.1: 1 (viel zu wenig um dem Kurs folgen zu können) – 4 (genau richtig) – 7 (alles mir schon bekannt, Besuch überflüssig) / Frage 7.1: 1 (viel zu leicht) – 4 (genau richtig) – 7 (viel zu schwer) / Frage

	Pre 10.1 Erwartete Anforderungen an das Modul Rhetorik	Pre 11.1 Lernerwartung Modul Rhetorik	Pre 12.1 Einschätzung: wesentlicher Wissenszuwachs Modul Rhetorik	Pre 20.1 Einschätzung: Das Modul Rhetorik wird interessant.
Median Pre	4	4	6	6
Quartilabstand	0	1	2	2
	Post 10.1 Anforderungen für das Modul Rhetorik	Post 11.1 Ich habe viel in der Veranstaltung gelernt: Modul Rhetorik	Post 12.1 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Rhetorik	Post 20.1 Das Modul Rhetorik wurde in interessanter Form gehalten.
Median Post	4	4	5	6
Quartilabstand	1	1	2	2
	Retrospektiv 10.1 Anforderungen für das Modul Rhetorik	Retrospektiv 11.1 Ich habe viel in der Veranstaltung gelernt: Modul Rhetorik	Retrospektiv 12.1 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Rhetorik	Retrospektiv 20.1 Das Modul Rhetorik wurde in interessanter Form gehalten.
Median Retrospektiv	4	4	5	6
Quartilabstand	1	0	2	1

Tabelle 4b: Überblick Median und Quartilabstand Modul Rhetorik²⁷³

Modul Stimmbildung und Sprecherziehung

	Pre 5.2 Vorwissen (Selbsteinschätzung) Modul Stimmbildung / Sprecherziehung	Pre 7.2 Einschätzung: Schwere des Stoffes Modul Stimmbildung und Sprecherziehung	Pre 8.2 Einschätzung: Stoffumfang Modul Stimmbildung und Sprecherziehung	Pre 9.2 Einschätzung: Kurstempo Modul Stimmbildung und Sprecherziehung
Median Pre	4	4	4	4
Quartilabstand	1	0	1	0
	Post 5.2 Vorwissen Modul Stimmbildung / Sprecherziehung	Post 7.2 Schwere des Stoffes Modul Stimmbildung und Sprecherziehung	Post 8.2 Stoffumfang Modul Stimmbildung und Sprecherziehung	Post 9.2 Kurstempo Modul Stimmbildung und Sprecherziehung
Median Post	4	4	4	4
Quartilabstand	1	1	1	0
	Retrospektiv 5.2 Vorwissen Modul Stimmbildung / Sprecherziehung	Retrospektiv 7.2 Schwere des Stoffes Modul Stimmbildung und Sprecherziehung	Retrospektiv 8.2 Stoffumfang Modul Stimmbildung und Sprecherziehung	Retrospektiv 9.2 Kurstempo Modul Stimmbildung und Sprecherziehung
Median Retrospektiv	5	4	4	4
Quartilabstand	1	1	1	1

Tabelle 5a: Überblick Median und Quartilabstand Modul Stimmbildung und Sprecherziehung

8.1: 1 (viel zu wenig) – 4 (genau richtig) – 7 (viel zu viel) / Frage 9.1: 1 (viel zu langsam) – 4 (genau richtig) – 7 (viel zu schnell)

²⁷³ Frage 10.1: 1 (viel zu niedrig) – 4 (genau richtig) – 7 (viel zu hoch) / Frage 11.1: 1 (nichts) – 4 (genau das, was für mich wichtig ist) – 7 (zuviel, nicht mehr zu verarbeiten) / Frage 10.1: 1 (viel zu niedrig) – 4 (genau richtig) – 7 (viel zu hoch) / Frage 12.1, 20.1: 1 (trifft nicht zu) – 7 (trifft völlig zu)

	Pre 10.2 Erwartete Anforderungen an das Modul Stimmbildung und Sprecherziehung	Pre 11.2 Lernerwartung Modul Stimmbildung und Sprecherziehung	Pre 12.2 Einschätzung: wesentlicher Wissenszuwachs Modul Stimmbildung und Sprecherziehung	Pre 20.2 Einschätzung: Das Modul Stimmbildung und Sprecherziehung wird interessant.
Median Pre	4	4	6	6
Quartilabstand	0	1	2	2
	Post 10.2 Anforderungen für das Modul Stimmbildung und Sprecherziehung	Post 11.2 Ich habe viel in der Veranstaltung gelernt: Modul Stimmbildung und Sprecherziehung	Post 12.2 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Stimmbildung und Sprecherziehung	Post 20.2 Das Modul Stimmbildung und Sprecherziehung wurde in interessanter Form gehalten.
Median Post	4	4	5	6
Quartilabstand	0	0	2	1
	Retrospektiv 10.2 Anforderungen für das Modul Stimmbildung und Sprecherziehung	Retrospektiv 11.2 Ich habe viel in der Veranstaltung gelernt: Modul Stimmbildung und Sprecherziehung	Retrospektiv 12.2 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Stimmbildung und Sprecherziehung	Retrospektiv 20.2 Das Modul Stimmbildung und Sprecherziehung wurde in interessanter Form gehalten.
Median Retrospektiv	4	4	5	7
Quartilabstand	0	0	3	1

Tabelle 5b: Überblick Median und Quartilabstand Modul Stimmbildung und Sprecherziehung²⁷⁴

Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

	Pre 5.3 Vorwissen (Selbsteinschätzung) Modul Präsentationstechniken / Gestaltung Kursunterlagen	Pre 7.3 Einschätzung: Schwere des Stoffes Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Pre 8.3 Einschätzung: Stoffumfang Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Pre 9.3 Einschätzung: Kurstempo Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen
Median Pre	4	4	4	4
Quartilabstand	2	1	0	1
	Post 5.3 Vorwissen Modul Präsentationstechniken / Gestaltung Kursunterlagen	Post 7.3 Schwere des Stoffes Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Post 8.3 Stoffumfang Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Post 9.3 Kurstempo Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen
Median Post	5	3	3	3
Quartilabstand	2	2	1	2
	Retrospektiv 5.3 Vorwissen Modul Präsentationstechniken / Gestaltung Kursunterlagen	Retrospektiv 7.3 Schwere des Stoffes Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Retrospektiv 8.3 Stoffumfang Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Retrospektiv 9.3 Kurstempo Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen
Median Retrospektiv	5	4	4	3
Quartilabstand	3	2	1	1

Tabelle 6a: Überblick Median und Quartilabstand Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen

²⁷⁴ Die Skalen sind analog zum Modul Rhetorik.

	Pre 10.3 Erwartete Anforderungen an das Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Pre 11.3 Lernerwartung Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Pre 12.3 Einschätzung: wesentlicher Wissenszuwachs Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Pre 20.3 Einschätzung: Das Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen wird interessant.
Median Pre	4	4	5	6
Quartilabstand	0	1	3	3
	Post 10.3 Anforderungen für das Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Post 11.3 Ich habe viel in der Veranstaltung gelernt: Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Post 12.3 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Post 20.3 Das Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen wurde in interessanter Form gehalten.
Median Post	3	3	4	4
Quartilabstand	2	2	2	3
	Retrospektiv 10.3 Anforderungen für das Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Retrospektiv 11.3 Ich habe viel in der Veranstaltung gelernt: Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Retrospektiv 12.3 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	Retrospektiv 20.3 Das Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen wurde in interessanter Form gehalten.
Median Retrospektiv	3	4	4	5
Quartilabstand	2	2	3	2

Tabelle 6b: Überblick Median und Quartilabstand Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen²⁷⁵

Modul kohärenter Sprachgebrauch

	Pre 5.4 Vorwissen (Selbsteinschätzung) Modul Kohärenter Sprachgebrauch	Pre 7.4 Einschätzung: Schwere des Stoffes Modul Kohärenter Sprachgebrauch	Pre 8.4 Einschätzung: Stoffumfang Modul Kohärenter Sprachgebrauch	Pre 9.4 Einschätzung: Kurstempo Modul Kohärenter Sprachgebrauch
Median Pre	4	4	4	4
Quartilabstand	1	1	0	0
	Post 5.4 Vorwissen Modul Kohärenter Sprachgebrauch	Post 7.4 Schwere des Stoffes Modul Kohärenter Sprachgebrauch	Post 8.4 Stoffumfang Modul Kohärenter Sprachgebrauch	Post 9.4 Kurstempo Modul Kohärenter Sprachgebrauch
Median Post	4	4	4	4
Quartilabstand	0	0	1	0
	Retrospektiv 5.4 Vorwissen Modul Kohärenter Sprachgebrauch	Retrospektiv 7.4 Schwere des Stoffes Modul Kohärenter Sprachgebrauch	Retrospektiv 8.4 Stoffumfang Modul Kohärenter Sprachgebrauch	Retrospektiv 9.4 Kurstempo Modul Kohärenter Sprachgebrauch
Median Retrospektiv	4	4	4	4
Quartilabstand	1	0	0	1

Tabelle 7a: Überblick Median und Quartilabstand Modul kohärenter Sprachgebrauch

²⁷⁵ Die Skalen sind analog zum Modul Rhetorik.

	Pre 10.4 Erwartete Anforderungen an das Modul Kohärenter Sprachgebrauch	Pre 11.4 Lernerwartung Modul Kohärenter Sprachgebrauch	Pre 12.4 Einschätzung: wesentlicher Wissenszuwachs Modul Kohärenter Sprachgebrauch	Pre 20.4 Einschätzung: Das Modul Kohärenter Sprachgebrauch wird interessant.
Median Pre	4	4	6	6
Quartilabstand	0	1	2	2
	Post 10.4 Anforderungen für das Modul Kohärenter Sprachgebrauch	Post 11.4 Ich habe viel in der Veranstaltung gelernt: Modul Kohärenter Sprachgebrauch	Post 12.4 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Kohärenter Sprachgebrauch	Post 20.4 Das Modul Kohärenter Sprachgebrauch wurde in interessanter Form gehalten.
Median Post	4	4	6	6
Quartilabstand	1	1	2	2
	Retrospektiv 10.4 Anforderungen für das Modul Kohärenter Sprachgebrauch	Retrospektiv 11.4 Ich habe viel in der Veranstaltung gelernt: Modul Kohärenter Sprachgebrauch	Retrospektiv 12.4 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Kohärenter Sprachgebrauch	Retrospektiv 20.4 Das Modul Kohärenter Sprachgebrauch wurde in interessanter Form gehalten.
Median Retrospektiv	4	4	6	6
Quartilabstand	0	1	3	2

Tabelle 7b: Überblick Median und Quartilabstand Modul kohärenter Sprachgebrauch²⁷⁶

Modul Lehrkompetenz

	Pre 5.5 Vorwissen (Selbsteinschätzung) Modul Lehrkompetenz	Pre 7.5 Einschätzung: Schwere des Stoffes Modul Lehrkompetenz	Pre 8.5 Einschätzung: Stoffumfang Modul Lehrkompetenz	Pre 9.5 Einschätzung: Kurstempo Kurses Modul Lehrkompetenz
Median Pre	4	4	4	4
Quartilabstand	2	0	0	0
	Post 5.5 Vorwissen Modul Lehrkompetenz	Post 7.5 Schwere des Stoffes Modul Lehrkompetenz	Post 8.5 Stoffumfang Modul Lehrkompetenz	Post 9.5 Kurstempo Kurses Modul Lehrkompetenz
Median Post	4	4	4	4
Quartilabstand	2	1	2	1
	Retrospektiv 5.5 Vorwissen Modul Lehrkompetenz	Retrospektiv 7.5 Schwere des Stoffes Modul Lehrkompetenz	Retrospektiv 8.5 Stoffumfang Modul Lehrkompetenz	Retrospektiv 9.5 Kurstempo Kurses Modul Lehrkompetenz
Median Retrospektiv	4	4	4	4
Quartilabstand	2	1	0	1

Tabelle 8a: Überblick Median und Quartilabstand Modul Lehrkompetenz

²⁷⁶ Die Skalen sind analog zum Modul Rhetorik.

	Pre 10.5 Erwartete Anforderungen an das Modul Lehrkompetenz	Pre 11.5 Lernerwartung Modul Lehrkompetenz	Pre 12.5 Einschätzung: wesentlicher Wissenszuwachs Modul Lehrkompetenz	Pre 20.5 Einschätzung: Das Modul Lehrkompetenz wird interessant.
Median Pre	4	4	5	7
Quartilabstand	0	1	2	3
	Post 10.5 Anforderungen für das Modul Lehrkompetenz	Post 11.5 Ich habe viel in der Veranstaltung gelernt: Modul Lehrkompetenz	Post 12.5 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Lehrkompetenz	Post 20.5 Das Modul Lehrkompetenz wurde in interessanter Form gehalten.
Median Post	4	4	4	4
Quartilabstand	2	2	4	3
	Retrospektiv 10.5 Anforderungen für das Modul Lehrkompetenz	Retrospektiv 11.5 Ich habe viel in der Veranstaltung gelernt: Modul Lehrkompetenz	Retrospektiv 12.5 Wesentlich höherer Wissensstand nach der Veranstaltung: Modul Lehrkompetenz	Retrospektiv 20.5 Das Modul Lehrkompetenz wurde in interessanter Form gehalten.
Median Retrospektiv	4	4	5	6
Quartilabstand	1	1	3	2

Tabelle 8b: Überblick Median und Quartilabstand Modul Lehrkompetenz²⁷⁷

²⁷⁷ Die Skalen sind analog zum Modul Rhetorik.

3.2.7 Fazit

Die Vielfalt, Komplexität und Fülle der im Rahmen der Evaluation erhobenen Daten wird in den Kapitel 3.2.4, 3.2.5 und 3.2.6 sehr detailliert beschrieben. Zieht man ein komplettes Resümee über die Einschätzungen und Beurteilungen der Teilnehmer über das Projekt PROFiL-TT, lässt sich ein sehr feinkörniges, positives und ausgewogenes Bild erkennen.

Die Erwartungshaltung der Teilnehmer an das Gesamtprojekt, die didaktische und methodische Vorgehensweise, die Rahmenbedingungen, Strukturierung und den Aufbau, die Modulinhalte und den individuellen Anspruch war dabei von vornherein sehr hoch. PROFiL-TT wurde diesen Erwartungen in großem Maßstab gerecht. Die positiven Einschätzungen zeigen sich sowohl auf Veranstaltungsebene in der Gesamtbeurteilung – berechnet man hierfür einen Notendurchschnitt, erhält man für den Wert 1,85 in der Post- Erhebung und für den Wert 1,88 in der Retrospektionserhebung –, in den Freitexteingaben der Teilnehmer (siehe Anhang IV), als auch in den Einschätzungen und Beurteilungen auf Modulebene. Der Praxisbezug und die Kombination von theoretischen Inhalten mit praktischer Umsetzung, die Zusammenstellung der Module, sowie die Kommunikation und Interaktion im Rahmen der Veranstaltung ergänzen dieses Bild. Darüber hinaus illustrieren die Rückmeldungen zur Stoffschwere, zum Kurstempo und zu den Anforderungen an die Teilnehmer die Ausgewogenheit und Angemessenheit des Veranstaltungskonzepts. In der Summe dieser Einschätzungen kann PROFiL-TT als stimmiges, harmonisch abgestimmtes und konsistentes Veranstaltungskonzept – in Theorie und praktischer Umsetzung – betrachtet werden.

Literatur- und Quellenverzeichnis

- Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung e.V. (Hrsg.). Berlin. QUEM-BULLETIN Jg. 1999 / Heft 5.
- Arnold R (1984). *Weiterbildung als Profession*. In: Zeitschrift für Weiterbildung in Rheinland Pfalz. Heft 4.
- Becker-Richter M, Habel E, Rinke B, Tegethoff HG (2002). *Evaluation von Studium und Lehre. Verfahren - Methoden - Erfahrungen*. Opladen : Leske + Budrich.
- Beywl W (1999). Abrufbar unter URL: http://www.leibi.de/takaoe/91_07.htm (Abgerufen: 20. Februar 2007).
- Beywl W (2004). Artikel *Partizipative Evaluation*. aus: Glossar wirkungsorientierte Evaluation, Univation-Institut für Evaluation Dr. Beywl & Associates GmbH, Köln 2004 online abrufbar unter URL: <http://www.univation.org/glossar/> (Abgerufen 02. März 2007).
- Beywl W, Geiter C (1997). *Evaluatoin - Controlling - Qualitätsmanagement in der Weiterbildung*. Bielefeld : Bertelsmann Verlag.
- Beywl W, Widmer T (2006). *Handbuch der Evaluationsstandards. Die Standards des "Joint Committee on Standards for Educational Evaluation" / Joint Committee on Standards for Educational Evaluation*. Sanders JR (Hrsg.). Übers. und für die dt. Ausg. erw. von Wolfgang Beywl und Thomas Widmer. Wiesbaden : VS, Verl. für Sozialwissenschaften.
- BGBI. I 931. §1 Berufsbildungsgesetz Abs. 4). Geänderte Fassung vom 23. März 2005.
- Binnendijk AL (1989). *Donor Agency Experience with the Monitoring and Evaluation of Development Projects*. In: Evaluation Review 13.
- BMAS (Hrsg.) (2006). Artikel *Der ESF in Deutschland*. In: Bundesministerium für Arbeit und Soziales. Berlin. Abrufbar unter URL: <http://www.bmas.bund.de/BMAS/Navigation/Europa-International/Europa/europaeischer-sozialfonds,did=97586.html> (Abgerufen: 26. Februar 2007).
- BMBF (Hrsg.) (2002). Artikel *Neue Formen der betrieblichen Weiterbildung*. In: Bundesministerium für Bildung und Forschung. Berlin. Abrufbar unter URL: http://www.berufsbildungsbericht.info/_htdocs/bbb1998/teil2/kap_5/teil2_5_2_3.htm (Abgerufen: 06. März 2007).
- BMBF (Hrsg.) (2006). Artikel *Lebenslanges Lernen*. In: Bundesministerium für Bildung und Forschung. Berlin. Abrufbar unter URL: <http://www.bmbf.de/de/411.php> (Abgerufen: 19. Februar 2007).
- BMBF (Hrsg.) (2006). Artikel *Professionalisierung in der Weiterbildung*. In: Bundesministerium für Bildung und Forschung. Berlin. Abrufbar unter URL: <http://www.bmbf.de/de/430.php> (Abgerufen: 8. Februar 2007).
- BMBF (Hrsg.) (2006). Artikel *Qualität in der Weiterbildung – Tests und mehr*. In: Bundesministerium für Bildung und Forschung. Berlin. Abrufbar unter URL: <http://www.bmbf.de/de/6238.php> (Abgerufen: 08. Februar 2007).
- BMBF (Hrsg.) (2006). Artikel *Weiterbildung*. In: Bundesministerium für Bildung und Forschung. Berlin. Abrufbar unter URL: <http://www.bmbf.de/de/1366.php> (Abgerufen: 16. Februar 2007).

- BMBF (Hrsg.) (2007). Artikel *Wissenschaftliche Weiterbildung*. In: Bundesministerium für Bildung und Forschung. Berlin. Abrufbar unter URL: <http://www.bmbf.de/de/349.php> (Abgerufen: 16. Februar 2007).
- BMUKK (Hrsg.) / BIFEB (Hrsg.) (2006). Artikel *Funktionsprofile*. In: Bundesministeriums für Unterricht, Kunst und Kultur / Bundesinstitut für Erwachsenenbildung. Wien. Abrufbar unter URL: <http://www.erwachsenenbildung.at/berufsfeld/berufsbild/funktionsprofile.php> (Abgerufen: 02. März 2007).
- Borg J (1981). *Anwendungsorientierte multidimensionale Skalierung*. New York : Springer.
- Brewer J, Hunter A (1989). *Multimethod Research. A Synthesis of Styles*. Newbury Park, CA : Sage.
- Campbell DT, Stanley JC (1963). *Experimental and Quasi-Experimental Designs for Research on Teaching*. In: Gage NL (Hrsg.). *Handbook of Research on Teaching*. Chicago.
- Chen HT, Rossi PH (1987). "The Theory-Driven Approach to Validity". *Evaluation and Program Planning* 10 : 95 - 103.
- Chen HT (1990). *Theory-driven evaluations*. Newbury Park : Sage.
- Crooks TJ (1988). *The impact of classroom evaluation practices on students*. *Review of Educational Research*. 58(4). 438 – 481.
- Deutsche UNESCO- Kommission e.V. (Hrsg.). Artikel *Lebenslanges Lernen in Europa*. Abrufbar unter URL: <http://www.unesco.de/451.html> (Abgerufen: 19. Februar 2007).
- Deutscher Bildungsrat (Hrsg.). *Empfehlungen der Bildungskommission. Strukturplan für das Bildungswesen* 13. Februar 1970.
- Didaktik Uni Jena (Hrsg.). Artikel *Jan Amos Komensky*. Abrufbar unter URL: http://www.didaktik.uni-jena.de/did_02/comenius.htm (Abgerufen 16. Februar 2007).
- Didaktik Uni Jena (Hrsg.). Artikel *Allgemeinbildung*. Abrufbar unter URL: http://www.didaktik.uni-jena.de/did_06/allgemein.htm (Abgerufen 16. Februar 2007).
- Dunkin MJ, Barnes J (1986). *Research on teaching in higher education*. In: Wittrock MC (Hrsg.). *Handbook of research on teaching* (3rd edition). New York : McMillan.
- Ebbinghaus H (1992). *Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie*. Neue, unveränd. und ungek. Ausgabe nach der 1. Aufl. 1885. Darmstadt : Wissenschaftliche Buchgesellschaft.
- Europäische Union (Hrsg.) (2006). Artikel *Schlüsselkompetenzen für lebenslanges Lernen*. Abrufbar unter URL: <http://europa.eu/scadplus/leg/de/cha/c11090.htm> (Abgerufen: 19. Februar 2007).
- EU-Verfassung, Artikel II-74, Ziffer 1.
- Faulstich P (2004). Artikel *Akkreditierung und Qualität wissenschaftlicher Weiterbildung*. Abrufbar unter URL: http://www.dgwf.net/docs/Akkreditierung_Qualitaet.pdf (Abgerufen: 16. Februar 2007).
- Fetterman D (1996). *Empowerment Evaluation. An introduction to theory and practice*. In: Fetterman DM, Kaffarian SJ, Wandersman A (Hrsgs.). *Empowerment Evaluation. Knowledge and tools for self-assessment & accountability*. Thousand Oaks : Sage.
- Fetterman D (2007). *Empowermentevaluation*. Abrufbar unter URL: <http://www.stanford.edu/~davidf/empowermentevaluation.html> (Abgerufen: 20. Februar 2007).

- Flick U (Hrsg.) (2006). *Qualitative Evaluationsforschung. Konzepte - Methoden - Umsetzung*. Orig.-Ausgabe. Reinbek bei Hamburg : Rowohlt.
- Grundgesetz der Bundesrepublik Deutschland, Artikel 1, §1.
- Guba EG, Lincoln YS (1981). *Effective evaluation. Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco : Jossey-Bass.
- Hager, Patry, Brezing (2000). *Zusammenfassung: Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien. Ein Handbuch*. Bern : Hans Huber. Abrufbar unter URL: http://www.fu-psychos.de/tiki-download_file.php?fileId=64 (Abgerufen: 20. Februar 2007).
- Hatwieger K (1999). Artikel *Weiterbildung der Weiterbildner. Unterstützung für eine sich wandelnde Profession*. In: Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung e.V. (Hrsg.). Berlin. QUEM-BULLETIN Jg. 1999 / Heft 5
- Haubrich K (2001). *Cluster-Evaluation – ein Modell für einen „dornigen“ Evaluationsgegenstand*; DeGEval-Jahrestagung, Block 16: „Modelle der Evaluation“; Speyer, 04.10.2001. Abrufbar unter URL: <http://www.degeval.de/calimero/tools/proxy.php?id=87> (Abgerufen: 20. Februar 2007).
- Hense J (2007). Online-Wörterbuch Evaluation. In: evolution.de - Evaluation und Qualitätssicherung im Bildungswesen. Abrufbar unter URL: <http://www.evaluation.de/glossary> (Abgerufen: 16. Januar 2007).
- Jütte W, Weber K (Hrsg.) (2005). *Kontexte wissenschaftlicher Weiterbildung. Entstehung und Dynamik von Weiterbildung im universitären Raum*. Münster ; New York ; München ; Berlin : Waxmann.
- Kempfert G, Rolff HG (2005). *Qualität und Evaluation. Ein Leitfaden für pädagogisches Qualitätsmanagement*. Weinheim ; Basel : Beltz.
- Kirkpatrick DL, Kirkpatrick JD (2006). *Evaluating Training Programs. The four levels (3rd edition)*. San Francisco, CA : Berrett-Koehler.
- Knapp G (Hrsg.) (2002). *Wissenschaftliche Weiterbildung im Aufbruch? Entwicklungen und Perspektiven*. Klagenfurt ; Ljubljana ; Wien : Hermagoras.
- Koch U, Wittmann W (Hrsg.) (1990). *Evaluationsforschung*. Berlin : Springer.
- Kraft S (2006). *Aufgaben und Tätigkeiten von Weiterbildner/inne/n – Herausforderungen und Perspektiven einer weiteren Professionalisierung in der Weiterbildung* [15.4.2006 zuletzt geändert am 02.05.2006] in: DIE- Bonn, Reports zur Weiterbildung. Abrufbar unter URL: http://www.die-bonn.de/esprid/dokumente/doc-2006/kraft06_02.pdf (Abgerufen: 12. Februar 2007).
- Kromrey H (2006). *Empirische Sozialforschung. Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung*. Stuttgart : Lucius und Lucius.
- Kuhlmann R (et.al.) (Hrsg.) (1982). *Hochschule und Weiterbildung. Bestandsaufnahme, Entwicklungen u. Zielvorstellungen aus gewerkschaftlicher Sicht*. Köln : Bund-Verlag.
- Kultusministerkonferenz (18. September 1998). *Überlegungen zur politischen Weiterbildung*. Abrufbar unter URL: <http://www.kmk.org/aktuell/miinf598.htm#polbild> (Abgerufen: 07. März 2007).
- Lecher T (1988). *Datenschutz und psychologische Forschung*. Göttingen : Hogrefe.

- Love AJ (2003). *Beyond the Black Box: Strengthening Performance Measurement through Implementation Evaluation*. Presentation to CES National Capital Chapter (November 26, 2003). Abrufbar unter URL: http://ncc.evaluationcanada.ca/e/PDFS/16_Arnold.pdf (Abgerufen: 02. März 2007).
- Lucke D (1995). *Akzeptanz. Legitimität in der „Abstimmungsgesellschaft“*. Opladen : Leske + Budrich.
- Martens H (1996). 26. Vortrag vor der Humboldt-Gesellschaft am 03. Juli 1996. *Anomie und Fundamentalismus*. Abrufbar unter URL: <http://www.humboldtgesellschaft.de/inhalt.php?name=anomie#B>. (Abgerufen: 02. März 2007).
- Müller U (2003). *Weiterbildung der Weiterbildner. Professionalisierung der beruflichen Weiterbildung durch pädagogische Qualifizierung der Mitarbeiter*. Hamburg : Kovac.
- Newman DL, Brown RD (1996). *Applied ethics for program evaluation*. Thousand Oaks : Sage.
- Nuissl E (2000). *Formazione degli Formatori. Ausbildung der Ausbilder in Deutschland*. Abrufbar unter URL: http://www.die-frankfurt.de/esprid/dokumente/doc-2000/nuissl00_02.doc Dokument aus dem Internet-Service des Deutschen Instituts für Erwachsenenbildung e. V. – <http://www.die-frankfurt.de/esprid> (Abgerufen: 12. Februar 2007).
- Olechowski R (1997). Abrufbar unter URL: <http://homepage.univie.ac.at/richard.olechowski/statements/allgemeinbildung.html> (Abgerufen: 07. März 2007).
- Opp K-D, Schmidt P (1976). *Einführung in die Mehrvariablenanalyse. Grundlagen der Formulierung und Prüfung komplexer sozialwissenschaftlicher Aussagen*. Reinbek bei Hamburg : Rowohlt.
- Owens T (1973). *Education evaluation by adversary proceeding*. In House ER (Hrsg.). *School evaluation. The politics and process*. Berkeley, CA : McCutcheon.
- Patry JL (Hrsg.) (1982). *Feldforschung*. Wien : Huber.
- Patton MQ (1997). *Utilization-focused evaluation*. (3rd edition). Thousand Oaks, CA ; London : Sage.
- Patton MQ (2002). *Qualitative Research & Evaluation Methods*. (3rd edition). Thousand Oaks, CA : Sage.
- Pehl K (2001). *Weiterbildungsstatistik – Datenlage und Perspektiven*. In: Deutsches Institut für Erwachsenenbildung (DIE). Abrufbar unter URL: http://www.die-frankfurt.de/esprid/dokumente/doc-2001/pehl01_01.pdf (Abgerufen: 06. März 2007).
- Peters R (1999). *Kompetenzen und Professionalität von Diplompädagogen mit der Studienrichtung Erwachsenenbildung/Weiterbildung*. In: Derichs-Kunstmann K, Faulstich P, Wittpoth J (Hrsg.). *Politik, Disziplin und Profession in der Erwachsenenbildung*. Beiheft zum REPORT. 92 – 102.
- Peters R (2005). *Jongleure ohne Profession*. In: *Erwachsenenbildung*. Heft 2.
- Posavac EJ, Carey RG (1997). *Program evaluation methods and case studies*. Upper Saddle River : Prentice Hall.
- PROFiL-TT. <http://www.profil-tt.lmu.de> (Abgerufen: 03. Januar 2007).
- Rindermann H (1996). *Untersuchung zur Brauchbarkeit studentischer Lehrevaluationen. Analysen der Validität und zu Auswirkungen ihres Einsatzes anhand des Heidelberger Inventars zur Lehrveranstaltungs-Evaluation (HILVE)*. Landau : Empirische Pädagogik.

- Rindermann H (1997). *Generalisierbarkeit studentischer Veranstaltungsbeurteilungen. Sind Lehrevaluationsresultate generalisierbar auf andere Veranstaltungen eines Dozenten oder auf inhaltsgleiche Parallelveranstaltungen verschiedener Dozenten?* Psychologie in Erziehung und Unterricht. 44(3). 216 – 234.
- Rindermann H (2000). *Evaluation*. In G. Wenninger (Hrsg.). Lexikon der Psychologie. Band 1.). Heidelberg : Spektrum.
- Rindermann H (2001). *Lehrevaluation. Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen ; mit einem Beitrag zur Evaluation computerbasierten Unterrichts*. Landau : Empirische Pädagogik.
- Rindermann H, Amelang M (1994). *Das Heidelberger Inventar zur Lehrveranstaltungs- Evaluation (HILVE)*. Handanweisung. Heidelberg : Ansanger.
- Rossi, Freeman, Lipsey (1999). *Evaluation. A Systematic Approach*. (6th edition). Thousand Oaks u.a. : Sage.
- Rossi, Freeman, Lipsey (2004). *Evaluation. A Systematic Approach*. (7th edition). Thousand Oaks u.a. : Sage.
- Sanders JR (1997). *Cluster evaluation*. In: Chelmsky E, Shadish WR (Hrsgs.). Evaluation for the 21st century. 396 - 404.
- Schlittgen R, Streitberg BHJ (2001). *Zeitreihenanalyse*. München ; Wien : Oldenbourg.
- Schmidt H (2002). Artikel *OCR – Was macht`s? Was bringt`s?* Berlin. Abrufbar unter URL: http://edoc.hu-berlin.de/e_rzm/23/schmidt-hannelore-2002-05-01/XML/15.xml (Abgerufen: 02. März 2007).
- Schnell R, Hill PB, Esser E (2005). *Methoden der empirischen Sozialforschung*. München ; Wien : Oldenbourg.
- Scriven M (1973). *Goal-free evaluation*. In: House ER (Hrsg.). School evaluation. The politics and process. Berkeley : McCutchan.
- Scriven M (1991). *Evaluation Thesaurus* (4th edition). Newbury Park, CA : Sage.
- Scriven M (2004). *Editorial. The Fiefdom Problem. Journal of MultiDisciplinary Evaluation* 1. Frei Abrufbar unter URL: <http://evaluation.wmich.edu/jmde/> (Abgerufen: 20. Februar 2007).
- Seel A (1997). *Handlungsorientiertes Lernen*. In Buchberger F, Eichelberger H, Klement K, Mayr J, Seel A., Teml H (Hrsg.). Seminar didaktik. Innsbruck : Studienverlag.
- Société Française de l'Évaluation (2003). *Charte de l'évaluation des politiques publiques et des programmes publiques*. Adoptée par l'assemblée générale de la SFE le 28 octobre 2003. s.l. : SFE.
- Sozialpsychologie Universität Hannover (Hrsg.). Artikel *Hawthorne Effekt*. Abrufbar unter URL: <http://www.sozpsy.uni-hannover.de/Marienthal/glossar/html/h8.htm> (Abgerufen: 06. März 2007).
- Stake RE (1975). *Program evaluation, particularly responsive evaluation*. Paper #5 in occasional paper series (Reports - Research/Technical No. ED163060). Western Michigan Univ. Kalamazoo. School of Education.
- Stake RE (1995). *The Art of Case Study Research*. Thousand Oaks, CA : Sage.
- Stamm M (1998). *Qualitätsevaluation und Bildungsmanagement im sekundären und tertiären Bildungsbereich*. Aarau : Sauerländer.

- Stockmann R (2006). *Evaluation und Qualitätsentwicklung*. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement [CEVAL, Centrum für Evaluation]. Münster ; New York ; München ; Berlin : Waxmann.
- Stufflebeam DL (1983). *The CIPP model for program evaluation*. In Madaus GF, Scriven M, Stufflebeam DL (Hrsg.). *Evaluation Models*. Boston : Kluwer-Nijhoff.
- Stufflebeam DL (1988). *The Personnel Evaluation Standards*. How to Assess Systems for Evaluating Educators. The Joint Committee on Standards for Educational Evaluation. Thousand Oaks, CA : Sage.
- Stufflebeam DL, Shinkfield AJ (1985). *Systematic evaluation*. Boston : Kluwer- Nijhoff.
- Sudman S, Bradburn NM (1982). *Asking questions*. San Francisco, CA : Jossey-Bass Publisher.
- Thompson RJ (1990). *Evaluators as Change Agents. The Case of a Foreign Assistance Project in Morocco*. In: *Evaluation and Program Planning* 13.
- Thorndike RL (1949). *Personnel selection. Test and measurement techniques*. New York : Wiley.
- Tietgens H (1988). *Professionalität für die Erwachsenenbildung*. In: Gieseke W (et. al). *Professionalität und Professionalisierung*. Bad Heilbrunn : Klinkhardt.
- Tippelt R (Hrsg.). (1999). *Handbuch Erwachsenenbildung / Weiterbildung*. 2. Auflage Opladen : Leske + Budrich.
- Tyler R (1949). *Basic Principles of Curriculum and Instruction*. Chicago : University of Chicago.
- Uhl A (1999). *Evaluation*. Abrufbar unter URL: <http://www.api.or.at/lbi/pdf/evaslex.pdf> (Abgerufen: 20. Februar 2007).
- Weinert FE (1996). *Für und Wider die „neuen Lerntheorien“ als Grundlagen pädagogisch-psychologischer Forschung*. *Zeitschrift für Pädagogische Psychologie*, 10(1).
- Weißeno G (2005). *Bildungsreformen. Standards für die politische Bildung*. Aus *Politik und Zeitgeschichte* (APuZ 12/2005). Abrufbar unter URL: http://www.bpb.de/publikationen/XXMG1M,0,Standards_f%FCr_die_politische_Bildung.html (Abgerufen 03.März 2007).
- Wikipedia. Artikel *Allgemeinbildung*. In: Wikipedia, Die freie Enzyklopädie. Bearbeitungsstand: 26. Februar 2007, 21:52 UTC. URL: <http://de.wikipedia.org/w/index.php?title=Allgemeinbildung&oldid=28389726> (Abgerufen: 27. Februar 2007, 10:20 UTC).
- Wikipedia. Artikel *Berufliche Weiterbildung*. In: Wikipedia, Die freie Enzyklopädie. Bearbeitungsstand: 9. Februar 2007, 16:41 UTC. URL: http://de.wikipedia.org/w/index.php?title=Berufliche_Weiterbildung&oldid=27594355 (Abgerufen: 27. Februar 2007, 10:09 UTC).
- Wikipedia. Artikel *Lebenslanges Lernen*. In: Wikipedia, Die freie Enzyklopädie. Bearbeitungsstand: 12. Dezember 2006, 17:00 UTC. URL: http://de.wikipedia.org/w/index.php?title=Lebenslanges_Lernen&oldid=25036658 (Abgerufen: 27. Februar 2007, 10:02 UTC).

- Wikipedia. Artikel *Politische Bildung*. In: Wikipedia, Die freie Enzyklopädie. Bearbeitungsstand: 19. Februar 2007, 22:46 UTC. URL:
http://de.wikipedia.org/w/index.php?title=Politische_Bildung&oldid=28061410 (Abgerufen: 27. Februar 2007, 10:22 UTC).
- Williams DD (Hrsg.). (1986). *Editor's notes. Naturalistic evaluation*. Vol. 30. New directions in program evaluation. San Francisco : Jossey-Bass.
- Wolf RL (1979). *The use of judicial evaluation methods in the formulation of educational policy*. *Educational Evaluation and Policy Analysis*. Vol. 1. No. 3 (May - Jun., 1979). 19 – 28.
- Worthen BR, Sanders JR (1973). *Educational evaluation. Theory and practice*. Belmont, CA : Wadsworth.
- Worthen BR, Sanders JR, Fitzpatrick JL (1997). *Program Evaluation. Alternative Approaches and Practical Guidelines*. White Plains, NY : Longman Inc.
- Wottawa H, Thierau H (2003). *Lehrbuch Evaluation*. Bern; Göttingen; Toronto; Seattle : Huber.

Anhang

Anhang I: Program Evaluation Standards des Joint Committee on Standards for Educational Evaluation²⁷⁸

Nützlichkeitsstandards

Die Nützlichkeitsstandards sollen sicherstellen, dass die Evaluation sich an den geklärten Evaluationszwecken sowie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausrichtet.

N1 Ermittlung der Beteiligten und Betroffenen

Die an einer Evaluation beteiligten oder von ihr betroffenen Personen sollten identifiziert werden, damit deren Interessen und Bedürfnisse berücksichtigt werden können.

N2 Glaubwürdigkeit der Evaluatorenin

Wer Evaluationen durchführt, sollte sowohl vertrauenswürdig als auch kompetent sein, damit bei den Evaluationsergebnissen ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.

N3 Umfang und Auswahl der Informationen

Die gewonnenen Informationen sollten von einem Umfang und einer Auswahl sein, welche die Behandlung sachdienlicher Fragen zum Programm ermöglichen und gleichzeitig auf die Interessen und Bedürfnisse des Auftraggebers und anderer Beteiligter und Betroffener eingehen.

N4 Feststellung von Werten

Die Perspektiven, Verfahren und Gedankengänge, auf denen die Interpretationen der Ergebnisse beruhen, sollten sorgfältig beschrieben werden, damit die Grundlagen der Werturteile klar ersichtlich sind.

N5 Klarheit des Berichts

Evaluationsberichte sollten das evaluierte Programm einschließlich seines Kontextes ebenso beschreiben wie die Ziele, die Verfahren und Befunde der Evaluation, damit die wesentlichen Informationen zur Verfügung stehen und leicht verstanden werden können.

N6 Rechtzeitigkeit und Verbreitung des Berichts

Wichtige Zwischenergebnisse als auch Schlußberichte sollten den vorgesehenen Nutzern so zur Kenntnis gebracht werden, daß diese sie rechtzeitig verwenden können.

N7 Wirkung der Evaluation

Evaluationen sollten so geplant, durchgeführt und dargestellt werden, daß die Beteiligten & Betroffenen dazu ermuntert werden, dem Evaluationsprozeß zu folgen, damit die Wahrscheinlichkeit steigt, daß die Evaluation genutzt wird.

Durchführbarkeitsstandards

Die Durchführbarkeitsstandards sollen sicherstellen, daß eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewußt ausgeführt wird.

D1 Praktische Verfahren

Die Evaluationsverfahren sollten praktisch sein, so daß Störungen minimiert und die benötigten Informationen beschafft werden können.

D2 Politische Tragfähigkeit

Evaluationen sollten mit Voraussicht auf die unterschiedlichen Positionen der verschiedenen Interessengruppen geplant und durchgeführt werden, um deren Kooperation zu erreichen und um mögliche Versuche irgendeiner dieser Gruppen zu vermeiden, die Evaluationsaktivitäten einzuschränken oder die Ergebnisse zu verzerren respektive zu mißbrauchen.

D3 Kostenwirksamkeit

Die Evaluation sollte effizient sein und Informationen mit einem Wert hervorbringen, der die eingesetzten Mittel rechtfertigt.

Korrekteitsstandards

Die Korrektheitsstandards sollen sicherstellen, daß eine Evaluation rechtlich und ethisch korrekt durchgeführt wird und dem Wohlergehen der in die Evaluation einbezogenen und auch der durch die Ergebnisse betroffenen Personen gebührende Aufmerksamkeit widmet.

K1 Unterstützung der Dienstleistungsorientierung (entfällt bei der DeGEval)

Die Evaluation sollte so geplant werden, daß Organisationen dabei unterstützt werden, die Interessen und Bedürfnisse des ganzen Zielgruppenspektrums zu berücksichtigen und ihre Tätigkeiten danach auszurichten.

K2 Formale Vereinbarungen (=F1)

Die Pflichten der Vertragsparteien einer Evaluation (was, wie, von wem, wann getan werden soll) sollten schriftlich festgehalten werden, damit die Parteien verpflichtet sind, alle Bedingungen dieser Vereinbarung zu erfüllen oder aber diese erneut zum Gegenstand von formalen Verhandlungen zu machen.

K3 Schutz individueller Menschenrechte (=F2)

²⁷⁸ Beywl W, Widmer T (2006)

Evaluationen sollten so geplant und durchgeführt werden, daß die Rechte und das Wohlergehen der Menschen respektiert und geschützt sind.

K4 Human gestaltete Interaktion (=F2)

Evaluatorinnen sollten in ihren Kontakten mit Anderen die Würde und den Wert der Menschen respektieren, damit diese nicht gefährdet oder geschädigt werden.

K5 Vollständige und faire Einschätzung (=F3)

Evaluationen sollten in der Überprüfung und in der Präsentation der Stärken und Schwächen des evaluierten Programms vollständig und fair sein, so daß die Stärken weiter ausgebaut und Problemfelder angesprochen werden können.

K6 Offenlegung der Ergebnisse (=F5)

Die Vertragsparteien einer Evaluation sollten sicherstellen, daß die Evaluationsergebnisse – einschließlich ihrer relevanten Beschränkungen – den durch die Evaluation betroffenen Personen ebenso wie all jenen, die einen ausgewiesenen Anspruch auf die Evaluationsergebnisse haben, zugänglich gemacht werden.

K7 Deklaration von Interessenskonflikten (in F4)

Interessenkonflikte sollten offen und aufrichtig behandelt werden, damit sie die Evaluationsverfahren und -ergebnisse nicht beeinträchtigen.

K8 Finanzielle Verantwortlichkeit (in D3)

Die Zuweisung und Ausgabe von Ressourcen durch [den Evaluator] sollte durch eine sorgfältige Rechnungsführung nachgewiesen werden und auch anderweitig klug sowie ethisch verantwortlich erfolgen, damit die Ausgaben verantwortungsbewusst und angemessen sind.

Genauigkeitsstandards

Die Genauigkeitsstandards sollen sicherstellen, daß eine Evaluation über die Güte und/oder die Verwendbarkeit des evaluierten Programms fachlich angemessene Informationen hervorbringt und vermittelt.

G1 Programmdokumentation

Das zu evaluierende Programm sollte klar und genau beschrieben und dokumentiert werden, so daß es eindeutig identifiziert werden kann.

G2 Kontextanalyse

Der Kontext, in dem das Programm angesiedelt ist, sollte ausreichend detailliert untersucht werden, damit mögliche Beeinflussungen des Programms identifiziert werden können.

G3 Beschreibung von Zielen und Vorgehen

Die Zwecke und das Vorgehen der Evaluation sollten ausreichend genau dokumentiert und beschrieben werden, so daß sie identifiziert und eingeschätzt werden können.

G4 Verlässliche Informationsquellen

Die in einer Programmevaluation genutzten Informationsquellen sollten hinreichend genau beschrieben sein, damit die Angemessenheit der Informationen eingeschätzt werden kann.

G5 Valide Informationen

Die Verfahren zur Informationsgewinnung sollten so gewählt oder entwickelt und dann umgesetzt werden, daß die Gültigkeit der gewonnenen Interpretationen für den gegebenen Zweck sichergestellt ist.

G6 Reliable Informationen

Die Verfahren zur Informationsgewinnung sollten so gewählt oder entwickelt und dann umgesetzt werden, daß die Zuverlässigkeit der gewonnenen Interpretationen für den gegebenen Zweck sichergestellt ist.

G7 Systematische Informationsüberprüfung

Die in einer Evaluation gesammelten, aufbereiteten und präsentierten Informationen sollten systematisch überprüft und alle gefundenen Fehler sollten korrigiert werden.

G8 Analyse quantitativer Informationen

Quantitative Informationen einer Evaluation sollten angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluation effektiv beantwortet werden.

G9 Analyse qualitativer Informationen

Qualitative Informationen einer Evaluation sollten angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluation effektiv beantwortet werden.

G10 Begründete Schlussfolgerungen

Die in einer Evaluation gezogenen Folgerungen sollten ausdrücklich begründet werden, damit die Beteiligten & Betroffenen diese einschätzen können.

G11 Unparteiische Berichterstattung

Die Verfahren der Berichterstattung sollten mittels Vorkehrungen gegen Verzerrungen durch persönliche Gefühle und Vorlieben irgendeiner Evaluationspartei geschützt werden, so daß Evaluationsberichte die Ergebnisse fair wiedergeben.

G12 Meta-Evaluation

Die Evaluation selbst sollte formativ und summativ in bezug auf die vorliegenden oder andere wichtige Standards evaluiert werden, so daß die Durchführung entsprechend angeleitet werden kann und damit die Beteiligten & Betroffenen bei Abschluß einer Evaluation deren Stärken und Schwächen gründlich überprüfen können.

Anhang II: Heidelberger Inventar zur Lehrveranstaltungsevaluation HILVE-II Dimensionen und Items²⁷⁹

Struktur

„Der inhaltliche Aufbau der Veranstaltung ist logisch / nachvollziehbar.“

„Die Veranstaltung war gut organisiert.“

Auseinandersetzung

„Der Stoff wird anhand von Beispielen veranschaulicht.“

„Die Bedeutung / Nutzen der behandelten Themen wird vermittelt.“

„Ein Bezug zwischen Theorie und Praxis / Anwendung wird hergestellt.“

Verarbeitung

„Zum Mitdenken und Durchdenken des Stoffes / Themas wird angeregt.“

„Die behandelten Themen werden kritisch / von verschiedenen Seiten beleuchtet.“

Lehrkompetenz

„Die Dozentin / Der Dozent spricht verständlich und anregend.“

„Die Dozentin / Der Dozent kann Kompliziertes verständlich machen.“

„Die Dozentin / Der Dozent fasst regelmäßig den Stoff zusammen.“

„Die Dozentin / Der Dozent wirkt gut vorbereitet.“

Dozentenengagement

„Die Dozentin / Der Dozent engagiert sich bei der Lehrtätigkeit und versucht Begeisterung zu vermitteln.“

„Der Dozentin / Dem Dozent ist es wichtig, dass die Teilnehmer etwas lernen.“

„Die Dozentin / Der Dozent motiviert die Teilnehmer.“

Klima

„Der/Die Dozent/in ist im Umgang mit den Studierenden freundlich.“

„Die Dozentin / Der Dozent ist kooperativ und aufgeschlossen.“

²⁷⁹ Vgl. Rindermann H (2001). Lehrevaluation. Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen ; mit einem Beitrag zur Evaluation computerbasierten Unterrichts / Heiner Rindermann. – Landau. Empirische Pädagogik. 384 – 387

Interessantheit

„Die Veranstaltung wird in interessanter Form gehalten.“

„Die Veranstaltung zieht sich schleppend dahin.“

Thema

„Ich habe mich schon vor dem Kurs sehr für die Themen interessiert.“

„Das Thema des Kurses als solches ist relevant (Beruf / Praxis[...]).“

Redundanz

„Mein Vorwissen:

1=zu wenig um Kurs folgen zu können,

4=genau richtig

7=alles mir schon bekannt gewesen, Besuch überflüssig“

„Es treten oft unnötige inhaltliche Überschneidungen mit anderen Kursen auf.“

Anforderungen

„Schwere des Stoffes als solches: viel zu leicht=1/genau richtig=4/viel zu schwer=7“.

„Umfang des Stoffes: viel zu wenig=1/genau richtig=4/viel zu viel=7.“

„Das Tempo des Kurses: viel zu langsam=1/genau richtig=4/ viel zu schnell=7.“

„Die Anforderungen sind viel zu niedrig=1/ genau richtig=4/ viel zu hoch=7.“

Lernen- quantitativ und qualitativ

Lernen- quantitativ

„Ich lerne viel in der Veranstaltung.“

„Mein Wissensstand ist nach der Veranstaltung wesentlich höher als vorher.“

Lernen- qualitativ

„Ich verfüge über ein grundlegendes Verständnis als vor dem Kurs.“

„Ich lerne etwas Sinnvolles und Wichtiges.“

Betreuung

„Die Lehrkraft gibt auf Beiträge der Teilnehmer hilfreiches Feedback.“

„Außerhalb der Veranstaltung findet eine gute Betreuung statt.“

Referate

(„Die Lehrkraft ergänzt Referate viel zu wenig=1/ genau richtig=4/ viel zu viel=7;“)

„Die Vortragsweise und Darbietung der Inhalte ist ansprechend / gut.“

„Die fachlich-inhaltliche Qualität der Inhalte ist hoch.“

„Ich lerne viel durch die Referate anderer Teilnehmer.“

Fleiß/ Arbeitshaltung

„Ich bereite den Kurs (z.B. Texte lesen) vor oder nach (1=sehr wenig / 7= sehr viel).“

„Mein Arbeitsaufwand ist verglichen mit anderen Veranstaltungen hoch.“

„Mein üblicher Arbeitsaufwand für den Kurs pro Woche (nicht Kursdauer): Stunden, Minuten“

Interaktionsmöglichkeiten

„Die Dozentin/Der Dozent fördert Fragen und aktive Mitarbeit.“

„Diskussionen werden gut geleitet (Anregungen von Beiträgen, Eingehen auf Beiträge, Zeiteinteilung, Bremsen von Vielrednern).“

Beteiligung

„Ich beteilige mich mit Wortbeiträgen/bei Diskussionen (1=nie/4=manchmal(7=oft)“

„Beim Einbringen eigener Beiträge fühle ich mich frei und äusserungsfähig. (Falls nie Beiträge bitte frei lassen.)“

Kommunikativen Unterrichtsformen

„Es finden ausreichend Diskussionen statt.“

„Es werden kommunikative Lehrformen eingesetzt (z.B. Gruppenarbeit).“

Anomie

„Unruhe, Reden oder Störungen durch Teilnehmer beeinträchtigen den Kurs.“

„Ich habe an mehreren Sitzungen gefehlt (0=keinmal/1=1x/2=2x/3=3x, usw.). Gründe (in Worten):“

Interessenförderung

„Die Veranstaltung fördert mein Interesse am Studium.“

„Der Kurs motiviert dazu, sich selbst mit den Inhalten zu beschäftigen.“

Allgemeineinschätzung

„Der Besuch der Veranstaltung lohnt sich.“

„Wenn man alles in eine Note zusammenfassen könnte, würde ich der Veranstaltung die folgende Note geben (Notenskala von sehr gut=1,0 bis ungenügend=6,0 inkl. Zwischennoten wie 2,5 oder 2- und 2+).“

Anhang III: Datenerhebungsinstrumente – Screenshots - Beispiele

Evaluationsbogen Profil TT -1- - Mozilla Firefox

http://www.profil-ht.lmu.de/evaluation/eval_vorher.php

EVALUATIONSBÖGEN PROFIL TT

Teilnahmetermin Lehrkompetenz:

Teilnahmetermin Sprachkompetenz:

Rahmenbedingungen (mehrere Antworten sind möglich, allerdings nicht notwendig)

Auf welchem Wege haben Sie von Profil TT erfahren?

Sind Sie momentan als Weiterbildner tätig?

In welchem Rahmen haben Sie vor, im Anschluß an das Seminar als Weiterbildner tätig zu sein?

Im folgenden Teil des Fragebogens werden Ihnen Fragen zu den einzelnen Modulen gestellt. Hierbei ist es für die Untersuchung sehr wichtig, daß Sie sich auf die Inhalte der einzelnen Module konzentrieren. Der Eindruck, den die Dozentin / der Dozent / die Dozenten bei Ihnen erweckt haben, sollte dabei weniger im Fokus stehen, da die Module dozentenunabhängig entwickelt wurden und auch als solche evaluiert werden. Eine separate Frage, die sich mit der Dozententätigkeit beschäftigt, finden Sie am Ende.

Antwortskala: "1 = trifft nicht zu" bis "trifft völlig zu = 7", "k.A." für "Keine Angabe"

Struktur	1	2	3	4	5	6	7	k.A.
1. Ein logischer/ nachvollziehbarer inhaltlicher Aufbau der Veranstaltung / der einzelnen Module ist mir wichtig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Eine gute Organisation ist mir wichtig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Auseinandersetzung	1	2	3	4	5	6	7	k.A.
3. Ich finde es wichtig, daß der Stoff anhand von Beispielen veranschaulicht wird.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Ein Bezug zwischen Theorie und Praxis / Anwendung ist für mich wichtig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Redundanz	1	2	3	4	5	6	7	k.A.
5. Mein Vorwissen (geschätzt): 1=zu wenig um Kurs folgen zu können, 4=genau richtig, 7=alles mir schon bekannt gewesen, Besuch überflüssig, "k.A." für "Keine Angabe"	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Rhetorik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Stimmbildung und Sprecherziehung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Kohärenter Sprachgebrauch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Lehrkompetenz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Für mich ist es wichtig, daß die Module nicht zu viele unnötige inhaltliche Überschneidungen mit den anderen Modulen aufweisen. Antwortskala: 1 = trifft nicht zu bis trifft völlig zu = 7, k.A. für Keine Angabe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anforderungen	1	2	3	4	5	6	7	k.A.
7. Erwartung: Schwere des Stoffes als solches: viel zu leicht=1/genau richtig=4/viel zu schwer=7, k.A. für Keine Angabe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fertig

Abbildung 5: Auszug Instrument Prospektive Erhebung

Evaluationsbogen Profil TT -2- - Mozilla Firefox

http://www.profil-tt.lmu.de/evaluation/eval_nachher.php

EVALUATIONSBOGEN PROFIL TT

Teilnahmeterrn Lehrkompetenz

Teilnahmeterrn Sprachkompetenz

Rahmenbedingungen (mehrere Antworten sind möglich, allerdings nicht notwendig)

Auf welchem Wege haben Sie von Profil TT erfahren?

Sind Sie momentan als Weiterbildner tätig?

In welchem Rahmen haben Sie vor, im Anschluß an das Seminar als Weiterbildner tätig zu sein?

Im folgenden Teil des Fragebogens werden Ihnen Fragen zu den einzelnen Modulen gestellt. Hierbei ist es für die Untersuchung sehr wichtig, daß Sie sich auf die Inhalte der einzelnen Module konzentrieren. Der Eindruck, den die Dozentin / der Dozent / die Dozenten bei Ihnen erweckt haben, sollte dabei weniger im Fokus stehen, da die Module dozentenunabhängig entwickelt wurden und auch als solche evaluiert werden. Eine separate Frage, die sich mit der Dozententätigkeit beschäftigt, finden Sie am Ende.

Antwortskala: "1 = trifft nicht zu" bis "trifft völlig zu = 7"; "k.A." bedeutet "keine Angabe"

Struktur	1	2	3	4	5	6	7	k.A.
1. Der inhaltliche Aufbau der Veranstaltung war logisch / nachvollziehbar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Die Veranstaltung war gut organisiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Auseinandersetzung	1	2	3	4	5	6	7	k.A.
3. Der Stoff wurde anhand von Beispielen veranschaulicht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Ein Bezug zwischen Theorie und Praxis / Anwendung wurde hergestellt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Redundanz	1	2	3	4	5	6	7	k.A.
5. Mein Vorwissen war: 1=zu wenig um Kurs folgen zu können, 4=genau richtig, 7=alles mir schon bekannt gewesen, Besuch überflüssig; "k.A." bedeutet "keine Angabe"	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Rhetorik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Stimm- und Sprechbildung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Kohärenter Sprachgebrauch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Lehrkompetenz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Es traten oft unnötige inhaltliche Überschneidungen mit den anderen Modulen auf. Antwortskala: 1 = trifft nicht zu bis trifft völlig zu = 7, k.A. für keine Angabe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anforderungen	1	2	3	4	5	6	7	k.A.
7. Schwere des Stoffes: viel zu leicht=1/genau richtig=4/viel zu schwer=7; "k.A." bedeutet "keine Angabe"	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Rhetorik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fertig

Abbildung 6: Auszug Instrument Post- Erhebung

Evaluationsbogen Profil TT -3- - Mozilla Firefox

http://www.profil-tt.lmu.de/evaluation/eval_retrospekt.php

EVALUATIONSBOGEN PROFIL TT

Teilnahmeterrn Lehrkompetenz

Teilnahmeterrn Sprachkompetenz

Rahmenbedingungen (mehrere Antworten sind möglich, allerdings nicht notwendig)

Auf welchem Wege haben Sie von Profil TT erfahren?

Sind Sie momentan als Weiterbildner tätig?

In welchem Rahmen haben Sie vor, im Anschluß an das Seminar als Weiterbildner tätig zu sein?

Im folgenden Teil des Fragebogens werden Ihnen Fragen zu den einzelnen Modulen gestellt. Hierbei ist es für die Untersuchung sehr wichtig, daß Sie sich auf die Inhalte der einzelnen Module konzentrieren. Der Eindruck, den die Dozentin / der Dozent / die Dozenten bei Ihnen erweckt haben, sollte dabei weniger im Fokus stehen, da die Module dozentenunabhängig entwickelt wurden und auch als solche evaluiert werden. Eine separate Frage, die sich mit der Dozententätigkeit beschäftigt, finden Sie am Ende.

Antwortskala: "1 = trifft nicht zu" bis "trifft völlig zu = 7"; "k.A." bedeutet "keine Angabe"

Struktur	1	2	3	4	5	6	7	k.A.
1. Der inhaltliche Aufbau der Veranstaltung war logisch / nachvollziehbar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Die Veranstaltung war gut organisiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Auseinandersetzung	1	2	3	4	5	6	7	k.A.
3. Der Stoff wurde anhand von Beispielen veranschaulicht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Ein Bezug zwischen Theorie und Praxis / Anwendung wurde hergestellt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Redundanz	1	2	3	4	5	6	7	k.A.
5. Mein Vorwissen war: 1=zu wenig um Kurs folgen zu können, 4=genau richtig, 7=alles mir schon bekannt gewesen, Besuch überflüssig; "k.A." bedeutet "keine Angabe"	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Rhetorik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Stimm- und Sprechbildung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Einsatz von Präsentationstechniken / Gestaltung von Kursunterlagen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Kohärenter Sprachgebrauch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Lehrkompetenz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Es traten oft unnötige inhaltliche Überschneidungen mit den anderen Modulen auf. Antwortskala: 1 = trifft nicht zu bis trifft völlig zu = 7, k.A. für keine Angabe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anforderungen	1	2	3	4	5	6	7	k.A.
7. Schwere des Stoffes: viel zu leicht=1/genau richtig=4/viel zu schwer=7; "k.A." bedeutet "keine Angabe"	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modul Rhetorik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fertig

Abbildung 7: Auszug Instrument Retrospektive Erhebung

Anhang IV: Verbatims

„Was ist besonders gut an der Veranstaltung?“

„Was ist schlecht?“

„Verbesserungsvorschläge: Kommentar zum Fragebogen / Erhebung“

Verbatims Post- Erhebung

Positive Aspekte
Tag 3 bis 5
Modul Sprecherziehung
Die Gelegenheit, den vermittelten Stoff regelmäßig mit Feedback direkt ausprobieren zu können.
Die einzelnen Module haben gut zusammengepasst. Es gab viele praktische Übungen. Die Inhalte konnten auf die verschiedenen Fachrichtungen, die die Teilnehmer vertraten, angewendet werden.
Die Kombination der Module hat mir gut gefallen. Die Übungen allein und in der Gruppe waren aktivierend und trugen dazu bei, sich das Gehörte anzueignen. Von Vorteil war hier auch die Gruppengröße.
Forschungsansatz ziele -erwartungen besprechen kompetenz gratis
Mir hat das Modul kohärenter Sprachgebrauch besonders gut gefallen. Er war methodisch und inhaltlich stimmig und anregend aufgebaut und vermittelte mir Dinge, von denen ich glaube, dass ich Sie in einem normalen \\\\\"Train the Trainer\\\\\\"-Seminar nicht zur Sprache gekommen wären.
Die Gruppengröße und die interdisziplinäre Zusammensetzung der Gruppe. Dass es verschiedene Dozenten waren. Die Zusammensetzung der Module.
Die Größe der Gruppe und damit die Möglichkeit, viel selber zu üben. Die Zusammenstellung der Module und Lerninhalte.
- Facettenreichtum der Kursinhalte - Balance zwischen Theorie und Praxis - kleine Gruppengröße, so dass viele Ideen vorgestellt werden konnten und jeder Teilnehmer / jede Teilnehmergruppe individuelles Feedback bekam - Bezug zum Thema Weiterbildung für Weiterbildung bes. in den Kursen außer \\\\"Lehrkompetenz\\\\" - Hilfestellungen bei der Bearbeitung von Aufgaben bes. beim \\\\"Kohärenten Sprachgebrauch\\\\" - detailliertes und hilfreiches Feedback in \\\\"Rhetorik\\\\" und \\\\"Kohärenter Sprachgebrauch\\\\"
Die Module Rhetorik und Stimmbildung waren angenehm und lehrreich, was auch dadurch ermöglicht wurde, dass innerhalb der Gruppe eine Atmosphäre gegenseitigen Vertrauens herrschte, die notwendig ist, sich offen und kritisch mit eigenen und fremden Beiträgen auseinanderzusetzen. In diesem Zusammenhang: Es müsste meiner Meinung nach noch mehr gruppendynamisch gearbeitet werden - es schien mit zuwenig (obwohl es in diesem Falle nicht unbedingt notwendig war).
Rhetorik Training Videoaufzeichnungen Größe der Gruppe
Die Möglichkeit zur Reflexion der eigenen Fähigkeiten bei eigenen Vorträgen, Seminaren etc. kombiniert mit praktischen Übungen und der Vermittlung der theoretischen Grundlagen von Elementen von Vorträgen, Seminaren und Kursen (z.B. \\\\"Rekurrenz, vertikale Integration, Unvollständigkeit\\\\") ist sehr wertvoll. Schon der Kurs \\\\"Profil\\\\" (ohne TT) war fuer die Verbesserung meines Vortragsstils erfolgreich, Profil-TT kann das Niveau von Profil halten. (Das Modul Stimmbildung/ Sprecherziehung war bei beiden enthalten und sehr aehnlich.)
- der interdisziplinäre Charakter durch die Teilnehmer - die kleine Gruppe - die vielen Übungen - wirklich gute Referenten
Die inhaltliche Zusammensetzung Die Vorbereitung der Dozenten Die Schulungsunterlagen kleine Gruppe viel Praktisches, Übungen
- übersichtliche Teilnehmerzahl - inetsives Eingehen auf Fragestellungen der Teilnehmer - arbeiten mit videoaufzeichnungen
Gute Strukturierung, gute Veranschaulichung durch Beispiele
-Wirklich sehr gut waren die Übungen im Modul Rhetorik -großartig: das gesamte Gruppenklima; IMMER WIEDER!!
Gruppenarbeiten, Diskussionen, angenehme Atmosphäre in der Gruppe
nette Atmosphäre

Die Organisation war perfekt. Die Gruppengröße ist äußerst angenehm. Die Idee des Programms an sich ist wundervoll.
Die anderen Teilnehmer des Kurses waren sehr nett. Die Woche war insgesamt ein großer Erfolg. Nochmals herzlichen Dank!

Konstruktive Anmerkungen
zu wenig Praxisbeispiele, zu viele eigenständige Arbeit, nur einige psychologische Faktoren bearbeitet, wichtige weggelassen.
Gruppenarbeit teilweise etwas langatmig. Zu häufige Präsentationen von Kursteilnehmern - Lerneffekt für Zuhörer und Präsentatoren nicht immer gegeben.
Zuspätkommen / Nichterscheinen der Teilnehmer
Das Modul Lehrkompetenz war zu sehr von Bezugnahme auf die den Konzepten zu Grunde liegenden wissenschaftlichen Arbeiten geprägt. Weniger Fachjargon (z.B. "Aufgaben, bei denen es auf Speed ankommt" statt "Aufgaben, bei denen es auf Geschwindigkeit ankommt") täte dem an sich informativen Modul sehr gut.
Im Teil Lehrkompetenz wurde gelegentlich etwas zu sehr in der Fachterminologie geschwelgt, was für unsere (praktischen) Zwecke nicht immer passte. Mit den Begrifflichkeiten hatten wohl vor allem diejenigen Teilnehmer etwas Schwierigkeiten, die nicht aus dem pädagogisch-psychologischen Bereich stammen. Außerdem wäre es in diesem Teil angenehm gewesen, für die Gruppenarbeiten etwas mehr Zeit zu haben. Es wäre der Qualität der einzelnen Präsentationen zu Gute gekommen und damit für die Beteiligten noch befriedigender gewesen.
M. E. wurde insgesamt zu wenig auf das Thema Weiterbildung eingegangen.
1. ausgewählte Texte passten nicht ganz zu dem Ziel 2. unterschiedliche Niveaus der Teilnehmer: einige konnten mit den Termini, die aus der Forschung kommen, nichts anzufangen 3. Planung: abweichendes Tempo (langsamer) der Gruppe schien nicht geplant/gewünscht zu sein: man kann nicht erwarten dass die Planung für alle Gruppen passt... trotzdem vielen Dank!!
Der Lehrkompetenz-Teil hatte nur wenig mit der Problemstellung (Train the trainer) zu tun.
Im Modul Lehrkompetenz war der Bezug zur Weiterbildung außerhalb der Universität nicht bis gar nicht ersichtlich. Die Dozenten in diesem Modul haben ihre eigenen Grundsätze oft nicht eingehalten, was ziemlich demotivierend wirkt. Der Input in diesem Modul war definitiv zu niedrig und sehr einseitig. Die anderen Module haben mir aber gut gefallen (inhaltlich, didaktisch und was die Dozenten betrifft).
Die orthographisch und grammatikalisch äußerst fehlerhaften Kursunterlagen für das Lehrkompetenz-Modul (aber auch in den anderen Teilen recht viele Trennungsfehler). Im Lehrkompetenz-Modul wäre es nicht schlecht gewesen, wenn die Beispiele für die Gruppenarbeit stärker vorgegeben gewesen wären (mangels Erfahrung der Teilnehmer kamen die gewählten Beispiele dann doch wieder meist aus dem Hochschulbereich).
- im Modul "Lehrkompetenz" fand zwar viel Gruppenarbeit statt, die aber eher inhaltlich bewertet wurde, statt sich anzuschauen, was man als Weiterbildner mit einer solche Aufgabe bezwecken will, wie sie verbessert werden könnte -> Reflexion fehlte in der Hinsicht oftmals - Aufgabenstellungen im Modul "Lehrkompetenz" waren oft unklar gestellt und die veranschlagte Zeit reichte meist nicht zur Bearbeitung - am Ende des Moduls "Lehrkompetenz" wurde herausgestellt, dass die Gestaltung von Postern ganz wichtig sei, allerdings wurden wir vorher nicht auf Postergestaltungsrichtlinien hingewiesen noch bekamen auf Nachfrage bei den Dozenten Literaturhinweise oder Beispiele für gute Poster. Es kann nicht davon ausgegangen werden, dass jeder schon Poster erstellt hat. In manchen Fachgebieten mag das gang und gäbe sein, aber nicht in allen. - Dozenten verließen manchmal während der Bearbeitung von Aufgaben den Raum und standen so nicht für Rückfragen zur Verfügung; nicht so gravierend, wenn 2 Dozenten da sind und einer im Raum bleibt, um Fragen zu beantworten - "Lehrkompetenz": auf Script wurde nicht weiter eingegangen; es wurde erwartet, dass Beispiele vor dem Kurs eingetragen werden, aber ohne die Einführung / Vertiefung in das Thema war das nicht ohne Weiteres möglich; Script ist nicht so ausführlich, dass man nur durch die Eigenlektüre in der Lage ist, das Thema vor dem Kurs hinreichend zu bearbeiten. Fehler im Script beeinträchtigten den Lesefluss, vor allem, weil die anderen Kursunterlagen weniger Fehler aufwiesen
Es fehlten in Modul "Lehrkompetenz" anregende Lernumgebungen, die didaktische Umsetzung der Inhalte - gemessen am eigenen Anspruch - war zu wenig durchdacht und nicht immer ideal durchgeführt.
Lehrkompetenz m.E. zu umfangreich

<p>Profil-TT ist so gut, dass die wenigen Verbesserungsmöglichkeiten vollstaendig genannt werden koennen: Im Modul Lehrkompetenz wurde mit viel Gruppenarbeit Begriffe der theoretischen empirischen Paedagogik erarbeitet. Manche Gruppenarbeit wirkte auf mich gekuenstelt, fast als Selbstzweck und sogar etwas nervend. Ein Abbau des Zeitanteils der Gruppenarbeiten im Modul Lehrkompetenz zugunsten eines noch hoeheren Informationsanteils ueber Ergebnisse und Studien der empirischen Paedagogik, kombiniert mit Erfahrungsberichten der Kursteilnehmer, haette mir persoendlich besser gefallen und waere wohl auch praxisnaeher.</p>
<p>Das Skript sollte man mit durchlaufenden Seitenzahlen und Inhaltsverzeichnis versehen.</p>
<p>PP wurde nicht am Laptop geübt, obwohl dies organisatorisch durchaus möglich ist, hier sollten mehr Möglichkeiten zum Üben von Handfertigkeiten vorhanden sein, weniger Kognitives Der Zeitplan macht sich etwas lang. Aus anderen Fortbildungen hat sich folgendes Schema mit je 2 Doppelstunden als günstig erwiesen: 9.30 - 11.00, 11.15 -12.45, 13.45 - 15.15, 15.30-17.00</p>
<p>-Zwar professionell gemacht aber insgesamt durch sehr formell war das Modul Lehrkompetenz -Äußerst träge und praktisch ausschließlich frontal präsentiert und mit kaum neuen Inhalten versehen war das Modul PRÄSENTATIONSTECHNIKEN</p>
<p>Zu kurz :-)</p>
<p>Das Engagement der Lehrenden war sehr unterschiedlich ausgeprägt. Die Dozenten sollten sich stärker an den Prämissen orientieren, die in dem Kurs insgesamt vermittelt werdem.</p>

Verbatims Retrospektions- Erhebung

Positive Aspekte
Kleine Lernergruppe -> hohe Teilnehmeraktivität
Die einzelnen Module waren sehr abwechslungsreich. Sie wurden gut zueinander in Beziehung gesetzt.
Die Veranstaltung hat die vielfältigen Verbesserungsmöglichkeiten aufgedeckt.
Das Modul Kohärenter Sprachgebrauch war selbst auf dem hohen Niveau des Gesamtkurses eine positive Ueberraschung.
Starke Einbeziehung der Kursteilnehmer Dadurch hoher Lerneffekt
Kollegen aus anderen Disziplinen zu treffen
- Modul Kohärenter Sprachgebrauch war am interessantesten und am besten gestaltet; dann die Module Rhetorik und Stimmbildung
Vor allem Stimmbildung, Lehrkompetenz und Rhetorik waren wirklich richtig gut & guter Zusammenhalt in der ganzen Gruppe: War wirklich sehr sehr angenehm!
nette Atmosphäre
die Auswahl und Reduzierung des Stoffes abwechslungsreiche Methodenwahl
Teilnehmerzahl von 8 Leuten praktische Übungen

Konstruktive Anmerkungen
Statt Weiterbildung stand die universitäre Lehre zu sehr im Vordergrund.
Es war zu kurz.
(siehe letzte Evaluation) Beim Modul Lehrkompetenz waere ein verstaerktes Eingehen auf aktuelle Forschungsergebnisse (Motivation nur als B-Faktor, Leistungsverbesserung allein durch Existenz einer Evaluation, etc.) sehr interessant gewesen. Statt thematisch sehr allgemeiner Gruppenarbeiten waeren Kurz\ "weiterbildungen\ " der Teilnehmer durch die Teilnehmer und Rueckkopplung durch andere Teilnehmer und des Dozenten vielleicht hilfreicher (\ "problemorientierter\ ") zur Verbesserung der Lehrkompetenz im eigenen Fachgebiet gewesen.
Anforderungen teilweise sehr hoch (vor allem kohärenter Sprachgebrauch)
Beim \ "Einsatz von Präsentationstechniken\ " war von Gruppenarbeit zu wenig zu sehen; der Dozent wurde dem Grundprinzip der ganzen Kursreihe \ "auf hohem Niveau Lehrkompetenzen zu vermitteln\ " nicht gerecht, da er sich selbst den teilweise wirklich hervorragend uns beigebrachten Techniken NICHT bediente.
sehr unterschiedliches Vorwissen bei den Seminarteilnehmern: starkes Gefälle zwischen Profis und Anfängern war teilweise hemmend

Lebenslauf

Name	Manfred Paul Kostolnik
Geburtsdatum, Ort	26.1.1979 in Wertingen
September 1986 bis Juli 1990	Grundschule Biberbach
September 1990 bis Juli 1997	Gymnasium Wertingen
August 1997 bis Juli 1998	Gymnasium bei St. Stephan, Augsburg
Juli 1998	Abschluss des Gymnasiums mit der Allgemeinen Hochschulreife
Oktober 1998 bis Juli 1999	Wehrdienst bei der II. Technischen Schule der Luftwaffe, Kaufbeuren
Oktober 1999 bis Februar 2004	Studium Magister Sprechwissenschaften (Psycholinguistik) an der LMU München Nebenfächer: Psychologie und Computerlinguistik
Februar 2004	Abschluss des Studiums mit dem Erwerb des akademischen Grades „Magister Artium“
April 2004	Aufnahme des Promotionsstudiums Sprechwissenschaft (Psycholinguistik) mit dem Nebenfach Psychologie an der LMU München
April 2004 bis Dezember 2004	Tätigkeit als Kommunikationsspezialist bei der BMW Group, Entwicklung Interieur
Seit Januar 2005	Tätigkeit als Senior Consultant bei Solution 24 GmbH & Co. KG