Jürgen Kohler

Empirische Überprüfung des Redeflusskompasses

Ein Beitrag zur Optimierung des frühzeitigen Erkennens von beginnendem Stottern

Wissenschaftliche Schriften
im Schulz-Kirchner Verlag

Reihe 13
Beiträge zur Gesundheits- und Therapiewissenschaft
Band 10

Jürgen Kohler

Empirische Überprüfung des Redeflusskompasses

Ein Beitrag zur Optimierung des frühzeitigen Erkennens von beginnendem Stottern

Schulz-Kirchner Verlag

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

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

Erstgutachter: Prof. Dr. Manfred Grohnfeldt
Zweitgutachter: Prof. Dr. Kurt Heller

Tag der mündlichen Prüfung: 01.07.2014

Besuchen Sie uns im Internet: www.schulz-kirchner.de

1. Auflage 2015
ISBN 978-3-8248-1192-2
e-ISBN 978-3-8248-0982-0
© Schulz-Kirchner Verlag GmbH, Idstein 2015
Mollweg 2, D-65510 Idstein
Vertretungsberechtigte Geschäftsführer:
Dr. Ullrich Schulz-Kirchner, Nicole Haberkamm
Druck und Bindung:
medienHaus Plump, Rolandsecker Weg 33, 53619 Rheinbreitbach
Printed in Germany

Die Informationen in diesem Buch sind von dem Verfasser und dem Verlag sorgfältig erwogen und geprüft, dennoch kann eine Garantie nicht übernommen werden. Eine Haftung der Verfasserin bzw. des Verlages und seiner Beauftragten für Personen-, Sach- und Vermögensschäden ist ausgeschlossen.

Vorwort

Diese Arbeit wurde als Dissertation an der Ludwig-Maximilians-Universität München im Juli 2014 angenommen.
Das Vorwort hat den Sinn, Entscheidungen zum Aufbau der Dissertation und Regeln der Nachvollziehbarkeit zu erläutern. Der inhaltliche Einstieg wird durch die Darstellung des Vorhabens (Kapitel 1.1) und dessen Begründung (Kapitel 1.2) in der Einleitung vollzogen.

Da in dieser Arbeit mit Berufsgruppen geforscht wurde, welche zum grössten Teil aus weiblichen Personen bestehen, wurde für diese Gruppen die **weibliche Form** benutzt (z.B. Logopäd**innen**). Bei allen anderen Gruppen wurde die **männliche Form** benutzt (z.B. Kinderärzte). Es sind aber immer beide Geschlechter gemeint und die jeweils nicht geschriebene Geschlechtsform bedeutet keine Abwertung dieses Geschlechts. Die Entscheidung für die eine oder andere Schreibweise soll lediglich die Lesbarkeit erleichtern.
Im empirischen Teil werden aus stilistischen Gründen verschiedene Begriffe für Personen verwendet, die den Redeflusskompass angewandt haben. Sie heissen Anwender, Rater, Einschätzer oder Tester. Diese Begriffe sind synonym.

Abkürzungen werden meist an den Textstellen erklärt, wo sie eingesetzt werden. Trotzdem gibt es vorweg ein Verzeichnis der häufigsten Abkürzungen.

Die zahlreichen in Klammern gesetzten Querverweise, welche auf Zusammenhänge zwischen Kapiteln oder zu Abbildungen und Tabellen hinweisen, können durch die **Nutzung des Inhaltsverzeichnisses** und des **Tabellen- und Abbildungsverzeichnisses** (vorletzter Teil dieser Dissertation) besser nachvollzogen werden.
Manchmal dienen die Anhänge nicht nur zur vertiefenden Nachvollziehbarkeit, sondern unterstützen oder erleichtern das Verständnis von Textteilen erheblich. Hier soll das **Verzeichnis der Anhänge** (letzter Teil dieser Dissertation) die schnelle Auffindbarkeit bringen. Die Anhänge selbst liegen in digitaler Form auf einem Datenträger vor und können beim Autor angefordert werden.

Die **Zitierweise** orientiert sich an den Richtlinien der Deutschen Gesellschaft für Psychologie (2007).

Die **Trennung von Theorie und Empirie** wird nicht so trennscharf gehalten wie das Inhaltsverzeichnis es suggeriert. Die theoretischen Grundlagen werden durch theoretische Spezifikationen im empirischen

Teil ergänzt und zwar dort, wo es inhaltlich angemessen bzw. gefordert ist.
Ebenfalls ist die **Trennung von Methodendarstellung und Ergebnisdarstellung** nicht so scharf wie im Inhaltsverzeichnis angegeben. Das gilt insbesondere für den quantitativen Teil, weil manche Fragestellung eine Spezifikation der Methodik erforderten.

Von den hier dargestellten Ergebnissen der Dissertation sind bisher nur partielle Vorüberlegungen durch einen Vortrag (Kohler, 2010b) und durch einen Zeitschriftenartikel (Kohler, 2011a) veröffentlicht worden. Es handelt sich um Vorüberlegungen zu der Befragung von Pädiatern und Logopädinnen (Kapitel 3.1 – 3.1.9). Diese Veröffentlichungen sind im Literaturverzeichnis gemäss §6 Abs.1 der Promotionsordnung der LMU vom 1. März 2005 angegeben.

Inhaltsverzeichnis

Vorwort VI

Inhaltsverzeichnis VIII

Abkürzungsverzeichnis XIX

1 Einleitung 1

1.1 Das Vorhaben 1

1.2 Begründung des Vorhabens 1

1.2.1 Notwendigkeit von evidenzbasierter Diagnostik 1

1.2.2 Prävention bei Stottern 2

2 Theoretischer Teil 6

2.1 Stottern in der frühen Kindheit 6

2.1.1 Phänomenologie: Normale vs. symptomatische Sprechunflüsigkeiten 6

2.1.2 Beginn, Verlauf und Remission 9

2.1.3 Ätiologie 12

2.2 Die Entscheidung für weiterführende Maßnahmen bei Verdacht auf beginnendes Stottern 16

2.3 Gegenstand der Untersuchung: Der Redeflusskompass 24

2.3.1 Der Redeflusskompass im Vergleich mit ähnlichen Instrumenten 28

2.4 Theorie der Entscheidung 31

3 Empirischer Teil 34

3.1 Nützlichkeit und Ökonomie: Subjektive Erfahrungen von Logopädinnen und Pädiaterinnen mit dem Redeflusskompass Version 1.0 36

3.1.1 Fragestellungen 37

3.1.2 Befragungen als Methode zur Erfassung von Nützlichkeit und Ökonomie: Planung und Durchführung der Datenerhebung und ihre Auswertung 38

3.1.3 Befragung der Pädiater mittels Fragebogen 39

3.1.3.1 Spezifikation Forschungsmethodik 39

3.1.3.2 Stichprobe 39

3.1.3.3 Auswertung und erste Interpretationen 39

3.1.3.3.1 Nutzer unter den Kinderärzten (Teil A) 40

3.1.3.3.2 Nicht-Nutzer unter den Kinderärzten (Teil B) 41

3.1.4 Gruppendiskussion Kinderärzte 43

3.1.4.1 Stichprobe und Spezifikation Forschungsmethodik 43

3.1.4.2 Auswertung und erste Interpretationen 43

3.1.6 Befragung der Logopädinnen mit Fragebogen 44

3.1.6.1 Spezifikation forschungsmethodisches Vorgehen 44

3.1.6.2 Stichproben 44

3.1.6.3 Auswertung und erste Interpretationen 45

3.1.6.3.1 Teil A (Gesamtstichprobe mit N = 52) 45

3.1.6.3.2 Teil B Nutzung des Redeflusskompasses (nur Nutzer mit N=10) 49

3.1.6.3.3 Teil C und D (Nutzer und Nicht-Nutzer mit N=23) 51

3.1.7 Befragung von Logopädinnen mittels Gruppendiskussion 51

3.1.7.1 Stichprobe und Spezifikation Forschungsmethodik 51

3.1.7.2 Auswertung und erste Interpretationen 52

3.1.8 Befragung von ausgewiesenen Stotterexperten 52
3.1.8.1 Stichprobe und Spezifikation Forschungsmethodik 52
3.1.8.2 Auswertung und erste Interpretationen 53
3.1.9 Zusammenfassung und weitere Interpretationen zur Befragung von Pädiatern und Logopädinnen 54
3.9.1.1 Weiterführende Interpretationen und Hypothesengenerierung 56
3.1.10 Optimierungsvorschläge für den Redeflusskompass auf Grund der subjektiven Erfahrungen von Pädiatern und Logopädinnen 58
3.1.10.1 Realisierung der Optimierungsvorschläge - Entstehung der Version 2.0 des Redeflusskompasses 61
3.2 Validität, Reliabilität, Spezifität und Sensitivität: Testdurchführungen als Grundlage zur Bestimmung der (Haupt-) gütekriterien 62
3.3 Das Re-Test-Verfahren als Setting zur Bestimmung der Gütekriterien 63
3.3.1 Durchführung 64
3.3.2 Auswertung für die Inter-Rater-Reliabilitäten 64
3.3.2.1 Ebene der Empfehlungen 64
3.3.2.2 Ebene der Gesamtpunktzahl 65
3.3.2.3 Ebene der Einzelitems 65
3.3.3 Auswertung für die Validitäten 65
3.3.3.4 Divergente Validität 65
3.3.3.5 Konvergente und konvergent-prognostische Kriteriumsvalidität 65
3.3.4 Auswertung für die Spezifität und Sensitivität 66
3.3.5 Tabellarischer Überblick zur Datenerhebung und Auswertung im Re-Test-Verfahren 67

3.3.6 Interpretation der Auswertung 68
3.3.7 Zustandekommen der Stichprobe 69
3.3.7.1 Experimentalgruppe 69
3.3.7.2 Vergleichsgruppe 70
3.3.7.3 Kontaktaufnahme mit Kooperationspartnern und Bezugspersonen 70
3.3.8 Kennzeichen der Gesamtstichprobe, Experimentalgruppe und Vergleichsgruppe 70
3.4 Bestimmung von Gütekriterien in der Zwischenauswertung 72
3.4.1 Stichproben und Spezifikationen der Durchführungsbedingungen 72
3.4.2 Bestimmung der Validitäten in der Zwischenauswertung 73
3.4.2.1 Inhaltsvalidität 73
3.4.2.2 Itemselektion und Inhaltsvalidität 73
3.4.2.3 Konvergente Kriteriumsvalidität 74
3.4.2.4 Divergente Kriteriumsvalidität 77
3.4.2.5 Zusammenfassung und Schlussfolgerungen 78
3.4.3 Die Inter-Rater-Reliabilitäten der Zwischenauswertung 80
3.4.3.1 Inter-Rater-Reliabilität auf Ebene der Empfehlungen 80
3.4.3.2 Inter-Rater-Reliabilität auf Ebene der Gesamtpunktzahl 81
3.4.3.3 Inter-Rater-Reliabilität auf Itemebene 84
3.4.3.4 Itemselektion mit Kennwerten der Inter-Rater-Reliabilität 86
3.4.3.5 Zusammenfassung und Interpretation 87
3.4.4 Spezifität der Zwischenauswertung 87
3.4.5 Zusammenfassung der Zwischenauswertung 89
3.5 Die Inter-Rater-Reliabilitäten im Re-Test-Verfahren 90

3.5.1 Deskription der Übereinstimmung und Berechnung der Kennwerte auf Ebene der Empfehlungen 90
3.5.2 Deskription der Übereinstimmung und Berechnung der Kennwerte auf Ebene der Gesamtpunktzahl 92
3.5.3 Deskription der Übereinstimmung und Berechnung der Kennwerte auf Ebene der Einzelitems 96
3.5.4 Rechnerische Itemselektion 106
3.5.5 Zusammenfassung und erste Interpretationen 107
3.6 Spezifität, Sensitivität und Prozentuale Übereinstimmung im Re-Test-Verfahren 110
3.6.1 Bestimmung von Spezifität und Prozentualer Übereinstimmung mit der zuvor gestellten logopädischen Diagnose 111
3.6.2 Bestimmung von Sensitivität, Spezifität sowie Prozentualer Übereinstimmung mit dem tatsächlichen Beratungs- und Diagnostikbedarf 112
Einschub: Spezifikation der Erhebungsmethode für die ProgVal 113
3.6.3 Zusammenfassung und erste Interpretationen 119
3.7 Kriteriumsvaliditäten: Konvergent-prognostische und divergente Validität im Re-Test-Verfahren 121
3.7.1 Konvergent-prognostische Kriteriumsvalidität 122
3.7.1.1 Berechnung der Kennwerte für die Ersttestung der Gesamtstichprobe 122
3.7.1.2 Berechnung der Kennwerte für die Zweittestung der Gesamtstichprobe 122
3.7.1.3 Berechnung der Kennwerte für die Ersttestung der Experimentalgruppe 123

3.7.1.4 Berechnung der Kennwerte für die Zweittestung der Experimentalgruppe 124
3.7.1.5 Zusammenfassung und erste Interpretationen 125
3.7.2 Divergente Validität 126
3.7.2.1 Deskription des Zusammenhangs und Berechnung der statistischen Kennwerte für die Ersttestung 126
3.7.2.2 Deskription des Zusammenhangs und Berechnung der statistischen Kennwerte für die Zweittestung 128
3.7.2.3 Zusammenfassung und erste Interpretationen zur divergenten Kriteriumsvalidität 129
3.7.3 Zusammenfassung und weitere Interpretationen zu den Kriteriumsvaliditäten 130
3.8 Einfluss der Vorerfahrung der Tester und Einfluss der Durchführungsbedingungen auf die Gütekriterien im Re-Test-Verfahren 132
3.8.1 Abhängigkeit von der Vorerfahrung 132
3.8.1.1 Merkmale der Tester 132
3.8.1.2 Prüfung der Vorerfahrung als Einfluss auf die Übereinstimmung zwischen den Empfehlungen von Erst- und Zweittestung 134
3.8.1.3 Prüfung der Vorerfahrung als Einfluss auf die Übereinstimmung zwischen den Empfehlungen der Ersttestung bzw. Zweittestung und der ProgVal 134
3.8.2 Abhängigkeit von den Durchführungsbedingungen 136
3.8.2.1 Deskription der Durchführungsbedingungen 136
3.8.2.2 Prüfung des Settings als Einfluss auf die Übereinstimmung zwischen den Empfehlungen der Ersttestung und Zweittestung 137

3.8.2.3 Prüfung des Settings als Einfluss auf die Übereinstimmung zwischen den Empfehlungen der Zweittestung und der ProgVal 137
3.8.2.4 Abhängigkeit von der Beobachtungsmöglichkeit 138
3.8.3 Zusammenfassung und erste Interpretationen 140
3.9 Videofalldarbietungen als quasi-experimentelles Setting zur Differenzierung der Gütekriterien 142
3.9.1 Beschreibung des Settings 142
3.9.1.1 Merkmale der Einschätzer 143
3.9.1.2 Merkmale der Videos 145
3.9.1.4 Rekrutierung der Rater (=Einschätzer) 145
3.9.1.5 Ablauf der Durchführung des quasi-experimentellen Settings 146
3.9.2 Die Übereinstimmung zwischen den Ratern im quasi-experimentellen Setting 146
3.9.2.1 Prozentuale Anteile der Übereinstimmung zwischen den Ratern insgesamt 148
3.9.2.2 Abhängigkeit der Inter-Rater-PAgesamt von der Individualität des Falles 150
3.9.2.3 Rangreihe der Itemzuverlässigkeit insgesamt 150
3.9.2.4 Zusammenfassung für die Inter-Rater-PAgesamt 151
3.9.2.5 Prozentuale Anteile der Übereinstimmung hinsichtlich der Erfahrung (Inter-Rater-PAerfahr) 151
3.9.2.6 Differenzierung der Inter-Rater-PAerfahr 153
3.9.2.7 Abhängigkeit der Inter-Rater-PAerfahr vom Fall und der Erfahrungsgruppe pro Einzelitem 154
3.9.2.8 Rangreihe der Itemzuverlässigkeit hinsichtlich der Erfahrung 156
3.9.2.9 Zusammenfassung für die Inter-Rater-PAerfahr 157

3.9.2.10 Prozentuale Anteile der Übereinstimmungen hinsichtlich der Berufszugehörigkeit (Inter-Rater-PAberuf) 157
3.9.2.11 Differenzierung der Inter-Rater-PAberuf 159
3.9.2.12 Abhängigkeit der Inter-Rater-PAberuf vom Fall und der Berufszugehörigkeit pro Einzelitem 159
3.9.2.13 Rangreihe der Itemzuverlässigkeit hinsichtlich der Berufe 160
3.9.2.14 Zusammenfassung für die Inter-Rater-PAberuf 161
3.9.2.15 Zusammenfassender Vergleich zwischen Inter-Rater-PAgesamt, Inter-Rater-PAerfahr und Inter-Rater-PAberuf 161
3.9.3 Intuitive Übereinstimmung der Rater zur Rangreihe des Schweregrades (Trefferquote der Intuition) 164
3.9.3.1 Intuitive Schweregradzuschreibung unabhängig von Erfahrung und Beruf (Intuitiv-Treffer-PAgesamt) 167
3.9.3.2 Differenzierung der Intuitiv-Trefferquote hinsichtlich der Erfahrung (Intuitiv-Treffer-PAerfahr) 170
3.9.3.3 Differenzierung der Intuitiv-Trefferquote hinsichtlich der Berufszugehörigkeit (Intuitiv-Treffer-PAberuf) 172
3.9.3.4 Zusammenfassung und erste Interpretationen 174
3.9.4 Modell zum kognitiven Verarbeitungsprozess bei der Einschätzung der Videos 176
3.9.4.1 Annahmen für die kognitiven Verarbeitungsprozesse der unterschiedlichen Erfahrungsgruppen 176
3.9.4.2 Interpretationen der Ergebnisse von Inter-Rater-PAerfahr und Intuitiv-Treffer-PAerfahr unter Annahme der verschiedenen kognitiven Verarbeitungsmechanismen 179
3.9.4.2.1 Interpretation für die Experten 180
3.9.4.2.2 Interpretation für die Personen mit mittlerem Vorwissen 181
3.9.4.2.3 Interpretation für die Personen ohne Vorwissen aber mit Kurzfortbildung 181

3.9.4.2.4 Interpretation für die Personen ohne Vorwissen 181
3.9.4.2.5 Interpretation für die Kinderärzte als Sonderfall der Personen ohne Vorwissen 182
3.9.4.3 Zusammenfassung 182
3.10 Einzelinterviews, Gruppendiskussion und Gruppenbefragung als Methoden zur Differenzierung der Validität und Klärung fehlender Übereinstimmung im Re-Test-Verfahren und quasi-experimentellen Setting 184
3.10.1 Konzeption einer gegenstandsangemessenen Inhaltsanalyse 185
3.10.2 Auswertung der Einzelinterviews mit den Experten 189
3.10.3 Auswertung der Einzelinterviews mit den Anwendern 191
3.10.4 Auswertung der Gruppenbefragungen 193
3.10.5 Auswertung Gruppendiskussion mit Experten 196
3.10.6 Integration – Kategorienvergleich 198
3.10.7 Validierung - Rücküberführung der Kategorien auf das Ausgangsmaterial 200
3.10.8 Zusammenfassung und erste Schlussfolgerungen 204

4 Diskussion: Zusammenfassende Integration aller Ergebnisse mit abschließender Interpretation 207
4.1 Der Stotterkompass - Konstruktion eines Instrumentes zur Identifikation des Abklärungsbedarfes bei beginnendem Stottern 218
4.1.1 Ziele und allgemeine Konstruktionsprinzipien 218
4.1.2 Das Instrument als ganzes 220
4.1.3 Das kommentierte Instrument 224

4.2 Der Stotterkompass *professionell* - Konstruktion eines Instrumentes zur Abklärung des Beratungs-, Diagnostik- und eventuell Therapiebedarfs bei beginnendem Stottern 234

4.2.1 Das kommentierte Instrument 235

5 Reflexion des Arbeitsprozesses und der angewandten Forschungsmethoden 245

5.1 Das „ich“ in einer wissenschaftlichen Arbeit 245

5.2 Zusammenspiel von externer und interner Validität.....245

5.3 Das Scheitern von Hypothesen im quasi-experimentellen Setting 246

5.4 Das übergreifende Prinzip der Falsifikation 246

5.5 Validierung und Komplementarität durch Triangulation 247

5.6 Unabhängige Durchführung, Auswertung und Interpretation.247

5.7 Ein möglichst großes N.....248

5.8 Was ich anders machen würde.....248

5.9 Zusammenarbeit mit Studierenden.....249

5.10 Statistik als Hilfsmittel 249

5.11 Verschiedene Rollen im Forschungsprojekt 251

5.12 Schriftliche Ausarbeitung als Abbild des Forschungsprozesses.. 252

5.13 Bedürfnis nach Genauigkeit.....253

5.14 Forschungsmethodologische Einordnung.....253

5.14.1 Empirische vs. theoretische Arbeit.....254

5.14.2 Quantitative vs. qualitative Sozialforschung 254

5.14.3 Grundlagenforschung vs. angewandte Forschung 254

5.15 Forschungsethik.....255

5.16 Beratung als Qualitätsmerkmal von Forschung 256

6 Ausblick 258

6.1 Expertendiskussion über Stotterkompass und Stotterkompass professionell 258

6.2 Rückmelden von Forschungsergebnissen als Möglichkeit zur Erhöhung der Güte bei der Identifikation des beginnenden Stotterns 258

6.3 Qualitative Erweiterung der Instrumente 259

6.4 Empirische Überprüfung von Stotterkompass und Stotterkompass professionell 259

6.5 Verlaufskontrolle mit *Stotterkompass* und *Stotterkompass professionell* systematisieren 260

6.6 Ausweitung von Fragestellungen und Konzeption einer best-practice der Anwendung 260

Literaturverzeichnis 265

Abbildungs- und Tabellenverzeichnis 271

Anhang 277

Verzeichnis der Anhänge 277

Datenträger mit allen Anhängen gemäß Verzeichnis (außer Anhang 3.9 A) 287

Datenträger mit Anhang 3.9 A – drei Videos mit sprechunflüssigen Kindern 288

Abkürzungsverzeichnis

Erklärung von benutzten Abkürzungen

AAUS - Aachener Analyse unflüssigen Sprechens

DIMDI - Deutsches Institut für Medizinische Dokumentation und Information

Empf - Empfehlungen des Redeflusskompass

IQWIG - Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

ICD-10 - Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme

Log Diag - Logopädische Diagnose

PA - Prozentuale Anteile (der Übereinstimmung)

PÜ - Prozentuale Übereinstimmung

ProgVal - Name der Variablen, die in der Untersuchung das tatsächliche Beratungs- und Diagnostikgeschehen repräsentiert

RfK - Redeflusskompass

RfK-Tot - Abkürzung für die Gesamtpunktezahl (=Summenscore) des Redeflusskompasses

SLS - Sreening List Stuttering

SPSS - Superior Performing Software System (Name eines Statistikprogramms)

SSI - Stuttering Severity Instrument (genormtes Instrument zur Bestimmung des Schweregrade bei Stottern)

WHO - Weltgesundheitsorganisation

1 Einleitung

1.1 Das Vorhaben

Der Redeflusskompass ist ein Instrument zur Früherfassung bei Verdacht auf beginnendes Stottern. Er ist als Entscheidungshilfe für Fachpersonen aus dem pädagogisch-medizinischen Handlungsfeld konzipiert.
Die systematische Begleitung und Erfassung des Instrumenteneinsatzes soll zur Überprüfung des Redeflusskompasses führen, um Validität, Reliabilität, Objektivität Sensitivität, Spezifität aber auch Nützlichkeit und Ökonomie empirisch zu erfassen und diskutieren zu können. Die Optimierung des Instrumentes ist ebenfalls Ziel des Vorhabens. Die Triangulation quantitativer und qualitativer forschungsmethodischer Strategien soll zu einer mehrdimensionalen empirischen Überprüfung führen. Mit den erhobenen Daten und ihrer Interpretation soll ein Diskussionsbeitrag geleistet werden, um die bisher im Bereich Früherfassung der Redeflussstörung „Stottern“ bestehende empirische Lücke im deutschsprachigen Raum zu schließen (Johannsen, 2001; Neumann & Schneider, 2011).
Im folgenden Abschnitt soll das Vorhaben unter Bezug auf die aktuelle Diskussion in der Fachwelt zum Thema „Identifikation von beginnendem Stottern“ differenziert begründet werden.

1.2 Begründung des Vorhabens

1.2.1 Notwendigkeit von evidenzbasierter Diagnostik

Evidenzbasiertes Handeln wird in der Sprachtherapie nicht nur für therapeutisches Handeln, sondern auch für diagnostische und präventive Interventionen gefordert (Beushausen, 2011, S. 121). Die klassischen testtheoretischen Gütekriterien Validität, Reliabilität und Objektivität aber nicht zuletzt auch die Spezifität und die Sensitivität bei Selektionsentscheidungen gehören zur Evidenz eines Verfahrens, welches vorgibt eine Entscheidungshilfe zu sein.
Im Abschlussbericht des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWIG) zur „Früherkennung von umschriebenen Störungen des Sprechens und der Sprache“ (IQWIG, 2006) wird zur Erfüllung dieser Forderung festgestellt:
„Gleichzeitig sind für die deutschsprachigen Instrumente gegenwärtig keine verlässlichen Gütekriterien berichtet, die eine Entdeckung von Kin-

dern mit umschriebenen Sprachentwicklungsstörungen in der Gesamtbevölkerung erlauben.“ (IQWIG, 2006 S.6)
Zwar schließt diese Aussage nicht explizit die Redeflussstörungen mit ein, doch zeigt der Überblick zu den in der einschlägigen Fachliteratur vorgeschlagenen Verfahren (s. Kapitel 2.2), dass auch im Bereich Stottern solche evidenzbasierten Instrumente fehlen. Es gibt offensichtlich keine systematisch-empirischen Untersuchungen zur Anwendung von kriteriengeleiteten Empfehlungen für die weiterführende Diagnostik und Beratung bei beginnendem Stottern.
Das Ziel der forschungsmethodisch systematischen Reflexion von Untersuchungsverfahren wird im Weiteren so benannt:„Sofern die methodischen Grundlagen geschaffen werden, eröffnen sie die Möglichkeit, den potenziellen Nutzen und potenziellen Schaden eines Screenings im Rahmen einer vergleichenden Studie zu untersuchen.“ (IQWIG, 2006, S.6)
Die vorliegende Untersuchung will einen Beitrag zur Diskussion der Evidenz des Redeflusskompasses liefern, indem der potentielle Nutzen genauso wie der potentielle Schaden der Anwendung erfasst wird. Außerdem sollen auch weiterführende und verallgemeinernde Schlussfolgerungen zur Identifikation des beginnenden Stotterns gezogen werden.

1.2.2 Prävention bei Stottern

Die Weltgesundheitsorganisation WHO ordnet in der ICD 10 GM (2011) das Stottern der Kategorie F90-F98 „Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend“ zu. Aus diesem Grunde werden vorsorgliche Maßnahmen gefordert. Der Begriff „Prävention“ stellt dabei die zentrale Kategorie dar und wird heute dann verwendet, wenn es darum geht, Krankheiten, Störungen, Beeinträchtigungen und Behinderungen
a) frühzeitig zu erkennen und
b) deren Folgen einzudämmen (Hurrelmann et al, 2007, S. 15).
Pädiater und Logopädinnen nehmen bei der Prävention von Redeflussstörungen eine entscheidende Rolle ein, da Bezugspersonen von vermeintlich betroffenen Kindern diese Fachleute aufsuchen, um Informationen zum Stottern zu erhalten. Andere Berufsgruppen wie z.B. Erzieherinnen spielen ebenfalls eine wichtige Rolle, weil sie oft intensiven Kontakt mit den Kindern haben und ihnen eine unflüssige Sprechweise auffällt. Die frühe Beurteilung von Sprechunflüssigkeit birgt die Chance, die Indikation von weiteren pädagogisch-therapeutischen Schritten zu erkennen und damit eine frühe Intervention bei Bedarf möglich zu machen. Bisherige Untersuchungen im deutschsprachigen Raum (Johannsen, 2001) legen nahe, dass die Remissionschancen umso größer sind, je früher geeignete Interventionen angegangen werden. Die Realität sieht

aber anders aus: „Derzeit werden stotternde Kinder in der Regel leider erst dann behandelt, wenn sich das Stottern schon über Jahre hinweg chronifiziert hat. Für die betroffenen Kinder erhöht sich so die Wahrscheinlichkeit, eine ausgeprägte Begleitsymptomatik und entsprechende Behandlungsresistenz zu entwickeln.“ (Sandrieser & Schneider, 2008, S. 18)
Bisher gibt es im Bereich Früherfassung der Redeflussstörung „Stottern“ keine empirisch abgesicherten Daten, welche kriteriengeleitet den Beratungs- bzw. Therapiebedarf erfassen würde (Johannsen, 2001) und das, obwohl quantitative also auch qualitative Kriterien bekannt sind, die pathologische Redeunflüssigkeiten von nicht-pathologischen schon im Vorschulalter unterscheiden können (s. Kapitel 2.1).
Trotz der vorhandenen Kriterien muss konstatiert werden, dass die Differentialdiagnose zwischen normaler Sprechunflüssigkeit und beginnendem Stottern kein einfaches Unterfangen ist. So stellen Ochsenkühn & Thiel fest: „Die Gefahr, dass beginnendes Stottern irrtümlich als altersgemäße Sprechunflüssigkeit eingeschätzt wird, ist durchaus gegeben, da es sowohl Anteile ganz normaler Sprechunflüssigkeiten als auch Symptome des manifesten Stotterns aufweist.“ (2005, S. 12)
Die empirische Absicherung der Früherkennung von Stottern erscheint nicht nur aufgrund der Komplexität des Gegenstandes notwendig, sondern auch, weil die Entscheidung zur Notwendigkeit einer therapeutischen Intervention oftmals gar nicht bei den Sprachtherapeuten liegt, sondern bei Kinderärzten, die auf dem Gebiet der Redeflussstörungen nicht über aktuelles und fundiertes Fachwissen verfügen (Baumann, Boltshauser & Braun, 2008). Die Bundesvereinigung Stotterer-Selbsthilfe (2009) in Deutschland vermutet gar: „Stottern wird bei der ärztlichen Untersuchung von Vorschülerinnen bisher meist nicht einmal hinterfragt.“ (2009, S. 279)
Aber auch in der einschlägigen pädagogisch-therapeutischen Praxis scheint es Unsicherheiten in Bezug auf die Differentialdiagnose zu geben. Natke (2005) formuliert dazu folgendes:

> Es war und ist teilweise immer noch verbreitet, das Auftreten stottertypischer Sprechunflüssigkeiten mit Begriffen wie Entwicklungsstottern, physiologisches Stottern, Entwicklungsunflüssigkeiten oder entwicklungsbedingte Sprechunflüssigkeiten zu belegen. Dies geschieht entweder, weil nicht zwischen stottertypischen und normalen Sprechunflüssigkeiten unterschieden wird oder weil eine Remission für sehr wahrscheinlich gehalten wird (...). Das Auftreten stottertypischer Sprechunflüssigkeiten vor diesem Hintergrund als Teil einer normalen Entwicklung darzustellen, wird dem Phänomen nicht gerecht und ist auch nicht hilfreich.“ (Natke, 2005, S. 44)

Zusammenfassend kann man die empirische Absicherung der Früherkennung von beginnendem Stottern hinsichtlich ihrer Praxisrelevanz mit den Worten der Bundesvereinigung Stotterer-Selbsthilfe (2009) begründen: „Erkennen von Stottern darf kein Zufall sein." (2009, S. 279)
Diese Aussage gilt insbesondere dann, wenn man bei der Entstehung von manifestem Stottern die psychosoziale Dimension als mitentscheidend und auch als pädagogisch-therapeutisch beeinflussbar begreift. Das sogenannte „Dynamische Modell" (Hansen & Iven, 2002, S.16) nimmt an, dass die Verstärkung der symptomatischen Stotteranteile aufgrund negativer Umweltreaktionen entsteht und/oder, weil das Kind sein Sprechen als unangenehm und schwierig empfindet.
Die Früherkennung von beginnendem Stottern hat bei dieser Annahme für die zweite Säule des Präventionsgedankens, der Eindämmung von unerwünschten Folgen, ein besonderes Gewicht, da negative Umweltreaktionen und Selbstwahrnehmungsprozesse als pädagogisch-therapeutisch beinflussbar gelten (Lattermann et al, 2008).
Auch die aktuelle Diskussion in der einschlägigen Fachwelt zur Prävalenz und Inzidenz (Yairi & Ambrose, 2013; Brocklehurst, 2013) zeigt unabhängig von den unterschiedlichen Positionen der Diskutanten die Notwendigkeit einer validen Früherkennung auf. Yairi & Ambrose (2013) formulieren den Bedarf an erhöhter Wachsamkeit für das beginnende Stottern so: „In spite of the fact that sooner or later many children experience natural recovery, their initial large number also requires greater availability of clinical services" (2013, S.83).
Brocklehurst (2013), der andere Zugänge zur Bestimmung von Prävalenz und Inzidenz fordert, formuliert aber für die Identifikation des beginnenden Stotterns Ähnliches wie Yairi & Ambrose oben: „Arguably, with respect to incipient stuttering, the most useful definitions are those that bestenable us to identify the children who are at increased risk of persisting in their stuttering and who are most likely to benefit from therapy." (2013, S.6)

Forschungen, die sich bisher mit der Identifikation von beginnendem Stottern beschäftigt hatten, legten vor allem Wert auf die Symptomatik der Redeunflüssigkeit an sich. Dabei ging es um die genaue Differenzierung von stottertypischen vs. normalen Unflüssigkeiten (z.B. Yairi, 1992 und 1997; Zebrowski, 1994; Sandrieser, 2004; Sawyer & Yairi, 2010) und den Zusammenhang mit linguistischen wie z.B. phonologischen oder syntaktischen Phänomenen (z.B. Byrd et.al, 2007; Ntourou, 2011) sowie um den Zusammenhang mit psychologischen, meist emotionalen Aspekten (z.B. Karras et.al, 2006 oder Ntourou et.al, 2013).
In der Regel wurden Experten eingesetzt, um das Gefährdungspotential solcher primären und sekundären Symptome für eine Chronifizierung zu erfassen. Auch wenn ausdrücklich Gütekriterien von Messverfahren bei

Stottern oder frühkindlichem Stottern bestimmt wurden, waren dabei meist extra geschulte Experten involviert (z.B. Hubbord, 1998; Bothe, 2008; Einarsdóttir & Ingham, 2009). Die vorliegende Untersuchung versucht eine andere Perspektive einzunehmen: Indem ein existierendes Instrument zur Früherfassung ob seiner Güte untersucht und dabei die externe Validität fokussiert wird, wird vielmehr die Frage nach der Anwendbarkeit der bis dato bekannten Kriterien in der Praxis gestellt. Es geht um die Frage, ob Personen, die bei der Prävention von Stottern und der Entscheidung zur Beratung bzw. Diagnostik mitwirken können, die bekannten Kriterien erfassen und tatsächlich als Entscheidungshilfe benutzen können. Diese Perspektive scheint im Lichte der aktuellen Datenlage im Forschungsbereich des frühkindlichen Stotterns ein logischer Schritt zu sein. So kann man feststellen, dass eine derzeit laufende Untersuchung von Neumann & Schneider (2011) mit dem Ziel, ein Screeninginstrument zur Identifikation von Stottern im Vorschulalter zu erschaffen, ausdrücklich mit Fachpersonen (=Kinderärzte und Erzieherinnen) arbeitet, die bei der Früherkennung von Stottern eine Rolle spielen. Leider liegen bis dato noch keine Ergebnisse dieser aufwändig angelegten Untersuchung vor.
Yairi (2013) formuliert als nächsten (oder übernächsten?) Schritt die Anwendbarkeit von genetischen Analysen oder bildgebenden neurophysiologischen Untersuchungen für die Vorhersagbarkeit der Chronifizierung des Stotterns. Solange diese Möglichkeiten jedoch noch nicht existieren, ist es wichtig, Entscheidungen zum Diagnostik-, Beratungs- und Therapiebedarf mit sozialwissenschaftlichen Methoden zu optimieren. Die vorliegende Untersuchung will dafür einen Beitrag liefern.

2 Theoretischer Teil

Nach der vorangegangenen differenzierten Begründung des Vorhabens werden für das Thema relevante theoretische Hintergründe dargestellt.
Dazu gehört eine Aufarbeitung des aktuellen Wissensstandes zum Stottern in der frühen Kindheit (zur Nachvollziehbarkeit der Literaturrecherche siehe Anhang 2. A). Eine ausführliche Darstellung des Redeflusskompasses erfolgt unter Bezugnahme auf den zuvor dargestellten aktuellen Wissensstand. Es werden außerdem ähnliche Verfahren zur Früherfassung des kindlichen Stotterns vorgestellt, und es wird der Begriff „Entscheidung" theoretisch reflektiert, da der Redeflusskompass ausdrücklich als *Entscheidung*shilfe gedacht ist.

2.1 Stottern in der frühen Kindheit

Aus der großen Fülle der empirischen Befunde zum frühkindlichen Stottern wähle ich diejenigen aus, die für die Entscheidung für eine Beratung oder Diagnostik bei Verdacht auf Stottern die größte Relevanz haben.

2.1.1 Phänomenologie: Normale vs. symptomatische Sprechunflüssigkeiten

Alle Menschen sprechen unflüssig. Es gibt keine 100% Flüssigkeit in der menschlichen Sprechweise. Insofern gibt es normale Sprechunflüssigkeiten, die sich folgendermaßen darstellen lassen: Lockere Wiederholungen von ein- oder mehrsilbigen Wörtern und Teil-Wortwiederholungen oder Satzteilen, Revisionen, Satzabbrüche, ungefüllte oder gefüllte Pausen (z.B. „ähm") und Einschübe. Auch lockere Dehnungen können dazu gehören, die meist als Betonungen interpretiert werden können. Bei jüngeren Kindern wird die Häufung dieser normalen Unflüssigkeiten auch als „entwicklungsbedingte Redeunflüssigkeiten" bezeichnet, deren Ursache man auf die Unreife des gesamten Sprachsystems zurückführt. Das beinahe synonyme Begriffspaar „funktionelle Redeunflüssigkeit" fokussiert die Unflüssigkeit als Folge von Unregelmäßigkeiten in der Sprachhandlungsplanung. Dann dienen die Unflüssigkeiten dem sprachpragmatischen Zeitgewinn für den Wortabruf, der syntaktischen Planung und der gedanklichen Strukturierung während der Sprachproduktion (Starkweather, 1987). Sie stellen einen erforderlichen Schritt beim Aufbau der hochautomatisierten Sprech- und Sprachsteuerungsabläufe dar. Funktionelle Unflüssigkeiten sind auch bei Erwachsenen zu beobachten.
Symptomatische Unflüssigkeiten hingegen unterscheiden sich quantitativ von den normalen Unflüssigkeiten durch eine allgemein größere Auftretenshäu-

figkeit, ihrer größeren Anzahl von Wiederholungseinheiten und ihrer längeren Dauer. Qualitativ sind die symptomatischen Unflüssigkeiten durch eine schnellere Repetition von kürzeren Einheiten (Silbe, Laut) sowie eventuell auch von Dehnungen oder Blockierungen gekennzeichnet. Manchmal ist ein höherer Grad an allgemeiner Spannung im Mund und Gesichtsbereich zu erkennen und es sind spannungsreiche ganzkörperliche Mitbewegungen zu beobachten
(u.a. Natke, 2005; Sandrieser, 2004; Sandrieser & Schneider, 2008). Nicht stotternde Kinder zeigen äußerst selten symptomatische Unflüssigkeiten, während bei Kindern, welche symptomatische Unflüssigkeiten zeigen auch normale Unflüssigkeiten vorkommen. Bei Kindern, welche schon länger symptomatische Unflüssigkeiten zeigen, sind die normalen Unflüssigkeiten wahrscheinlich sogar häufiger als bei nicht stotternden Kindern (Sandrieser, 2004, S.136). Die Grenzen zwischen normalen vs. stottertypischen Unflüssigkeiten sind allerdings fließend. Guitar (2014, S.113ff) schlägt daher den Begriff „borderline stuttering“ vor, der die Überlappung von normalen Unflüssigkeiten und beginnendem Stottern treffen soll. Dieser Typus ist für eher jüngere Kinder (1,5 bis 3,5 Jahre) reserviert und grenzt sich gegenüber den normalen Unflüssigkeiten insbesondere durch die größere Auftretenshäufigkeit der Laut-, Silben- und Teilwortwiederholungen insgesamt und der Anzahl der Wiederholungen innerhalb des Stottereregnisses ab. Guitar (2014) erklärt den Terminus „borderline stuttering“ mit der Schwierigkeit einer eindeutigen Differenzierung zwischen normalen Unflüssigkeiten und dem beginnenden Stottern aufgrund der Fluktuation der Symptomatik:

> (...) I also agree that there is much overlap between the disfluencies of stuttering children and the disfluencies of normally disfluent children. Moreover (...) these children often go back and forth between stuttering and normal disfluency over a period of months. For this reason, we use the term „borderline“ to indicate that these children are neither entirely normally disfluent nor definititly stuttering. (Guitar, 2014, S.120)

Im deutschsprachigen Raum findet man meist nur binäre Gegenüberstellungen von normalen Unflüssigkeiten vs. stottersymptomatischen Unflüssigkeiten, wobei die Autoren in der Regeln auf den Grenzbereich zwischen diesen beiden Kategorien hinweisen, ihn aber nicht mit einer extra Kategorie explizit hervorheben. Eine Ausnahme ist Bosshardt (2008), der audrücklich von „Sprechunflüssigkeiten bei Grenzfällen“ spricht. Seine tabellarische Auflistung zeigt dementsprechend drei Kategorien, wie folgende Abbildung zeigt: