



Einsatz von RedeflussKompass 3.0 zur Früherkennung und StotterKompass zur Ersterfassung/Erstberatung*

Jürgen Kohler & Wolfgang Braun

Zusammenfassung

Hintergrund: Auf dem Hintergrund von wenig befriedigenden Ergebnissen bei der Evaluation von Screeninginstrumenten zur Identifikation des beginnenden Stotterns im deutschsprachigen Raum wurde eine optimierte Version des RedeflussKompasses und der neu geschaffene StotterKompass einer intensiven Überprüfung unterzogen.

Fragestellung/Ziele: Die Untersuchung zielte auf die Erkundung von Güte Merkmalen der genannten Instrumente ab und auf die Schaffung einer Datenbasis zur Optimierung von Früherkennung, Ersterfassung und Erstberatung bei beginnendem Stottern.

Methodik: Die Ergebnisse der Durchführungen bei N = 60 Fällen wurden mit den wahren Werten des Bedarfs an weiteren Interventionen konfrontiert, um quantitative Kennwerte der Güte beider Instrumente zu ermitteln. Außerdem wurde eine breit angelegte Dokumentenanalyse durchgeführt, deren Daten vornehmlich aus einer qualitativen Befragung der Anwender und Beobachtung der Durchführung generiert wurden.

Ergebnisse: Es zeigte sich eine sehr gute Sensitivität und gute Spezifität des RedeflussKompasses, sowie sehr gute Werte für seine Kriteriumsvalidität. Die Sensitivität der StotterKompasses war durchaus befriedigend, jedoch konnte seine Spezifität nicht überzeugen. Seine Kriteriumsvalidität hingegen kann als sehr gut bezeichnet werden. Die nur für den StotterKompass berechnete Inter-Rater-Reliabilität hängt stark von der Berechnungsebene ab. Für die Ebene der Empfehlungen ist sie mittelmäßig, für die Ebene der Gesamtpunktzahlen sehr gut.

Die qualitativen Analysen erbrachten reichhaltige Einblicke in die Praxis der Anwendung beider Instrumente, wobei im Beitrag die gleichermaßen positiven als auch kritischen Ergebnisse beispielhaft dargestellt werden.

Schlussfolgerungen: Insgesamt darf man die quantitativ bestimmten Gütekriterien als gut bis befriedigend werten. Insbesondere die qualitative Analyse zeigte, dass der komplementäre Einsatz beider Kompass für die Ersterfassung und Erstberatung nützlich und wertvoll zu sein scheint. Außerdem konnte vornehmlich aus der qualitativen Analyse die Optimierung beider Instrumente abgeleitet werden. So war es möglich, digitale Versionen der Kompass zu schaffen.

Schlüsselwörter

Beginnendes Stottern, Früherkennung, Ersterfassung, Erstberatung

Abstract

Background: Based on unsatisfactory results from the evaluation of screening instruments to identify beginning stuttering in the German speaking area, an optimised version of the RedeflussKompass as well as the newly developed StotterKompass were subject to an intensive review.

Question/goals: The analysis aimed to explore quality characteristics of the above-mentioned tools as well as to create a data base to optimise the early detection, first recording and first consultation in the case of beginning stuttering.

* Dieser Beitrag hat das Peer-Review-Verfahren durchlaufen.

Methodology: The results of the execution with n = 60 cases were confronted with the demand for further interventions to identify quantitative quality indicators of both instruments. Furthermore, a wide document analysis was conducted, data for which was mostly generated through a qualitative interview with users as well as observations of the conduction.

Results: A very good sensitivity and a good specificity of the RedeflussKompass, as well as very good values for its criteria validity are shown. The sensitivity of the StotterKompass is highly satisfactory. However, the specificity is not convincing. The criteria validity can be labelled as very good. The inter-rater-reliability which is only calculated for the StotterKompass depends strongly on the calculation level. For the recommendation level the reliability is average, for the overall score level it is very good.

The qualitative analysis generated substantial insights into the practice of the application of both instruments. The article portraits equally the positive as well as the sceptical results using examples.

Conclusion: Overall, the quantitatively defined quality criteria can be rated as good to satisfactory. Especially the qualitative analysis highlights that the complimentary use of both compasses seems to be highly useful and valuable for the first recording and first consultation. Furthermore, an optimisation for both instruments can be deduced mostly from the qualitative analysis. This allowed to develop digital versions of both compasses.

Keywords

Beginning stuttering, early detection, first recording, first consultation

1 Einleitung mit theoretischem Hintergrund

1.1 Kompakte Annäherung an das Phänomen Stottern

1.1.1 Was ist Stottern?

Eine Million Menschen stottern im deutschsprachigen Raum (D, CH, A). Stottern ist neben den Störungsbildern Poltern und Mutismus dem Bereich Störungen der Rede zu zuordnen. Stottern kennzeichnet sich durch unfreiwillige Wiederholungen von Lauten, Silben und einsilbigen Wörtern, Dehnungen und Blockierungen. Diese Unflüssigkeiten werden als Kernsymptome des Stotterns bezeichnet. In der S3-Leitlinie Redeflussstörungen (Neumann, Euler, Bosshardt, Cook, Sandrieser, Schneider, Sommer & Thum, 2017) wird Stottern wie folgt definiert: eine „zentralnervöse Störung des Sprechens und seiner Planung, die in der Kindheit hauptsächlich auf Grund einer genetischen Disposition zustande kommt. Es umfasst eine Kernsymptomatik mit stottertypischen Sprechunflüssigkeiten und eine Begleitsymptomatik mit vegetativen, motorischen und emotionalen Reaktionen“ (S. 25). Als quantitativer Schwellenwert für die Diagnose Stottern wird eine Stotterhäufigkeit von 3 % aller Silben in einer repräsentativen Sprechprobe genannt (Neumann et al. 2017, Sandrieser 2017). Dieser Wert ist aber als Parameter für das weitere Vorgehen und die Einschätzung eines Stotterschweregrades nur sehr bedingt relevant. Auch bei einer niedrigeren Stotterrate können die emotionale Belastung durch die Kommunikationsbeeinträchtigung, die Qualität des Symptoms (z. B. der Grad der Anspannung im Symptom oder die Dauer des einzelnen Symptoms) sowie eine sich entwickelnde Begleitsymptomatik (Vermeideverhalten, Mitbewegungen) eine Indikation für Beratung und Therapie sein.

1.1.2 Ätiologie

Für das Stottern gibt es keine endgültige Erklärung für seine Entstehung. „Eine Reihe von Theorien wurde über die Verursachung des Stotterns aufgestellt. Zu jeder Theorie scheinen sich stotternde Personen finden zu lassen, für die diese Theorie plausibel ist. Umgekehrt kann bislang keine Theorie für sich beanspruchen, auf alle stotternden Personen zuzutreffen“ (Natke, 2005, S. 67). Tab. 1 versucht die zentralen Erklärungen der gängigsten Theorietypen kompakt darzustellen:

Tab. 1: Überblick der gängigen Theoriegruppen zur ätiologischen Erklärung des Stotterns

Theoriegruppen			
Stottern als neurotische Reaktion	Lerntheorien	Breakdown-Theorien	Multikausale- bzw. faktorielle Theorien
Unbewusste Bedürfnisse = Stottersymptome haben einen Bezug zu unbewussten Zielen	Stottern wird erlernt: Operantes oder klassisches und operantes Konditionieren in Zusammenspiel mit kognitiven Bewertungsprozessen	Zusammenbruch des hochautomatisierten Sprechvermögens bei Überlastung des Systems im Sinne eines „angelegten“ neurophysiologischen Defizits	Disponierende-auslösende-aufrechterhaltende Bedingungen = Ideografische Sichtweise

Traditionell sind Lerntheorien (antizipierende Anstrengungsreaktion im Sinne von ‚Stottern ist der Versuch, nicht zu stottern‘), Breakdowntheorien (Mühen in der hoch automatisierten Planung und Realisation des Sprechens), multifaktorielle Theorien (Aufteilung in disponierende, auslösende und aufrechterhaltende Bedingungen) sowie ‚Stottern als neurotische Reaktion‘ (mit tiefenpsychologischen Wurzeln) als Erklärungsansätze der Entität Stottern aufzuführen (vgl. Natke, 2005).

Das komplizierte wechselseitig-abhängige Bedingungsgefüge entspricht am ehesten den multikausalen Theorien und verbindet lerntheoretische Erklärung mit denen der Breakdown-Theorien. Die tiefenpsychologische Erklärung tritt dagegen in aktuellen Veröffentlichungen zur Ätiologie des Stotterns in den Hintergrund (vgl. Natke, 2005, S.67f).

Wichtig in der Ätiologiediskussion ist die Unterscheidung von disponierenden, auslösenden sowie aufrechterhaltenden Bedingungen (vgl. Johannsen, 2010). Die Ursache des Stotterns wird mit einer genetischen Veränderung, die zu einer zentralnervösen Störung des Sprechens z. B. in Form einer neuromotorischen Koordinationsstörung führt, erklärt. Die *genetische Disposition* (Prägung) für Stottern (Epigenetik) wurde mittels verhaltensgenetischer Untersuchungen (Zwillings-, Adoptions-, Familienstudien) und DNA-Analysen aufgezeigt. Genetiker postulieren, dass beim aktuellen Forschungsstand inzwischen „bei 70 % der Gesamtgruppe der stotternden Kinder“ das Auftreten von Stottersymptomen genetisch erklärbar ist (Sandrieser, 2017, S. 377; vgl. auch Überblick bei Neumann et al., 2017, S. 31 ff). Als *auslösende Bedingungen* werden Ereignisse, Umstände benannt, die dem Erstauftreten von Stottern kurz vorhergehen. Dies kann ein institutioneller Übergang sein (Übertritt in den Kindergarten, in die Schule), eine für das Kind herausfordernde familiäre Veränderung (Geburt eines Geschwisters, Wohnortwechsel) oder in seltenen Fällen eine traumatische Erfahrung. Diese Bedingungen sind nicht als die eigentliche Ursache zu verstehen, sondern als der ‚Tropfen, der das Fass zum Überlaufen‘ bringt.

Aufrechterhaltende Faktoren sind dann emotionale, mentale und kommunikative Verunsicherung des Kindes selber oder seiner Umgebung sowie damit einhergehende veränderte Reaktionsweisen der Umgebung.



Abb 1: Zeitliche Ordnung der Entstehung und Aufrechterhaltung von Stottern

Diese zeitliche Abfolge differenziert die Entstehung und die Aufrechterhaltung von Redeunflüssigkeiten über einen längeren Zeitraum hinweg. Hingegen versucht das folgende 3-Faktoren-Modell *Momente des Stotterns* zu erklären. Dabei wird argumentiert, dass alle ursächlichen Faktoren in jedem Moment des Stotterns wirksam sein müssen.

Die Erkenntnis, dass ein neurologisches Defizit in der Verarbeitung von gesprochener Sprache Grundvoraussetzung für das Auftreten von Stottern ist, wurde von Packmann und Attanasio (2010) mit dem bekannten Anforderungs- und Kapazitätenmodell (Starkweather, 1997) verknüpft und es resultierte das *3-Faktoren-Modell* (Packmann & Attanasio, 2010; Packmann, 2012; Sandrieser, 2017). In diesem Modell wird nachvollziehbar versucht, die Momente des Stotterns zu erklären und Mechanismen für verschiedene Behandlungsausrichtungen abzuleiten.

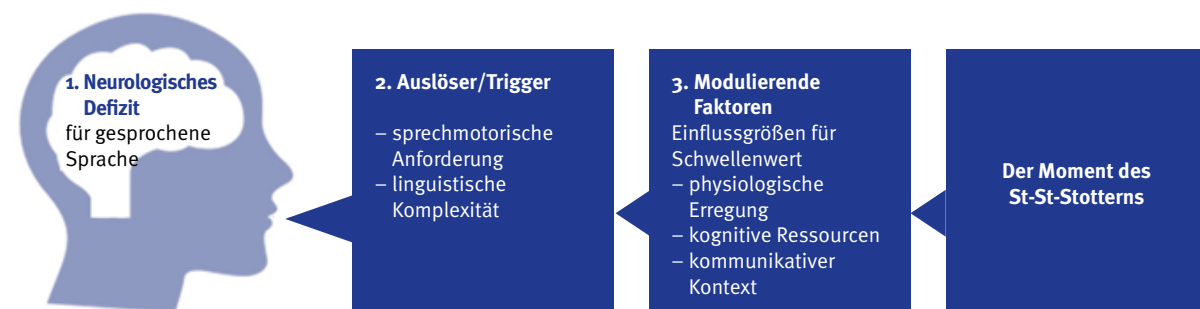


Abb 2: 3-Faktoren-Modell in Anlehnung an Packman (2012)

Das oben beschriebene neurologische Defizit (bspw. die gestörte oder verzögerte Entwicklung von Myelin) ist der Ausgangspunkt des Modells (Packman, 2012). Stotterer stottern aber nicht bei jeder Silbe und auch nicht in jeder Situation – es spielen im Moment des Stotterns Auslöser (=2.) und individuelle modulierende Faktoren (=3.) eine gewichtige Rolle.

Als *Auslöser* können erhöhte artikulatorische Anforderungen sowie die linguistische Komplexität wirken. Vereinfacht gilt der Grundsatz, je komplizierter die Äußerung, desto mehr Aufwand wird von unserem Gehirn und der anzusteuern, ausführenden Artikulation verlangt, um diese zu produzieren. Bloodstein & Bernstein Ratner (2008) weisen in einer Studie darauf hin, dass Stottern tendenziell in Äußerungen auftritt, die sprachlich komplexer sind. Im 3-Faktoren-Modell wird der Auslösemechanismus u.a. durch körperliche Faktoren moduliert. So kann das Ausmaß der physiologischen Erregung bei einem Individuum die Schwelle maßgeblich verändern, ab der ein Moment des Stotterns ausgelöst wird. Je höher die Erregung desto niedriger die Schwelle für unflüssiges Sprechen. Der Körper hat die Bereitschaft, auf belastende innere und äußere Einflüsse zu reagieren. Die Erregung ist u.a. bedingt durch Leistungsangst, Angst vor negativer Bewertung oder Aufregung. Bezogen auf die Kommunikationsstörung Stottern weiß man, dass Stottern situationsabhängig ist und die stotternde Schwere je nach (kommunikativen) Kontext variieren kann. Die modulierenden Faktoren sind für jedes Individuum einzigartig und sind „neben dem Temperament auch durch Erfahrungen (z. B. Hänself), Einstellungen (z. B. negative Bewertung des Stotterns) und die aktuelle Stottersymptomatik selbst beeinflusst“ (Sandrieser 2017, S. 400). Modulierende Faktoren können innerhalb von Individuen als Hauptfaktor für die Variabilität des Stotterns über kommunikative Kontexte hinweg angesehen werden. Sie stellen eine Interaktion zwischen dem körperlichen Aufbau des Betroffenen, den Lebenserfahrungen und seiner Umgebung dar.

1.2 Beginnendes Stottern

1.2.1 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 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 Sprach-Handlungs-Planung. 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 Auftretenshäufigkeit, 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 und es sind spannungsreiche ganzkörperliche Mitbewegungen sowie eine spannungsreiche Stimmgebung zu beobachten (vgl. u.a. Natke, 2005; Sandrieser, 2004; Sandrieser & Schneider, 2015). Nichtstotternde 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 nichtstotternden Kindern (Sandrieser, 2004, S.136).

Die folgende Tab. 2 fasst in Form einer Gegenüberstellung normale und symptomatische Sprechunflüssigkeiten zusammen.

Tab. 2: Unterscheidung normale versus symptomatische Sprechunflüssigkeiten

Beobachtbare Symptome	Normale, entwicklungsbedingte Sprechunflüssigkeiten	Verstärkte Sprechunflüssigkeiten – beginnendes Stottern
Wiederholungen		
Wiederholungen einsilbiger Wörter («Ich – ich – ich muss dir was erzählen!»)	•	
Wiederholungen mehrsilbiger Wörter («Ich hatte – hatte keine Angst beim Zahnarzt.»)	•	
Satzteilwiederholungen («Ich kann – ich kann das nicht so gut.»)	•	
Lautwiederholungen («A-a-a-aber ... ich bin jetzt dran!»)		•
Häufig mehr als zwei Silbenwiederholungen («Ich ma – ma – ma – ma – mache die Tür auf.»)		•
Dehnungen		
Vokaldehnungen von mehr als einer Sekunde Dauer («Ich kaaaaaaaaann schon Rad fahren!»)		•
Pausen		
Stille Pausen zwischen einzelnen Wörtern/Satzteilen zur inhaltlichen Planung («Wo ist kannst du mir das geben?»)	•	
Blockierungen		
mit sichtbarer Anstrengung, Mund oder Gesicht sind beim Hängenbleiben verspannt		•
Begleitsymptome		
Vermeidung von Wörtern («Der wohnt in der Hundehü – hü – haus.»)		•
Verspannungen von Lippen, Gesicht, Hals in Verbindung mit der Unflüssigkeit		•
Mitbewegungen des Körpers in Verbindung mit der Unflüssigkeit		•

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. Diese Begrifflichkeit macht großen Sinn, da Kinder in diesem Stadium Entwicklungen in beide Richtungen (Stottern oder normaler Redefluss) vollziehen können.

Als charakteristische Merkmale von *Borderline Stottern* werden folgende Items genannt (ebd.):

- Mehr als 10 Unflüssigkeiten pro 100 Wörter
- Frequenz der Wiederholungen oft über 2 (Bsp. Ba – Ba – Ball)
- Mehr Wiederholungen und Dehnungen als Verbesserungen oder unvollständige Phrasen
- Unflüssigkeiten sind locker und spannungsarm
- Kinder mit Borderline Stottern reagieren selten auf ihre Unflüssigkeiten

Der Typus *Borderline Stottern* 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 Stottereignisses ab. Guitar 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 ist 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. stotterympptomatischen Unflüssigkeiten, wobei die Autoren in der Regel auf den Grenzbereich zwischen diesen beiden Kategorien hinweisen, ihn aber nicht mit einer extra Kategorie explizit bezeichnen. Eine Ausnahme ist Bosshardt (2008), der ausdrücklich von „Sprechunflüssigkeiten bei Grenzfällen“ (S. 24) spricht oder Natke (2005), der das *borderline stuttering* als „grenzwertiges Stottern“ (S. 50) übersetzt.



Abb. 3: Kontinuum der Sprechunflüssigkeiten

1.2.2 Beginn, Verlauf und Remission

Der Beginn des Stotterns in der Kindheit zeigt eine altersmäßige Streuung zwischen dem Sprechbeginn (vor dem 2. Lebensjahr) und dem 5. Lebensjahr. Die linksschiefe Verteilung zwischen diesen beiden Polen hat ihren Höchstwert bei 2,6-jährigen Kindern (Yairi & Ambrose, 2005, S. 54).

Es gibt sowohl einen plötzlichen Beginn als auch einen eher allmählichen, wobei die Mehrheit der Kinder spontan und eher plötzlich Stottersymptome zeigen oder eine schnelle graduelle Entwicklung innerhalb von wenigen Wochen aufweisen. Der langsam-allmähliche Stotterbeginn scheint in einem Zusammenhang mit einer ungewöhnlich schnellen Wortschatzentwicklung zu stehen.

Sowohl zu Beginn als auch im weiteren Verlauf gibt es einen klaren Geschlechtereffekt: Es sind mehr Jungen als Mädchen betroffen und zwar im Verhältnis von ca. 3:1 (Natke, 2005, S. 11).

Es scheint auch einen Zusammenhang zwischen dem Stotterbeginn und subjektiv empfundenen Belastungssituationen der betroffenen Kinder zu geben. Bei 80 % der Kinder hat man Belastungen emotionaler, physischer oder sprachlicher Natur gefunden. Solche Belastungen werden auch zur Erklärung von Aufrechterhaltung, Chronifizierung, Variabilität und Intensität der Symptomatik herangezogen (Starkweather & Gottwald, 1990). Allerdings sind die meisten der betroffenen Kinder sich ihres Stotterns am Beginn nicht bewusst und es stört sie zunächst auch nicht. Dies ändert sich jedoch für viele betroffenen Kinder mit der anhaltenden Dauer der Symptomatik. Bloodstein (1960) nennt Phasen der emotionalen Reaktionen auf Stottern, die auch schon in der frühen Kindheit einsetzen können: von vereinzelt Reaktionen auf das Stottern ohne emotionale Reaktion geht es über Reaktionen wie z. B. Ärger, Wut oder Frustration hin zu Angst oder gar Panik.

Lange Zeit wurde das Verhalten der Bezugspersonen als alleiniger auslösender Faktor betrachtet und beherrschte die Meinung in der Fachwelt durch die diagnosogene Theorie des Stotterns (Johnson, 1955). Heute geht man davon aus, dass sich das Verhalten der Eltern stotternder Kinder nicht von dem Verhalten unterscheidet, welches andere Eltern ihren nicht-stotternden Kindern gegenüber zeigen (Yairi & Ambrose, 2005, S. 282).

Guitar (2014, S. 112ff) schlägt Entwicklungsstadien des beginnenden Stotterns vor, die eine Kombination aus altersabhängigen und symptombezogenen Kriterien darstellen. Er begründet dies mit der Nützlichkeit für mögliche Interventionen, die sich sowohl am Kindesalter als auch an der Qualität der Symptomatik orientieren sollen.

Für die Stufe der „normal disfluency“ gibt Guitar (2014, S. 115) eine Lebensalterszeitspanne zwischen 1,5 und 6 Jahren an, für das „borderline stuttering“ von 1,5 bis 3,5 Jahren, für das „beginning stuttering“ von 3,5 bis 6 Jahren, für das „intermediate stuttering“ von 6 bis 13 Jahren und für das „advanced stuttering“ ab 14 Jahren

Yairi & Ambrose (2005) kritisieren die von Guitar vorgeschlagenen Entwicklungsstufen als zu vereinfachend und auch die Diskussion zur Prävalenz und Inzidenz des Stotterns wird kontrovers geführt. Yairi und Ambrose (2013) gehen davon aus, dass die bisher genannte Prozentzahl von 5 % für die Inzidenz höher angesetzt werden muss (nämlich zwischen 7 % und 9 %) und die Prävalenz für Stottern i. a. wahrscheinlich niedriger als 1 % der Gesamtbevölkerung ist. Das

macht das Problem der Identifikation von Fällen des beginnendem Stotterns mit Beratungs- und Diagnostikbedarf nicht kleiner, weil die Anzahl der Risikokinder aufgrund dieser Zahlen höher ist, als bisher angenommen (entspricht den oben formulierten 7 bis 9%), auch wenn die Remissionsrate größer ist, als bisher angenommen: Ca. 90 % aller Kinder, die Symptome des frühkindlichen Stotterns zeigen, verlieren diese Symptomatik wieder (Yairi & Ambrose, 2005, S. 152f und Yairi & Ambrose, 2013), wobei hier nicht zwischen spontaner Remission und Remission aufgrund pädagogisch-therapeutischer Maßnahmen unterschieden wird. Brocklehurst (2013) kritisiert das Zustandekommen dieser Zahlen ganz grundsätzlich, weil er die unterschiedlichen Definitionen als verantwortlich für die Daten betrachtet.

Ohne auf diese Diskussion näher einzugehen sind die Faktoren, welche eine Remission wahrscheinlicher bzw. unwahrscheinlicher machen als Prognosefaktoren interessant:

Mädchen remittieren häufiger als Jungen. Klar scheint zu sein, dass solche Kinder, die eine anhaltende oder zunehmende Quantität und Intensität der Stottersymptomatik aufweisen, weniger remittieren, als solche Kinder, deren Symptomatik nach dem ersten Auftreten innerhalb der ersten zwölf Monate kontinuierlich abnimmt. Im Allgemeinen ist bei langanhaltender Dauer der Symptomatik (länger als zwei Jahre) eine Remission unwahrscheinlicher. Auch der späte Beginn (nach dem dritten Geburtstag) geht mit einem erhöhten Risiko der Chronifizierung einher. Wenn es chronisch stotternde Familienangehörige gibt, fällt die Remissionswahrscheinlichkeit ebenfalls. Schlechte phonologische Fähigkeiten gehen wahrscheinlich einher mit einer verminderten Remission. Auf den ersten Blick dazu widersprüchlich erscheint der Zusammenhang zwischen überdurchschnittlichen sprachlichen Fähigkeiten und erhöhtem Risiko von persistierendem Stottern. Allerdings bezieht sich die Überdurchschnittlichkeit nicht auf die Aussprache oder phonologische Fähigkeiten, sondern auf die Differenziertheit im Wortschatz, der erhöhten Mittleren Äußerungslänge (MLU) sowie der auditiven Verstehensleistung im Vergleich zu Ihren unauffälligen Altersgenossen (Yairi & Ambrose, 2005, S. 238ff). Ältere Studien (z. B. Andrews, Craig, Feyer, Hoddinott, Howie & Neilson, 1983) zeigen dagegen, dass stotternde Kinder häufiger von Sprachentwicklungsverzögerungen betroffen sind. Insbesondere artikulatorische Probleme seien bei stotternden Kindern häufiger zu beobachten im Vergleich zu den nichtstotternden Altersgenossen (z. B. Bloodstein, 1958). Yairi & Ambrose (2005, S. 282) gehen davon aus, dass das sprechmotorische System von stotternden Kindern weniger stabil und flexibel ist als das der unauffälligen Altersgenossen. Im Überblick ergeben die Studien zum Zusammenhang zwischen Sprachentwicklung und beginnendem Stottern ein uneinheitliches Bild (vgl. Natke, 2005, S. 56), wobei der Zusammenhang zwischen phonetisch-phonologischen Defiziten und beginnendem Stottern einigermaßen gesichert zu sein scheint.

Tab. 3: Übersichtstabelle Remissionswahrscheinlichkeit

Remissionswahrscheinlichkeit	
höher	geringer
Mädchen	Jungen
Quantitative Verringerung der Symptome innerhalb der ersten 12 Monate nach Erstauftreten	Anhaltende, zunehmende Quantität der Symptome
Erstauftreten vor dem 3. Lebensjahr	Zunehmende Intensität der Symptomatik
Kein chronisch stotterndes Familienmitglied	Langanhaltende Symptomatik (länger als zwei Jahre)
Keine artikulatorischen Auffälligkeiten	Phonetisch-phonologische Defizite

1.3 Identifikation des beginnenden Stotterns

Es werden in der Fachliteratur verschiedene Vorschläge gemacht, um das beginnende Stottern zu identifizieren. Sie unterscheiden sich jedoch hinsichtlich der Gewichtung einzelner Bereiche der Symptomatik. So werden z. B. manchmal Auffälligkeiten der Atmung ganz explizit als Beobachtungskriterium aufgenommen (Ochsenkühn & Thiel, 2005) oder die Häufigkeit der Symptomatik hervorgehoben bzw. einzelne Aspekte unterschiedlich bepunktet (Sandrieser & Schneider, 2015 oder Kuckenbergh & Lattermann, 2011) und damit gewichtet.

Auch hinsichtlich der Zielgruppen und der Durchführungsart solcher Entscheidungshilfen gibt es Unterschiede. Manche richten sich an Eltern und können als paper-pencil bzw. Online Version genutzt werden (Sandrieser & Schneider, 2015) andere sind eher für Therapeutinnen gedacht

(Lattermann, 2011 oder Ochsenkühn & Thiel, 2005) oder explizit für Kinderärztinnen (Kuckenberg & Lattermann, 2011).

Auch die meisten Ratgeber zum Thema kindliches Stottern listen die Primär- und teilweise auch die Sekundärsymptomatik des Stotterns auf, um den Ratsuchenden eine Veranschaulichung des Phänomens zu geben (z. B. Hansen & Iven, 2004, Brügge & Mohs, 2006 oder Schneider, 2013).

Eines der bekanntesten Vorschläge ist die Screening Liste Stottern (SLS). Die SLS ist ein Fragebogen, der Eltern vorgelegt wird und zu einer Empfehlung hinsichtlich des weiteren Vorgehens führt (Sandrieser & Schneider, 2015, Original von Riley & Riley, 1989).

Ab 2010 wurde im deutschsprachigen Raum an der Schaffung eines Screeninginstrumentes zur Identifikation des beginnenden Stotterns gearbeitet. Zwei Forschergruppen haben unabhängig voneinander dieses Projekt betrieben. Das *Bochum-Aachener Stotterscreening (BASS)* (Neumann, Euler & Schneider, 2014) war als Universalscreening konzipiert und sollte flächendeckend bei der Vorsorgeuntersuchung durch Kinderärzte eingesetzt werden. Der *RedeflussKompass 2.0* (Braun et al. 2011, Kohler 2015b) war als Risikoscreening gedacht und sollte nur bei Verdacht auf Stottern eingesetzt werden. Beide wurden hinsichtlich ihrer Gütekriterien untersucht und beide konnten aus unterschiedlichen Gründen nicht überzeugen (vgl. Kohler, 2017). Für den RedeflussKompass 2.0 (Braun, Baumann & Boltshauser, 2011) wurde insbesondere die fehlende Passung an das Zielpublikum (Laien vs. Fachpersonen) als Grund der unbefriedigenden Güte identifiziert. Die Schlussfolgerungen aus der Evaluation des RedeflussKompasses 2.0 führten zu einer Differenzierung des Instrumentariums. Da diese Differenzierung der Hauptgegenstand vorliegender Untersuchung darstellt, soll sie im Folgenden genauer erläutert werden. Es werden die beiden Instrumente RedeflussKompass 3.0 und StotterKompass beschrieben.

1.4 Konzeption von RedeflussKompass 3.0 und StotterKompass

Der RedeflussKompass 3.0 (Braun & Kohler, 2014) in seiner Weiterentwicklung zur Version 3.0 (= RfK 3.0) wurde an ein Laienpublikum angepasst, welches eine wichtige Rolle bei der Früherkennung des Stotterns spielen kann. Das sind insbesondere die Bezugspersonen der betroffenen Kinder wie Eltern, (Früh-)Erzieherinnen, Spielgruppenleiterinnen etc.. Sie sollten eine für diese Laiengruppe verständliche Entscheidungshilfe erhalten. Der RfK 3.0 soll im Sinne einer Früherkennung helfen zu klären, ob eine weitere Ersterfassung und Erstberatung durch eine Fachperson notwendig ist. Für diese Aufgabe wurde der StotterKompass (Kohler & Braun, 2014) geschaffen. Er steht den Fachleuten (z. B. Logopädinnen, Sprachtherapeutinnen, Sprachheilpädagoginnen etc.) zur Verfügung, die zusammen mit den Bezugspersonen die Entscheidung für eine intensivere Diagnostik, weitere Beratung oder die Notwendigkeit einer Therapie treffen müssen. Abb. 4 veranschaulicht den so konzipierten Entscheidungsablauf.



Abb. 4: Entscheidungsablauf und Zielgruppen von RedeflussKompass und StotterKompass

Beide Instrumente¹ gleichen einem Fragebogen. Es werden Fragen gestellt und Beispiele gegeben, welche in einem inhaltlichen Zusammenhang zum beginnenden Stottern stehen. Die daraus ab-

¹ Die beste Nachvollziehbarkeit erreicht man, wenn man die Instrumente im Original vorliegen hat. Aus Platzgründen können die Instrumente hier nicht vollständig abgebildet werden. Beide Kompassse können über der Website der Hochschule für Heilpädagogik (HfH) kostenlos heruntergeladen werden. Den schnellsten Zugang hat man durch die Eingabe *HfH Früherkennung* bei der Suchmaschine von Google. Der oberste Treffer ist eine Website der Hochschule auf deren linken Seite die Kompassse zum Download bereitstehen. Die Vorlage der Kompassse ist auch für die folgenden Ausführungen des vorliegenden Artikels sehr nützlich.

geleiteten Antworten sollen angekreuzt werden und stellen eine Skala der Häufigkeit oder der Intensität des Zutreffens dar. Z. B. wird die Frage gestellt *Wiederholt das Kind einzelne Wörter oder Wortteile?* Ein Beispiel dazu ist *Ich ma ma ma mache die Türe auf.* Die anzukreuzende Antwort ist, ob diese Unflüssigkeiten *nie, selten, manchmal* oder *häufig* vorkommen.

Der Einsatz der beiden Instrumente RfK 3.0 und SK ist komplementär gedacht. Die Bezugsperson bearbeitet zunächst den RedeflussKompass 3.0. Falls das Ergebnis für eine Ersterfassung und eventuelle Erstberatung spricht, soll ein Gespräch mit einer Fachperson unter Vorlage des StotterKompass und des ausgefüllten RedeflussKompasses die Grundlage für Ersterfassung und Erstberatung sein. Unter komplementär ist dabei folgendes zu verstehen: Da alle Items des RedeflussKompasses 3.0 auch im StotterKompass wiederauftauchen, haben die Fachpersonen die Möglichkeit, diese Inhalte zu validieren, indem sie nachfragen, vertiefende Beispiele geben (insbesondere durch Vormachen der Symptomatik!) und/oder die Symptomatik durch die Bezugspersonen vormachen lassen. Dieses Vorgehen entspricht einer *kommunikativen Validierung* der Items des RedeflussKompasses 3.0.

Der StotterKompass ist zudem differenzierter in seiner Funktion der Ersterfassung. Es werden dort weitere Items aufgeführt, die über das rudimentäre Niveau des RedeflussKompasses 3.0 hinausgehen. Diese Items erfassen die Unflüssigkeiten selbst und die vermeintliche Sekundärsymptomatik genauer, ohne allerdings an eine intensive und hochdifferenzierte logopädische Diagnostik heranzureichen.

Zudem bietet der StotterKompass „allgemeine Regeln“ für pointierte Zusammenhangsaussagen innerhalb von Symptomgruppen an, die für die Erstberatungstätigkeit benutzt werden können. Die allgemeinen Regeln sind die auf den Punkt gebrachten empirisch gesicherten Erkenntnisse zum beginnenden Stottern (vgl. 1.2.1 bis 1.2.2). Als Beispiel soll an dieser Stelle die allgemeine Regel zum Zusammenhang zwischen der Primärsymptomatik Blockade und dem Schweregrad des Stotterns gegeben werden: „Je länger die Blocks dauern, je häufiger sie auftreten und je höher der Grad der Anspannung dabei, desto höher der Schweregrad des Stotterns“ (Kohler & Braun, 2014, S.4).

Das Setting zum Einsatz des StotterKompasses ist niederschwellig gedacht. Das heißt, es wird angenommen, dass zu diesem Zeitpunkt die Befragung als Methode der Informationsgewinnung und das Gespräch als zentrale Methode für die Erstberatung (Bachmair, Faber, Hennig, Kolb & Willig, 2007) ausreicht.

Um RedeflussKompass 3.0 und StotterKompass im Detail in ihrer Entsprechung und Unterschiedlichkeit nachvollziehen zu können, lohnt es sich, sie zu vergleichen.

Tab. 4 zeigt einen tabellarischen Vergleich der beiden Instrumente. Die farbig markierten Begriffe kennzeichnen den höheren inhaltlichen Differenzierungsgrad des StotterKompasses und damit den Unterschied zum RedeflussKompass 3.0.

Tab. 4: Darstellung der Unterschiede von RedeflussKompass 3.0 und StotterKompass

RedeflussKompass 3.0	StotterKompass	Allgemeine Regeln, Erklärungen
Wiederholungen: Wörter Wortteile	Wiederholungen: Satzteile, Wörter, Wortteile (Laute, Silben), Rhythmus, Tonhöhe	
Dehnungen: Wörter, Wortteile	Dehnungen mit Zeitangabe: Wörter, Wortteile. Rhythmus, Tonhöhe	
Pausen (Blockierungen): Wörter	Pausen (Blockierungen): Wörter	
Dauer	Dauer und Beginn	
Verlauf	Verlauf	
Reaktion des Kindes	Reaktion des Kindes: Vermeidung, Blickkontakt, Sprechweise	
Reaktion der Umwelt	Reaktion der Umwelt: Hilflosigkeit, Ärger, Aggression	
	Begleiterscheinungen: Mitbewegungen, Stimme, allg. Anspannung	
	Risikofaktoren: Familiäre Disposition, Sprachentwicklung, Geschlecht	

Durch den StotterKompass soll vor einer intensiven Diagnostik oder weiterführenden Beratung und Therapie eine Ersterfassung mit Erstberatungsmöglichkeit geschaffen werden. Inhaltlich

Der direkte Link zur Website ist <https://www.hfh.ch/de/unser-service/expertenwissen-online/logopaedie-und-praevention/frueherkennung>

referiert diese Zwischenstufe mit dem in der Theorie des beginnenden Stotterns formulierten „borderline stuttering“ (Guitar, 2014, S. 120), welche eine fehlende Eindeutigkeit und damit einhergehende Unsicherheit im Umgang mit normalen vs. stottersymptomatischen Unflüssigkeiten proklamiert (vgl. oben unter 1.2.1 bzw. Bosshardt, 2008 oder Guitar, 2014).

2 Fragestellung und Zielsetzung

Der Einsatz von RedeflussKompass 3.0 und StotterKompass wurde in der Praxis begleitet, systematisch dokumentiert und einer eingehenden Prüfung unterzogen.

Es wurden die Empfehlungen des Redeflusskompasses 3.0 mit dem wahren Bedarf an Früherfassung und Erstberatung konfrontiert und so die Gütekriterien Sensitivität, Spezifität und Validität quantitativ bestimmt. Die Empfehlungen des StotterKompasses wurden mit dem wahren Bedarf an weiterer Beratung, Diagnostik und/oder Therapie konfrontiert, so dass ebenfalls seine Sensitivität, Spezifität und Validität sowie darüber hinaus durch eine unabhängige Zweidurchführung dessen Interrater-Reliabilität berechnet werden konnte.

Die weitere Analyse des Instrumenteneinsatzes lieferte Erkenntnisse zur Praxis der Durchführung. Es fanden sich so auch Erklärungen zum Zustandekommen der Gütekriterien. Nicht zuletzt gab es Hinweise zur Optimierung beider Instrumente und deren Einsatz.

Die übergeordneten Fragestellungen lauteten im Einzelnen wie folgt:

- Welche Sensitivität, Spezifität und Validität weisen der RedeflussKompass 3.0 und der StotterKompass auf?
- Welche Interrater-Reliabilität lässt sich für den StotterKompass bestimmen?
- Diese beiden Fragen zielen auf die Güte der Entscheidungen ab, die mit dem Instrumentarium getroffen wurde.
- Die folgenden Fragen zielen auf die Art und Weise ab, mit der die Entscheidungen getroffen wurden und wie die Ersterfassung und die Erstberatung gelang.
- Welche Erfahrungen machen Bezugspersonen bei dem Einsatz des RedeflussKompasses 3.0?
- Welche Erfahrungen machen Fachpersonen bei dem Einsatz des StotterKompasses?
- Welches Potenzial zur (Erst-)Beratung hat der kombinierte Instrumenteneinsatz?
- Welche Schlussfolgerungen lassen sich aus der Berechnung der Gütekriterien und den Erfahrungen für die Optimierung der beiden Instrumente und deren Einsatz ziehen?

Die folgende Abb. 5 soll das Forschungsdesign kompakt veranschaulichen. Mit den quantitativen Ergebnissen der Durchführungen wurden die Güte Merkmale bestimmt. Die Pfeile, die von der grün unterlegten, qualitativen Dokumentenanalyse ausgehen, stehen für die Erfahrungen der Anwender und deren Potenzial zur Erklärung und Optimierung des Instrumenteneinsatzes.

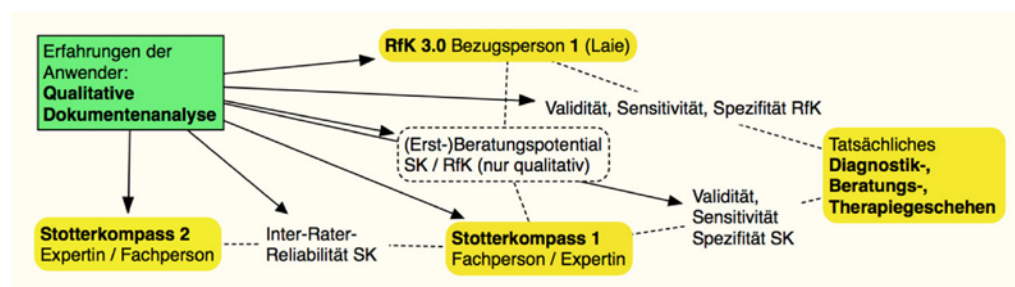


Abb. 5: Grobes Forschungsdesign zur kompakten Illustration von Fragestellungen und Zielsetzung

3 Methoden

Der methodische Ansatz zur Instrumentenprüfung kann als mixed-method Design (Kuckartz, 2014) bezeichnet werden. Es kamen sowohl quantitative als auch qualitative Methoden zum Einsatz. Die Bestimmung der Kennwerte für die Gütekriterien Sensitivität, Spezifität und der Interrater-Reliabilität des StotterKompasses folgte einer quantitativen Strategie. Die systematische Analyse der Erfahrungen der Anwender geschah mit einem schwerpunktmäßig qualitativen Ansatz.

Als übergeordnetes Paradigma dieses Forschungsprojektes kann der Begriff „Praxisforschung“ verwandt werden (Moser, 2001 & 2015). Im Rahmen der Praxisforschung sind alle Methoden der Sozialwissenschaften gebräuchlich, es wird jedoch besonderen Wert auf ein Setting gelegt, welches dem berufspraktischen Alltag nahe ist. Insofern ist die Praxisforschung dem Begriff Feldforschung nahe und bringt eine hohe externe Validität mit sich. Der Nachteil gegenüber einem Setting, welches eher einer Laborforschung nahesteht, ist die verminderte Variablenkontrolle und damit die Gefahr einer geringen internen Validität. Konkret heißt das, dass die kooperierenden Praxisstellen mit unterschiedlicher Intensität am Forschungsprojekt beteiligt waren und insofern die Nachvollziehbarkeit pro Fall eine unterschiedliche Güte aufweist. Dieser Nachteil galt vor allem für den qualitativen Teil des Forschungsprojektes und wurde versucht durch ein hohes Maß an Triangulation (Flick, 2008) aufzufangen. Es wurde daher eine „breit“ angelegte Datenerhebung und -auswertung bevorzugt. Das heißt, dass eine Vielzahl von Erhebungs- und Auswertungsmethoden in unterschiedlichem Differenzierungsgrad eingesetzt wurden. Die entsprechenden methodischen Vorgehensweisen werden unter den Punkten 3.2 und 3.3 beschrieben.

3.1 Stichprobenkonstruktion

Für die Durchführung der Kompass wurden Logopädinnen und Sprachtherapeutinnen aus der Praxis als Kooperationspartner gewonnen. Sie sind die potenziellen Anwenderinnen des StotterKompasses und haben Fälle für dessen Einsatz ausgewählt. Diese Fachpersonen waren insofern auch die Schnittstelle zu den Bezugspersonen, welche den RedeflussKompass 3.0 bearbeitet haben. Falls die Bereitschaft zur Kooperation auch bei den ausgewählten Bezugspersonen bestand, lag eine Situation vor, die der Realität in der Praxis weitestgehend entspricht.

Außerdem wurde versucht, Bezugspersonen, die sich in der Therapie-Lehr-Praxis der Hochschule für Heilpädagogik in Zürich mit dem Ansinnen auf Abklärung hinsichtlich beginnenden Stotterns gemeldet hatten, als Kooperationspartner zu gewinnen. Bei der Bereitschaft zur Kooperation boten wir den Termin für die Ersterfassung und Erstberatung für diese Personengruppe kostenlos an. Auch diese Situation bildet die Realität nahe zu 100 % ab.

Auf diese Art und Weise kamen bei 86 Fällen eine Durchführung mit den Kompassen zustande. Es wurden davon 60 in die Datenanalyse eingeschlossen:

3.1.1 Einschlusskriterien für die Stichprobe der Fälle

Must to have:

- I. Für diese N=60 Fälle war auf irgendeine Art und Weise bei den Kindern ein Verdacht auf Stottern geäußert worden. Die Stärke des Verdachts variierte von schwach bis stark und wurde von den verschiedensten Personen aus dem engen oder weiten sozialen Umfeld des Kindes geäußert (Eltern, Großeltern, Erzieher, Lehrer, Verwandte, Bekannte, Logopädin etc.). Die Stichprobe der Fälle entspricht dementsprechend den Kriterien für die Anwendung eines Risikoscreenings.
- II. Es konnte ein „Goldstandard“ ermittelt werden: Der tatsächliche Beratungs-, Diagnostik- oder Therapiebedarf konnte sicher bei den ausgewählten Fällen nachvollzogen werden und wurde in drei Stufen operationalisiert: kein Bedarf – Entwicklungsbeobachtung – Bedarf.
- III. Das Alter der Kinder lag zum Zeitpunkt der Durchführung zwischen 2,6 und 6,2 Jahren (siehe Abb. 6). Im Durchschnitt waren die Kinder 4,5 Jahre alt. Die Altersspanne entspricht der relevanten Zielgruppe für beginnendes Stottern. Auch wenn der Mittelwert von 4,5 Jahre dem Mittelwert aus der Grundlagenforschung (=2,6 Jahre) nicht entspricht, erscheint er doch hinsichtlich dem Altersdurchschnitt, mit dem Kinder einem Risikoscreening für beginnendes Stottern unterzogen werden, realistisch zu sein.

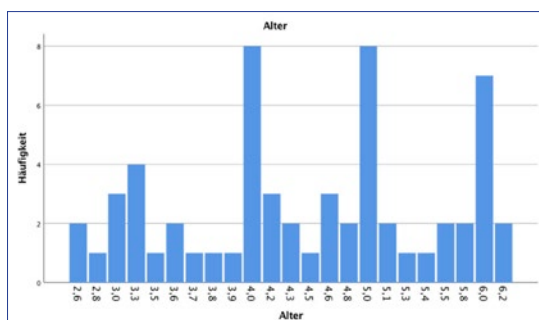


Abb. 6: Altersverteilung der Stichprobe der Kinder

IV. Es lag die Durchführung des RedeflussKompasses 3.0 (in 54 Fällen) oder die Durchführung des StotterKompasses (in 57 Fällen) vor. In den meisten Fällen (=52) konnte sowohl die Durchführung des RfK 3.0 als auch des StotterKompasses realisiert werden.

Nice to have:

V. Zwar wurde eine unabhängige Zweitudurchführung des RedeflussKompasses i.d.R. versucht, jedoch gelang dies in nur fünf Fällen wirklich unabhängig. Es bearbeiteten nicht selten zwei Bezugspersonen gemeinsam den RedeflussKompass 3.0 (meist Mutter und Vater), jedoch nicht unabhängig voneinander. Diese gemeinsame Bearbeitung wurde dann zum Gegenstand der qualitativen Analyse, wegen der fehlenden Unabhängigkeit jedoch nicht zum Gegenstand der quantitativen Bestimmung von Gütekriterien.

VI. Ebenfalls *Nice to have* war eine unabhängige Zweitudurchführung des StotterKompasses. Dies gelang immerhin in 17 Fällen. Allerdings ist die Zweitudurchführung in der Regel in einem anderen Setting durchgeführt worden als die Erstdurchführung. Die Erstdurchführung wurde i.d.R. face-to-face (50 von 57) durchgeführt. Die Zweitudurchführung dagegen mehrheitlich am Telefon (11 von 17). Da alle Zweitudurchführungen am Telefon von einem Experten durchgeführt wurden, kann das Telefonsetting für diese Fälle als dem face-to-face-Setting gleichwertig (also grundsätzlich gleich valide und reliabel) betrachtet werden (Kohler, 2015a).

Die Gesamtstichprobe der Fälle unterteilt sich in die drei Teilstichproben A), B) und C):

A) Die Stichprobe der Kinder, deren Redeunflüssigkeiten zentraler Gegenstand des Einsatzes der Kompass waren, lässt sich in Ergänzung zu den schon gemachten Ausführungen wie folgt beschreiben:

Von den insgesamt 60 Kindern waren 45 männlich und 15 weiblich. Die Geschlechterverteilung entspricht mit 3:1 dem für das Phänomen (beginnendes) Stottern bekannten Verhältnis (Natke, 2005).

Sechsendvierzig Kinder waren ohne, zehn Kinder mit Migrationshintergrund (=18 %). Bei vier Kindern gab es dazu keine Angaben. Der Migrationsanteil der Stichprobe unterschreitet nicht wesentlich denjenigen, welcher für die Population im deutschsprachigen Gebiet bekannt ist (= 23,6 % im Jahr 2017. Vgl. Bundeszentrale für politische Bildung, 2018).

B) Die Stichprobe der Anwender des RedeflussKompasses waren enge Bezugspersonen zum Kind. Sie lassen sich bei einem N-Bezug=54 in drei Gruppen unterteilen:

- 37 Eltern (Mutter oder Vater bzw. Mutter und Vater) = 68,5 %,
- 15 Kindergartenlehrpersonen = 28 %
- Zwei Spielgruppenleiterinnen = 3,5 %

Diese Stichprobe erscheint in ihrer Zusammensetzung als repräsentativ für die Population, die den RfK 3.0 in der Praxis anwenden. Insbesondere der hohe Anteil an den engsten Bezugspersonen (Eltern mit 68,5 %) ist realistisch.

C) Die Stichprobe der Fachpersonen der Erstanwendung des SK lässt sich bei einem N=57 der relevanten Fälle hinsichtlich ihres unterschiedlichen Erfahrungshintergrundes wie folgt beschreiben:

- Bei drei Fällen waren Logopädinnen oder Sprachtherapeutinnen im ersten Berufsjahr (= 5 %) die Anwenderinnen.
- Bei 30 Fällen waren Logopädinnen oder Sprachtherapeutinnen ab dem zweiten Berufsjahr (= 53 %) die Anwenderinnen.
- Bei 24 Fällen waren Logopädinnen oder Sprachtherapeutinnen mit Schwerpunkt im logopädischen Bereich der Redeflussstörungen (= 42 %) im Einsatz.

Diese Stichprobe erscheint in ihrer Zusammensetzung als repräsentativ für die Population, die den SK in der Praxis anwenden. Insbesondere die Anzahl der Personengruppe (= 5 % + 53 % = 58 %) mit dem Erfahrungsgrad einer Novizin oder einer Logopädin mit Berufserfahrung ab einem Jahr ist die Zielgruppe für die Anwendung des SK. Außerdem wird vermutet, dass bei Stottern oft schon von Beginn an Expertinnen in die Abklärung involviert sind, so dass der relativ hohe Anteil dieser Personengruppe (42 %) realistisch ist. Das Gesamt-N der Fälle entspricht nicht dem Gesamt-N der Stichprobe der Anwender, da einige Anwender zwei oder mehr Fälle bearbeiteten. Das Gesamt-N der Anwender ist NAnwender = 33.

3.2 Dokumentenanalyse – Erhebungsverfahren

Im Folgenden werden die Typen von Dokumenten dargestellt, die für die Analyse zur Verfügung standen. Zudem werden Spezifikationen einzelner Erhebungsverfahren dargestellt, damit das Zustandekommen der Dokumente nachvollzogen werden kann. Die aufgeführten Zahlen referieren auf die jeweilige Anzahl der vorhandenen Dokumente.

3.2.1 Ausgefüllte RedeflussKompass 3.0 und StotterKompass

Die Analyse umfasste 54 RfK 3.0 für die Erstdurchführung und 5 für die Zweitudurchführung. Für den StotterKompass standen bezogen auf die Erstdurchführung 57 ausgefüllte Exemplare und für die Zweitudurchführung 17 Exemplare zur Verfügung.

Quantitative Analyse: Punktwerte, die durch Ankreuzen der Items entstanden. Hier wurde insbesondere die Gesamtpunktwerte betrachtet und die dreistufigen Empfehlungen, die auf Basis der Summenbildung aller Punktwerte zustande kam. Der ebenfalls dreigestufte, „wahre Wert“ wurde durch Befragung von Fachleuten erhoben, welche die weiterführenden Maßnahmen durchgeführt hatten.

Qualitativ-quantitative Analyse: Erfassung der Randnotizen und deren Häufigkeit.

3.2.2 Video- oder Audioaufnahmen der Durchführungen

Es standen 14 Durchführungen des StotterKompasses in Form von Video- oder Aufnahmen mit und ohne Transkription zur Verfügung.

3.2.3 Protokolle und Audioaufnahmen der Interviews mit Fachpersonen

Diese Interviews mit den Fachpersonen zur Durchführung des StotterKompasses (35) wurden immer vom Leiter des Forschungsprojektes durchgeführt. Oftmals am Telefon, aber auch in face-to-face-Situationen und zwar auf Basis eines schwach strukturierten Gesprächsleitfadens mit offenem Einstieg, Verständnissicherung und Nachfragen.

3.2.4 Schriftliche Reflexionen der Fachpersonen

Einige Mitglieder des Forschungsprojektes arbeiteten im Rahmen einer wissenschaftlichen Arbeit am Projekt mit. In der Regel führten diese Personengruppe Einzelfallstudien zum Einsatz der Kompass durch und dokumentierten ihre Reflexionen (9) zur Durchführung schriftlich.

3.2.5 Dokumentation der Interviews mit den Bezugspersonen

Grundlage waren unterschiedlich stark strukturierte Gesprächsleitfäden. Diese Interviews (17) wurden zum Teil von den Fachpersonen selbst oder zum Teil von Mitgliedern des Forschungsprojektes durchgeführt. Das auszuwertende Material (Protokolle, Transkriptionen und Audioaufnahmen) lag dementsprechend in unterschiedlich starker Informationsdichte und unterschiedlichem Reduktionsgrad vor.

3.2.6 Protokolle von zwei Gruppendiskussionen mit Experten

Es gab insgesamt zwei Gruppendiskussionen mit Experten. Diesen Experten lagen die Kompass vor und sie wurden von der Projektleitung in die komplementäre Durchführung der Kompass per Videomaterial eingeführt. Es wurden Ausschnitte aus einer Durchführung des StotterKompasses gezeigt und so das mögliche Vorgehen exemplarisch veranschaulicht. In der einen Gruppe hatten die Experten selbst alle schon Erfahrung mit dem Instrumentarium gesammelt. In der anderen Gruppe waren das die Minderheit. Von den eher offenen und schwach strukturierten Diskussionsrunden wurden die Expertenmeinungen protokolliert.

3.2.7 Protokolle von vier Gruppenbefragungen mit Laien

Es wurden mit Früherzieherinnen, Spielgruppenleiterinnen, Kindergartenlehrpersonen und Primarschullehrkräften Gruppenbefragungen durchgeführt. Diese Berufsgruppen hatten an einer Fortbildung zum Einsatz der Kompass teilgenommen und sich insofern eine Meinung zur Einsatzfähigkeit insbesondere des RedeflussKompasses 3.0 bilden können.

3.3 Dokumentenanalyse – Verfahren der Aufbereitung und Auswertung

3.3.1 Aufbereitung und Auswertung der Qualitativen Daten

Da die Aufbereitung und Auswertung von qualitativen Daten in diesem Projekt eng miteinander verwoben sind, werden sie miteinander dargestellt. Das Vorgehen lehnt sich an die Vorschläge von Mayring (2015) zur qualitativen Inhaltsanalyse an.

Die grundlegende Vorgehensweise geschah pro Fall in vier Durchgängen A) bis D). Die zwei Gruppendiskussionen und die vier Gruppenbefragungen wurden bei der Aufbereitung und Auswertung technisch wie je ein Fall betrachtet.

A) Erster Durchgang mit induktiv-qualitativem Charakter

Die Daten mit unterschiedlicher Verarbeitungstiefe wurden über einen relativ großen Zeitraum gesammelt (= 4 Jahre). Schon während dieser Erhebungsphase wurde in einem Forschungstagebuch die zentralen Schlussfolgerungen und Ergebnisse pro Fall gesammelt. Dies kann als erster Durchgang der Aufbereitung und inhaltlichen Reduktion betrachtet werden. Die Reduktion bestand in einer vorläufigen Formulierung von inhaltlichen Kategorien.

B) Zweiter Durchgang mit induktiv-qualitativem Charakter:

Zentrale Funktion des zweiten Durchgangs war die endgültige Formulierung der Kategorien. Dies geschah nach Ende der Datenerhebungsphase, indem alle vorliegenden Dokumente pro Fall noch einmal vollständig gesichtet und mit dem Tagebucheintrag verglichen wurden.

Dabei geschah nicht nur eine weitere Reduktion und Verdichtung des Materials zu den endgültigen Kategorien, sondern auch eine Zuschreibung von Wertigkeit zu den Kategorien. Beispielsweise wurde der Einsatz von Farben der Kompasskontrovers diskutiert. Für die Aufbereitung bedeutete dies, dass nicht nur erfasst wurde, ob bei einem Fall die Farben als solche diskutiert wurde, sondern auch, ob sie positiv oder negativ betrachtet wurden (=Wertigkeit). Viele Kategorien wurden aber nur ob ihres Vorkommens erfasst, weil die Wertigkeit nur in einer Richtung vorkam. Z. B. wurden unter der Kategorie „Begriffe“ nur das Vorkommen von Äußerungen zu missverständlichen bzw. uneindeutigen Begriffen erfasst, oder es wurden unter der Kategorie „Ökonomie“ nur Äußerungen subsumiert, die den RfK ausschliesslich in einem positiven Kosten-Nutzen-Verhältnis darstellen.

C) Dritter Durchgang mit deduktiv-quantitativem Charakter

In einem dritten Durchgang wurden die Dokumente pro Fall den verdichteten Kategorien quantitativ zugeordnet, um einen Überblick zur Häufigkeit einzelner Elemente in einer Kategorie zu erhalten. Der dritte Durchgang hatte insofern einen deduktiven Charakter, weil die zuvor induktiv gefundenen Kategorien nun als Theorie betrachtet wurden und die Fälle dahingehend abgesucht wurden. Die Häufigkeiten entschieden über die Aufnahme in das endgültige Kategorienraster. Es wurden nur solche Kategorien gebildet, deren Elemente mindestens zehnmal genannt wurden.

D) Vierter Durchgang – Analyse des Scheiterns

Im vierten Durchgang wurden nur diejenigen Fälle noch einmal gründlich analysiert, bei denen es eine fehlende Übereinstimmung zwischen Empfehlung der Kompass und dem wahren Wert gegeben hatte. Es wurde versucht, die Gründe für die fehlende Übereinstimmung genau zu erfassen.

Diese Form der Dokumentenanalyse stellt also eine Mischung aus induktivem und deduktivem Vorgehen dar. Sie ist schwerpunktmässig qualitativ angelegt, hat aber durch die Relevanz der Häufigkeiten auch eine quantitative Ebene.

3.3.2 Aufbereitung der quantitativen Daten

Die Punktwerte der vorliegenden Kompass, die daraus abgeleiteten dreistufigen Empfehlungen, der dreigestufige wahre Wert des Beratungs-, Diagnostik- oder Therapiebedarfs, die codierten personenbezogenen Daten sowie das Vorkommen von Elementen einer qualitativen Kategorie wurden zunächst fallweise in eine Exceltabelle eingetragen. Diese wurde für die statistische Verrechnung in das Softwareprogramm SPSS importiert.

4 Ergebnisse

Das Kapitel 4 beschränkt sich auf die bloße Darstellung der Ergebnisse. Deren Bedeutung wird erst in Kapitel 5 diskutiert.

4.1 Sensitivität und Spezifität der beiden Kompass

Zunächst werden die wichtigsten quantitativen Auswertungen dargestellt. Es handelt sich um die Bestimmung der Gütekriterien Sensitivität und Spezifität getrennt für beide Kompass.

Die Sensitivität gibt die Fähigkeit des Verfahrens wieder, den prozentualen Anteil der in Wahrheit auffälligen Personen zu entdecken. Die Spezifität gibt hingegen die Fähigkeit des Verfahrens an, den prozentualen Anteil der in Wahrheit unauffälligen Personen zu identifizieren. Grundle-gend für die Berechnung dieser Gütekriterien sind die folgenden Kreuztabellen. Zunächst wird die Auswertung für den RedeflussKompass 3.0, danach für den StotterKompass dargestellt. Bei der Berechnung für den RedeflussKompass 3.0 wurde nur die Erstdurchführung herangezogen, weil die Zweitudurchführung mit einem N = 5 vom Umfang her zu gering war.

Tab. 5: Kreuztabelle der dreigestuften RfK Empfehlungen und des dreigestuften wahren Erstberatungs-, und Ersterfassungsbedarf

		Erstberatungs- und Ersterfassungsbedarf bei beginnendem Stottern			
		kein Bedarf	Entwicklungsbeobachtung	Bedarf	Gesamt
RfK 3.0 Empfehlung	keine Maßnahme	9	0	0	9
	Entwicklungsbeobachtung	1	1	0	2
	Weiterweisung	1	6	36	43
	Gesamt	11	7	36	N = 54

Für die Berechnung der Sensitivität und Spezifität des RedeflussKompasses 3.0 ist es notwendig, diese Kreuztabelle zu binärisieren. Da es sich bei dem RedeflussKompass 3.0 um ein Instrument handelt, welches die Entscheidung für den Erstberatungs- und Ersterfassungsbedarf bringen soll, wird so binärisiert, dass die Stufe Entwicklungsbeobachtung und Weiterweisung bzw. Bedarf zusammengefasst werden. Daraus ergibt sich folgende zweigestufte Kreuztabelle:

Tab. 6: Kreuztabelle der binärisierten RfK 3.0 Empfehlungen und des binärisierten wahren Erstberatungs- und Ersterfassungsbedarfs

		Erstberatungs- und Ersterfassungsbedarf bei beginnendem Stottern		
		kein Bedarf	Bedarf	Gesamt
RfK 3.0 Empfehlung	keine Maßnahme	9	1	10
	Weiterweisung	2	42	44
	Gesamt	11	43	N = 54

Die Sensitivität des RfK 3.0 lässt sich daraus wie folgt berechnen: $42 : 43 \times 100\% = 97,6\%$

Die Berechnung der Spezifität des RfK 3.0 ist wie folgt: $9 : 11 \times 100\% = 81,8\%$

Für die Berechnung der Gütekriterien beim StotterKompass wurden ebenfalls nur die Empfehlungen der Erstdurchführung herangezogen, weil die Anwender der Erstdurchführung in ihrer Zusammensetzung (= Logopädinnen mit und ohne Spezialisierung für Stottern) der Realität in der Praxis viel besser entsprachen als dies bei der Zweitudurchführung (nur Experten für Stottern) der Fall war (vgl. auch unter 3.1.1 Punkt VI).

Tab. 7: Dreigestufte Übereinstimmung zwischen den Empfehlungen von Erstdurchführung (= SK1) und dem wahren Beratungs-, Diagnostik oder Therapiebedarf bei beginnendem Stottern

		Beratungs-, Diagnostik- oder Therapiebedarf bei beginnendem Stottern			
		kein Bedarf	Entwicklungsbeobachtung	Bedarf	Gesamt
SK1 Empfehlung	Entwicklungs-Beobachtung	7	0	0	7
	Diagnostik, Beratung	2	0	0	2
	Therapie wahrscheinlich	3	9	36	48
	Gesamt	12	9	36	N = 57

Für die Berechnung der Sensitivität und Spezifität des StotterKompasses ist es ebenso notwendig, die Kreuztabelle zu binärisieren. Da es sich bei dem StotterKompass um ein Instrument handelt, welches den weiteren Beratungs-, Diagnostik oder Therapiebedarf entdecken soll, wird so binärisiert, dass die Empfehlungen des SK zur Diagnostik, Beratung mit der Wahrscheinlichkeit zur Therapie zusammengefasst werden und beim wahren Wert dementsprechend die Entwicklungsbeobachtung und der Bedarf zusammengefasst werden. Daraus ergibt sich folgende zweigestufige Kreuztabelle:

Tab. 8: Binärisierte Übereinstimmung zwischen den Empfehlungen von Erstdurchführung (= SK1) und dem wahren Beratungs-, Diagnostik oder Therapiebedarf bei beginnendem Stottern

		Beratungs-, Diagnostik- oder Therapiebedarf bei beginnendem Stottern		
		kein Bedarf	Bedarf	Gesamt
SK1 Empfehlung	Entwicklungsbeobachtung	7	0	7
	Diagnostik, Beratung, Therapie wahrscheinlich	5	45	50
	Gesamt	12	45	N = 57

Hieraus ergibt sich folgende Berechnung der beiden Gütekriterien für den StotterKompass:

- Sensitivität des StotterKompasses = $45 : 45 \times 100\% = 100\%$
- Spezifität des StotterKompasses = $7 : 12 \times 100\% = 58\%$

4.2 Inter-Rater-Reliabilitäten des StotterKompasses

Dank einer ausreichend hohen Anzahl an unabhängigen Zweidurchführungen kann für den StotterKompass auch die sogenannten Inter-Rater-Reliabilität berechnet werden. Prinzipiell ist die Bestimmung auf drei Ebenen möglich. Die inhaltlich relevanteste, aber auch größte ist die Ebene der Empfehlungen. Die zweite Möglichkeit besteht darin, die Gesamtsummenwerte zwischen Erst- und Zweidurchführung zu korrelieren. Die dritte Ebene ist die Übereinstimmung auf Einzelitemebene. Letzteres macht aus mehreren Gründen wenig Sinn (vgl. dazu Kohler, 2017). Im Folgenden werden daher die Kennwerte für die Ebene der Empfehlungen und der Gesamtsummenwerte dargestellt. Wirtz & Caspar (2002) schlagen für die Berechnung der Inter-Rater-Reliabilität bei Ordinaldaten u.a. den Spearman Rho vor. Vorgängig wird der jeweilige deskriptive Hintergrund veranschaulicht. Für die Ebene der Empfehlungen geschieht dies mit einer Kreuztabelle und für die Ebene der Gesamtsummenwerte mit einem einfachen Streudiagramm.

Tab. 9: Deskription der Übereinstimmung zwischen den Empfehlungen von Erstdurchführung (= SK1) und Zweidurchführungen (= SK2) des StotterKompasses

		SK2 Empfehlung			Gesamt
		Entwicklungsbeobachtung	Diagnostik, Beratung	Therapie wahrscheinlich	
SK1 Empfehlung	Entwicklungsbeobachtung	2	0	0	2
	Diagnostik, Beratung	0	0	0	0
	Therapie wahrscheinlich	2	1	12	14
Gesamt		4	1	12	N = 17

Wirtz und Caspar (2002) empfehlen die Angabe von mehreren Maßen der Übereinstimmung, wenn das Skalenniveau nicht eindeutig bestimmt werden kann. Daher befindet sich in der Tabelle mit dem Cramers V auch ein Kennwert für das Nominalskalenniveau für die Ebene der Empfehlungen. Außerdem wird der Kappa angegeben, der als Goldstandard der Übereinstimmungsmessung gilt. Es bietet sich allerdings an, dem Spearman Rho die höchste Aufmerksamkeit zu schenken, weil die Werte mit den anderen Berechnungen (z.B zur Inter-Rater-Reliabilität auf Ebene der Gesamtsummenwerte), welche klar Ordinaldatenniveau haben, damit besser vergleichbar sind.

Tab. 10: Zusammenhang bzw. Übereinstimmung zwischen den Empfehlungen von Erstdurchführung (= SK1) und Zweidurchführungen (= SK2) des StotterKompasses

Inter-Rater-Reliabilität des StotterKompasses auf Ebene der Empfehlungen				
		Wert	Standardfehler	Signifikanz
Nominal- bzgl. Nominalmaß	Cramer-V	,658		,025
Ordinal- bzgl. Ordinalmaß	Korrelation nach Spearman	,607	,177	,010
Maß der Übereinstimmung	Kappa	,495	,227	,009
Anzahl der gültigen Fälle		N = 17		

Die Berechnung mit dem Spearman Rho führt zu dem Koeffizienten von $r = 0.607^{**2}$ bei dem Standardfehler von $s_e = 0.177$ für die Inter-Rater-Reliabilität des StotterKompasses auf der Ebene der Empfehlungen.

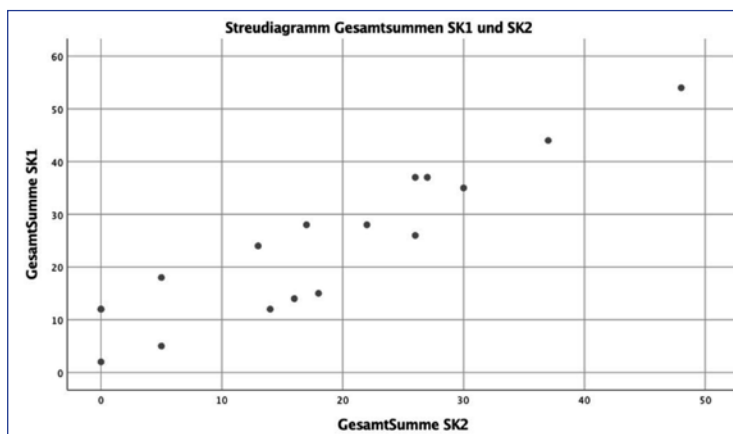


Abb. 7: Zusammenhang zwischen Gesamtsummenwerte der Erst- und Zweidurchführung beim StotterKompass (N = 17)

Die Berechnung mit dem Spearman Rho führt zu dem Koeffizienten von $r = 0.902^{**}$ bei dem Standardfehler von 0.046 für die Inter-Rater-Reliabilität des StotterKompasses auf der Ebene der Gesamtpunktzahlen.

4.3 Kriteriumsvaliditäten der Kompassse

Die hier bestimmten Kriteriumsvaliditäten beziehen sich auf die dieselben Datengrundlagen wie die zuvor berechnete Sensitivität und Spezifität. Die Empfehlungen der Kompassse werden mit dem wahren Wert als konvergentes Kriterium mittels schließender Statistik auf Zusammenhang bzw. Übereinstimmung geprüft. Es gelten daher die oben dargestellten Deskriptionen (vgl. Tab. 5 und 7) und es werden an dieser Stelle nur die statistischen Kennwerte als Ergebnisse der Auswertung dargestellt.

Tab. 11: Maße des Zusammenhangs bzw. der Übereinstimmung zwischen den dreigestuften Empfehlungen des RfK 3.0 und dem dreigestuften wahren Wert zum Bedarf an Beratung, Diagnostik oder Therapie

Kriteriumsvalidität des Redeflusskompasses				
		Wert	Standardfehler	Signifikanz
Nominal- bzgl. Nominalmaß	Cramer-V	,669		,000
Ordinal- bzgl. Ordinalmaß	Korrelation nach Spearman	,808	,069	,000^c
Maß der Übereinstimmung	Kappa	,656	,102	,000
Anzahl der gültigen Fälle		N = 54		

Wie oben unter 4.2 begründet werden auch hier mehrere Kennwerte nebeneinander angegeben. Es befindet sich in der Tabelle wieder der Cramers V als ein Kennwert für das Nominalskalenniveau. Außerdem wird der Kappa angegeben, der als Goldstandard der Übereinstimmungs-

^c2*. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

messung gilt. Es bietet sich auch hier an, dem Spearman Rho die höchste Aufmerksamkeit zu schenken. Der Kennwert für die Kriteriumsvalidität des Redeflusskompasses 3.0 kann durch den Spearman Rho mit dem hochsignifikanten Wert von 0.808** bei einem Standardfehler von 0.069 angegeben werden.

Tab. 12: Masse des Zusammenhangs bzw. der Übereinstimmung zwischen den dreigestuften Empfehlungen des SK1 und dem dreigestuften wahren Wert zum Bedarf an Beratung, Diagnostik oder Therapie

Kriteriumsvalidität des StotterKompasses		Wert	Standardfehler	Signifikanz
Nominal- bzgl. Nominalmaß	Cramer-V	,593		,000
Ordinal- bzgl. Ordinalmaß	Korrelation nach Spearman	,689	,080	,000
Maß der Übereinstimmung	Kappa	,438	,104	,000
Anzahl der gültigen Fälle		N = 57		

Auch für die Kriteriumsvalidität des StotterKompasses bietet es sich an, dem Spearman Rho die höchste Aufmerksamkeit zu schenken. Der Kennwert für die Kriteriumsvalidität des StotterKompasses kann durch den Spearman Rho mit dem hochsignifikanten Wert von 0.689** bei einem Standardfehler von 0.080 angegeben werden.

4.4 Erfahrungen der Anwender

Wie unter 3.3 beschrieben wurden die Erfahrungen der Bezugspersonen und der Fachpersonen mit dem Einsatz der Kompassse durch eine breit angelegte Dokumentenanalyse systematisch aufbereitet. Die Ergebnisse werden im Folgenden auf relativ hoher Abstraktionsebene in komprimierter Tabellenform dargestellt (siehe Tab. 13). Die Häufigkeiten ihrer Nennung waren letztendlich für die Kategorienkonstruktion ausschlaggebend und werden deshalb pro Kategorie angegeben.

Tab. 13: Kategorien der Erfahrungen, welche die Anwender beim Einsatz der Kompassse gemacht haben.

Code	Kategorie	Kategorienbeschreibung	Häufigkeit der Elemente bei N=66
Kategorien mit kritischer bzw. negativer Wertigkeit			
A	Variabilität	Variabilität des Stotterns wird im SK oder RfK 3.0 nicht angemessen erfasst	35
B	Begriffe	Begriffe im RfK 3.0 oder SK sind uneindeutig und missverständlich	31
C	Dimension Zeit	Es fehlt die Unterscheidung von Qualität und Quantität der Symptomatik auf der Dimension Zeit („früher und jetzt“) bei SK und RfK 3.0	25
D	Gesamtpunktzahl	hohe Gesamtpunktzahl bei RfK 3.0 und SK oder „scharfer“ Übergang zwischen den Empfehlungen irritiert	25
E	Unlogisch/Abbruch	Die Anwender kreuzen unlogisch an oder sie plädieren für oder gebrauchen ein Abbruchkriterium beim SK oder RfK 3.0	18
Kategorien mit positiver Wertigkeit			
F	Beispiele	Bezugspersonen loben die Beispiele im RfK 3.0	42
G	hilfreich & nützlich	Bezugspersonen und Fachpersonen schätzen den RfK 3.0 bzw. den SK als hilfreich und nützlich	37
H	Ökonomie RfK	Bezugspersonen und Fachpersonen finden RfK 3.0 ökonomisch	34
Kategorien mit differenzierter Wertigkeit			
I	Skalierung	Skalierung im RfK 3.0 oder SK	Positiv: 5 Negativ: 32
J	Farben	Farbsystem beim RfK 3.0 und SK	Positiv: 10 Negativ: 16
K	Ökonomie SK	Einschätzung der Ökonomie des SK durch Fachpersonen	Positiv: 30 Negativ: 2
L	einfach RfK	Einschätzung der Einfachheit des RfK 3.0 durch Bezugspersonen	Positiv: 23 Sowohl positiv als auch negativ: 2
M	Empfehlung	Handeln gemäß Empfehlungen beim RfK 3.0 oder SK	Positiv: 24 Negativ: 3 Weitere Abklärung: 7
Kategorien ohne eindeutige Wertigkeit			
N	Erstberatung	Eine Erstberatung hat beim Einsatz des StotterKompasses stattgefunden	50
O	Randnotiz SK	Fachpersonen machen Randnotizen auf dem SK	45
P	Randnotiz RfK	Bezugspersonen machen Randnotizen auf dem RfK 3.0	15
Q	Allgemeine Regel	Die Allgemeine Regel wurde in der Erstberatung benutzt	12
R	Diskussion Bezugspersonen	Die Durchführung des RfK löst eine Diskussion bei den Bezugspersonen aus	11

Eine konkretere Darstellung der qualitativen Ergebnisse wird im Zuge der nun folgenden Diskussion und Schlussfolgerungen an einigen Beispielen gegeben.

5 Diskussion und Schlussfolgerungen

5.1 Spezifität und Sensitivität

In der Literatur werden keine allgemein gültigen Richtlinien zur Interpretation der Prozentzahlen von Sensitivität und Spezifität angegeben, weil die Prozentzahlen je nach inhaltlicher Fragestellung sehr individuell interpretiert werden müssen.

Hundert Prozent ist natürlich ein hervorragender Wert für die Sensitivität des StotterKompasses und auch die Sensitivität des RfK 3.0 mit 97,6 % gehört in diese Klasse der Bewertung. Bei den Spezifitäten sieht es anders aus. Die Spezifität des RfK 3.0 mit 81,8 % kann als befriedigend bezeichnet werden aber die Spezifität des StotterKompasses mit 58 % kann nicht befriedigen.

Verantwortlich für die schlechte Spezifität des StotterKompasses sind die Fälle fehlender Übereinstimmung zwischen Empfehlungen des SK mit dem wahren Wert. Die Gründe hierfür liefern insbesondere die qualitativen Ergebnisse (vgl. Tab. 13) und zwar die Kategorien mit kritischer bzw. negativer Wertigkeit. Hierzu sei die Kategorie D beispielhaft diskutiert:

Die von den Anwendern kritisch bemerkten hohe Werte bei der Gesamtpunktzahl kommen u.a. aufgrund eines Kumulationseffekt durch die Mehrfachoperationalisierung von Anspannung zustande (beim StotterKompass durch die fünf Items 1.6, 2.3, 3.3, 4.2, 4.5). Dieser Effekt hat wahrscheinlich die größten Auswirkungen auf die Spezifität und führt zu ihrer Unterschätzung.

Die Konstruktion der Zwischenstufe beim StotterKompass ist ein weiterer Grund für den schlechten Wert der Spezifität des StotterKompasses. Dies kann man direkt aus der dreigestuften Deskription (vgl. Tab. 7) ablesen.

Man sieht, dass in Wahrheit neun Fälle einer Entwicklungsbeobachtung zugeführt wurden, für die Zwischenkategorie des SK aber nur zwei Fälle empfohlen wurden und für die Übereinstimmung in der Zwischenkategorie finden sich null Fälle. Das ist ein Zeichen für den Konstruktionsfehler der Zwischenkategorie im StotterKompass.

5.2 Kriteriumsvalidität

Bettet man die berechneten Kennwerte für die Kriteriumsvaliditäten in die Interpretationsvorschläge aus der einschlägigen Fachliteratur ein, so kann man folgendes feststellen:

Der maßgebliche Kennwert für die Kriteriumsvalidität des RedeflussKompasses 3.0 von 0.808** (Spearman Rho) gilt als sehr gut (Bühner, 2011, S.81).

Der maßgebliche Kennwert für die Kriteriumsvalidität des StotterKompasses mit dem Wert von 0.689** (Spearman Rho) gilt ebenfalls als sehr gut.

5.3 Inter-Rater-Reliabilität des StotterKompasses

Die Berechnung mit dem Spearman Rho führt zu dem Koeffizienten von $r = 0.607^{**3}$ bei dem Standardfehler von 0.177 für die Inter-Rater-Reliabilität des StotterKompasses auf der Ebene der Empfehlungen und für die Ebene der Gesamtpunktzahlen zu dem Koeffizienten von $r = 0.902^{**}$ bei dem Standardfehler von 0.046.

In der Literatur zur Testdiagnostik werden grobe Richtlinien zur vorsichtigen Bewertung von Reliabilitäten gegeben (Bühner, 2011). Demnach ist der Wert für die Ebene der Empfehlungen als „eher niedrig“, für die Ebene der Gesamtpunktzahl als „sehr gut“ zu bezeichnen. Wirtz und Caspar (2002) beschäftigen sich differenzierter mit der Inter-Rater-Reliabilität als Gütekonstrukt und mahnen eindringlich zur Vorsicht bei der Anwendung von solchen allgemeinen Richtlinien bei der Interpretation statistischer Kennwerte. Sie empfehlen den Vergleich mit inhaltsähnlichen Untersuchungen.

Im Vergleich zu Messungen der Inter-Rater-Reliabilität anderer Instrumente mit ähnlichem Geltungsbereich fallen die Werte positiv aus. In der Untersuchung des RedeflussKompasses 2.0 (Kohler, 2015a) wurde bei vergleichbarer Stichprobe (= nur Verdacht auf Stottern in der Stichprobe der Fälle) für die Ebene der Empfehlungen ein $r = 0.396^*$ angegeben und für die Ebene der Gesamtpunktzahl ein $r = 0.360^*$ gemessen. Der Erfolg der Optimierung von der Version 2.0 zur Version 3.0 des RedeflussKompasses und die Konzeption des StotterKompasses scheinen hier einen quantitativen Ausdruck zu finden.

3**. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). * Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig).

5.4 Qualitative Ergebnisse

Die qualitativen Ergebnisse geben vielerlei Auskünfte über den Einsatz des Instrumentariums. Sie erhellen das „Wie“ der Anwendung sehr konkret und differenziert. Die positiven Aspekte bestätigen die Güte (vgl. Tab. 13, die Kategorien mit positiver Wertigkeit). So beurteilen viele Anwender den Instrumenteneinsatz grundsätzlich als positiv. Sie finden dafür Worte wie hilfreich, nützlich oder klärend. Diese Bewertung deutet direkt auf eine erfolgreiche Informationsgabe und damit Erstberatung hin, die mit Hilfe der Instrumente gelingt. Die Kategorien mit kritischer oder negativer Wertigkeit (vgl. Tab. 13) liefern reichhaltige Impulse zur Optimierung der Kompassse. So ist klar, dass die Instrumente durch das Ersetzen von missverständlichen Begriffen verbessert werden können. Hierzu ein Beispiel: Ein häufig genanntes Problem war die Skalierung der Symptomatik nach Häufigkeit oder Intensität. Von den Anwendern wurde mehrheitlich bemängelt, dass die Unterscheidung zwischen „manchmal“ und „selten“ semantisch nicht trennscharf ist. Für die Weiterentwicklung der Instrumente bietet es sich an, auf diese verbal abgestufte Häufigkeits- oder Intensitätsbezeichnung zu verzichten und stattdessen eine möglichst stufenlose Bewertungsmöglichkeit zwischen den Polen „nie“ und „sehr häufig“ zu schaffen.

5.5 Methodenkritik und -diskussion

Die wesentliche Kritik an den verwandten Methoden wurde schon im Kapitel 3 angedeutet: Es geht um das Problem einer relativ geringen Variablenkontrolle, wenn man mit Kooperationspartnern in der Praxis zusammenarbeitet. Die Projektleiter dieser Untersuchung hatten oftmals keine direkte Kontrolle über das Geschehene. Die mittels Triangulation gewonnenen Mehrfachperspektiven auf die Datenerhebung und Auswertung ist ein pragmatischer Kompromiss, nicht mehr, aber auch nicht weniger. Ein grundsätzliches Problem dieser Art von Praxisforschung ist die Selektivität der Stichprobe. Ohne die Bereitschaft der freiwilligen Teilnahme sowohl der Fachpersonen als Nutzer des StotterKompasses als auch der Bezugspersonen als Anwender des RedeflussKompasses wäre diese Untersuchung unmöglich gewesen. Der Stichprobe fehlt dementsprechend das wichtige Merkmal der Randomisierung, um die Ergebnisse generalisieren zu dürfen.

Ein weiteres Problem der Stichprobe stellt die angemessene Anzahl der Fälle dar. Insbesondere wenn Methoden der schliessenden Statistik Anwendung finden, sollte man eine so berechnete und damit angemessene Anzahl von Fällen aufweisen können, weil das N die Berechnung der Signifikanz unmittelbar beeinflusst. Mit der Power-Analyse (G* Power, 2009) können die notwendige Anzahl der Fälle berechnet werden. Es ist aber gar nicht so leicht, eine angemessene Effektstärke dafür festzulegen. Für die Kriteriumsvalidität wurde eine hohe Effektstärke (=0,5) angenommen. Für die Inter-Rater-Reliabilität wurde ein Korrelationskoeffizient von 0,8 als Effektstärke angenommen, weil dies für Reliabilitäten im Allgemeinen so gefordert wird (vgl. Bühner, 2011). Für die Berechnung der Kriteriumsvalidität ergab sich durch die Power-Analyse ein N von 66 Fällen, für die Berechnung der Inter-Rater-Reliabilität ein N = 28 als optimale Stichprobengrösse. Beide Grössen wurden nicht ganz erreicht. Da die tatsächlich gefundenen Effektgrössen aber grösser waren und insofern zu einem geringeren notwendigen Stichprobenumfang führen, fällt die Verfehlung des planerischen Stichprobenumfang nicht so ins Gewicht.

Im vorliegenden Forschungsprojekt wurde der wirkliche Beratungs-, Diagnostik oder Therapiebedarf durch Befragung der involvierten Fachleute erhoben. Dies geschah, um herauszukriegen, ob wirklich eine Beratung, Diagnostik oder Therapie stattgefunden hat und inhaltlich gerechtfertigt war.

Es lassen sich vielfältige Einwände gegen diese Form der Operationalisierung finden (vgl. Kohler, 2017). Die Befragung ist eine indirekte Form der Erfassung des Phänomens. Der Forschende sieht und hört das Geschehen nicht direkt, die Befragten können sich irren etc. Das Problem der schon erwähnten verminderten Variablenkontrolle ist hier offensichtlich. Auch die erwünschte Unabhängigkeit zwischen den Ergebnissen der Kompassse und den wahren Werten war in einigen Fällen nicht wirklich gewährleistet, und zwar dann, wenn die Fachleute, welche den StotterKompass durchgeführt hatten, auch in demselben Fall für die Beratung, Diagnostik und Therapie verantwortlich waren.

5.6 Zusammenfassung und Ausblick

Abschließend kann festgestellt werden, dass die gefundenen statistischen Kennwerte der Gütekriterien für ein insgesamt gutes bis befriedigendes Instrumentarium sprechen. Bei der Wei-

terentwicklung der Instrumente sollten aus den qualitativen Ergebnissen die möglichen Optimierungen abgeleitet werden. Dann ist eine weitere Erhöhung der messbaren Güte naheliegend. Eine weiterführende kritische Diskussion der Gütekriterien und deren Zustandekommen im Zusammenhang mit dem Instrumenteneinsatz sprengt leider den Rahmen dieses Beitrages. Dies wird aber bei Kohler & Braun (in Druck) nachvollzogen werden können. Auch werden dort die qualitativen Ergebnisse viel ausführlicher dargestellt und diskutiert. Im Zuge der immer weiter voranschreitenden Digitalisierung wurde eine internetbasierte Webapplikation der Kompassse realisiert (www.hfh.ch/stotternerkennen). Dies erleichtert den Zugang zu den Kompassen und zielt auf eine weitere Ökonomisierung, Optimierung und Vereinfachung in der Handhabung des Instrumentariums.

Schlussbemerkung

Der vorliegende Artikel ist auf Basis eines Forschungsberichtes verfasst worden, welcher in Kürze im Schulz-Kirchner-Verlag als e-book erscheinen wird. Siehe auch unten bei den Literaturangaben: Kohler, J. & Braun, W. (im Druck).

Literatur

- Andrews, G., Craig, A. Feyer, A.-M., Hoddinott, S. Howie, P. & Neilson, M.D. (1983). Stuttering: A review of research findings and theories circa 1982. *Journal Speech Hearing Disorders*, 48, 226-246.
- Bachmair, S., Faber, J., Hennig, C., Kolb, R. & Willig, W. (2007). *Beraten will gelernt sein*. Weinheim: Beltz.
- Bloodstein, O. & Bernstein Ratner, N. (2008). *A handbook of stuttering*. New York: Thomson Delmar Learning.
- Bloodstein, O. (1958). Stuttering as an anticipatory struggle reaction. In: *Stuttering. A Symposium*. New York: Harper & Brothers, 1-69.
- Bloodstein, O. (1960). The development of stuttering. Changes in nine basic features. *Journal of Speech Hearing Disorders*, 25, 219-237.
- Bosshardt, H.G. (2008). *Stottern*. Göttingen: Hogrefe.
- Braun, W., Baumann, U. & Boltshauser, M. (2011). *RedeflussKompass 2.0. Entscheidungshilfe für Fachpersonen bezüglich Beratung und Ablärungsbedarfs bei Sprechunflüssigkeiten für Kinder im Alter von 2 bis 6 Jahren*. Zürich: Hochschule für Heilpädagogik.
- Braun, W. & Kohler, J. (2011). *RedeflussKompass 3.0. Entscheidungshilfe für Fachpersonen bezüglich Beratung und Ablärungsbedarfs bei Sprechunflüssigkeiten für Kinder im Alter von 2 bis 6 Jahren*. Zürich: Hochschule für Heilpädagogik.
- Brügge, W. & Mohs, K. (2006). *Wenn ein Kind anfängt zu stottern: Ratgeber für Eltern und Erzieher*. München: Reinhardt.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson-Studium.
- Bundeszentrale für politische Bildung (2018). <https://www.bpb.de/nachschlagen/zahlen-und-fakten/soziale-situation-in-deutschland/61646/migrationshintergrund-i> Zugriff am 17.6.19.
- Flick, U. (2008). *Triangulation*. Wiesbaden. Verlag für Sozialwissenschaften.
- Guitar, B. (2014). *Stuttering: an integrated approach to its nature and treatment*. 4th edition. Baltimore: Lippincott Williams & Wilkins.
- G*Power (2009). *Programm u.a. für die Berechnung des notwendigen Stichprobenumfangs*. Version 3.1.9.4. Free Download unter <http://www.gpower.hhu.de>. Zugriff am 19.6.19.
- Hansen, B. & Iven, C. (2004). *Stottern bei Kindern: Ein Ratgeber für Eltern und pädagogische Berufe*. Idstein: Schulz-Kirchner.
- Johannsen, H. (2010). Idiografisches Ursachen-, Diagnose- und Therapiekonzept zum Stottern und Einordnung aktueller Forschungsergebnisse aus verschiedenen Fachdisziplinen. In C. Iven & B. Kleissendorf (Hrsg.). *St-t-tt-tottern. Aktuelle Impulse für Diagnostik, Therapie und Evaluation*. Idstein: Schulz Kirchner.
- Johnson, W. (1955). A study of the onset and development of stuttering. In W. Johnson (Hrsg), *Stuttering in Children and Adults*. Minneapolis: University of Minnesota Press.
- Kohler, J. & Braun, W. (2014). *Stotterkompass (Beta-Version)*. Zürich: Hochschule für Heilpädagogik.
- Kohler, J. (2015a). *Empirische Überprüfung des RedeflussKompasses*. Idstein: Schulz-Kirchner Verlag.
- Kohler, J. (2015b). Die Früherfassung des beginnenden Stotterns. *Frühförderung interdisziplinär*. 34 Jg., 32-45.
- Kohler, J. (2017). Evidenz bei der Identifikation des beginnenden Stotterns (2017). *Praxis Sprache* 3, 147-154.
- Kohler, J. & Braun, W. (im Druck). *Früherkennung, Ersterfassung und Erstberatung bei beginnendem Stottern*. Idstein: Schulz-Kirchner.
- Kuckartz, U. (2014). *Mixed-Methods*. Heidelberg: Springer VS.
- Lattermann, C. (2011): Frühkindliches Stottern: Abwarten oder sofort behandeln? *FORUM Logopädie*. 2, S. 6-11.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse*. Weinheim: Beltz.
- Moser, H. (2001): Einführung in die Praxisforschung. In T. Hug (Hrsg.): *Wie kommt Wissenschaft zu Wissen. Bd. 3 Einführung in die Methodologie der Sozial- und Kulturwissenschaften*. Hohengehren: Schneider-Verlag.
- Moser, H. (2015). *Instrumentenkoffer für die Praxisforschung*. Freiburg: Lambertus.
- Natke, U. (2005). *Stottern – Erkenntnisse, Theorien, Behandlungsmethoden*. Bern: Verlag Hans Huber.
- Neumann, K., Euler, H.A. & Schneider P. (2014): *Identifikation von Stottern im Vorschulalter*. Köln: Demosthenes-Verlag der Bundesvereinigung Stottern & Selbsthilfe e.V.
- Neumann, K., Euler, H., Bosshardt, H.-G., Cook, S., Sandrieser, P., Schneider, P., Sommer, M. & Thum, G. (2017). *Stottern und Poltern: Entstehung, Diagnose, Behandlung*. Frankfurt: Peter Lang Academic Research.
- Packman, A. & Attanasio, J. (2010). *A Model of the Mechanism Underspinning Early Intervention for Stuttering*. Presentation Annual convention ASHA Philadelphia.
- Packman, A. (2012). Theory and therapy in stuttering: A complex relationship. *Journal of Fluency disorder* 37, 225 – 233.

- Ochsenkühn, C. & Thiel, M. (2005). *Stottern bei Kindern und Jugendlichen*. Heidelberg: Springer Verlag.
- Sandrieser, P. (2004). *Eine zeitliche Analyse der Unflüssigkeiten im Sprechen von stotternden und nichtstotternden Kindern*. Universitätsbibliothek Aachen: Dissertation.
- Sandrieser, P. & Schneider, P. (2015). *Stottern im Kindesalter* (4. Auflage). Stuttgart: Thieme Verlag.
- Sandrieser, P. (2017). Stottern. In A. Mayer & T. Ulrich. *Sprachtherapie mit Kindern*. München: Ernst Reinhardt Verlag.
- Schneider, P. (2013). *Stottern bei Kindern erfolgreich bewältigen. Ratgeber für Eltern und alle, die mit stotternden Kindern zu tun haben*. Neuss: Natke-Verlag.
- Starkweather, C. W. (1987). *Fluency and stuttering*. Englewood Cliffs, N. J.: Prentice-Hall.
- Starkweather, C.W. & Gottwald, S.R. (1990). The demands and capacities model: Clinical Applications. In: *Journal of Fluency Disorders* 15, 143-157.
- Starkweather, C.W. & Givens-Ackerman, J. (1997). *Stuttering*. Austin, Texas: Pro Ed.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen, Bern: Hogrefe Verlag für Psychologie.
- Yairi, E. & Ambrose, N.G. (2005). *Early Childhood Stuttering*. Austin, Texas: pro-ed.
- Yairi, E. & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of fluency disorders* Vol. 38, 66-87.

Zu den Autoren

Prof. Dr. phil. Jürgen Kohler arbeitet als Senior Lecturer an der Interkantonalen Hochschule für Heilpädagogik in Zürich (HfH) mit Schwerpunkten in der Lehre von Forschungsmethoden, Redeflussstörungen, Entwicklungs- und Lernpsychologie sowie Diagnostik. Sein aktueller Schwerpunkt in der Forschung liegt im Bereich der Redeflussstörungen. Er ist logopädisch-praktisch tätig im Bereich Stottern und Aphasie in der Lehr-Therapie-Praxis der HfH sowie in einer logopädischen Praxis in Konstanz.

Prof. Wolfgang G. Braun arbeitet als Senior Lecturer an der Interkantonalen Hochschule für Heilpädagogik in Zürich (HfH) mit Schwerpunkten in der Lehre von Redeflussstörungen, Diagnostik, Logopädie im Frühbereich sowie Therapie-didaktik. Ferner leitet er seit 20 Jahren die Stotterintensivtherapiewoche ‚Stotterchamp‘. Er ist logopädisch-praktisch tätig im Bereich Stottern und Kindersprache in der Lehr-Therapie-Praxis der HfH, die er auch leitet.

Korrespondenzadresse

Prof. Dr. Jürgen Kohler
Dozent Studiengang Logopädie
Institut für Verhalten, sozio-emotionale
und psychomotorische Entwicklung (IVE)
Interkantonale Hochschule für Heilpädagogik
Schaffhauserstrasse 239
Postfach 5850
CH-8050 Zürich
juergen.kohler@hfh.ch

Prof. Wolfgang Braun
Dozent Studiengang Logopädie
Institut für Sprache und Kommunikation
unter erschwerten Bedingungen
Interkantonale Hochschule für Heilpädagogik
Schaffhauserstrasse 239
Postfach 5850
CH-8050 Zürich
wolfgang.braun@hfh.ch

→ www.hfh.ch/stotternerkennen