

Aus der Klinik und Poliklinik für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie der Ludwig-Maximilians-Universität München  
Direktor: Prof Dr. Gerd Schulte-Körne

# Beurteilung von Arztgesprächen mit Jugendlichen

Dissertation zum Erwerb des Doktorgrades der Medizin an der Medizinischen Fakultät der Ludwig Maximilians-Universität München

Vorgelegt von Jennifer Englmeier  
aus Ulm-Söflingen  
2011

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Berichterstatter: Prof. Dr.med. R. Frank

Mitberichterstatter: Priv. Doz. Dr Irene Schmid  
Priv. Doz. Dr Alarcos Cieza  
Priv. Doz. Dr Susanne Stüber

Dekan: Prof. Dr. med. Dr. h.c. M. Reiser, FACR, FRCR

Tag der mündlichen Prüfung:06.10.2011

*Inhaltsverzeichnis*

1	Einleitung .....	5
2	Theoretischer Hintergrund: Stand der Forschung.....	7
2.1	Arzt-Patient-Kommunikation.....	7
2.1.1	Patientenzentrierte Gesprächsführung.....	7
2.1.2	Verbesserung des ärztlichen Kommunikationsverhaltens durch Training? 7	
2.1.3	Kommunikation-Schlüsselfertigkeit im Medizinstudium?.....	9
2.1.4	Einfluss des Geschlechts auf die Arzt-Patient-Kommunikation.....	10
2.2	Evaluation kommunikativer Fertigkeiten .....	11
2.2.1	Problemstellung .....	11
2.2.2	Evaluationsmöglichkeiten kommunikativer Fertigkeiten.....	11
2.2.3	Videotechnik und Beobachtung.....	13
2.3	Überblick über Bewertungsinstrumente.....	14
2.3.1	Anforderungen an Bewertungsinstrumente .....	14
2.3.2	Beispiele für Bewertungsinstrumente .....	15
2.3.3	Zusammenfassung:.....	20
2.4	Kommunikation in der Kinder-und Jugendpsychiatrie.....	23
2.4.1	Die Besonderheiten der Kinder- und Jugendpsychiatrie.....	23
2.4.2	Gespräch mit Jugendlichen.....	23
2.4.3	Notfallsituation Suizidalität .....	25
2.4.4	Umgang von Medizinstudenten und Ärzten mit Suizidalität .....	26
2.5	Gesprächstraining für Medizinstudenten an der Ludwigs-Maximilians- Universität-München.....	27
2.5.1	Pflichtveranstaltungen im Lehrplan .....	27
2.5.2	Das Wahlfach Kinder- und Jugendpsychiatrie an der Ludwigs- Maximilians-Universität .....	27
2.6	Gesprächstraining für Ärzte mit Berufserfahrung.....	30
3	Zielsetzung.....	32
4	Material und Methoden.....	34

4.1	Auswahl der Videos .....	34
4.2	Beurteilungsinstrumente.....	35
4.2.1	Modifizierte Aufklärungsgesprächsbewertungsskala – mAGBS-Jugendliche.....	35
4.2.2	Verlaufscharakteristik von Gesprächen .....	42
4.3	Rater und Raterschulung .....	44
4.4	Statistische Auswertung.....	46
4.4.1	Güte der Daten .....	46
4.4.2	Mittelwerte und Standardabweichungen.....	47
4.4.3	Reliabilität, Varianzanalysen .....	47
5	Ergebnisse .....	51
5.1	Deskriptive Darstellung .....	51
5.1.1	Qualität der Videoaufnahmen.....	51
5.1.2	Ratingzeiten .....	51
5.1.3	Beurteilungsskala – Einzelitems der mAGBS .....	52
5.1.4	Rollenidentifikation als Patient (Item 24) .....	65
5.1.5	Geschlechtsverteilung der Darsteller.....	66
5.1.6	Zusammenfassung: Rangreihe .....	67
5.1.7	Themenblöcke des Gespräches.....	68
5.2	Reliabilitätsanalyse .....	70
5.2.1	Bewertungstendenz der Rater.....	70
5.2.2	Trennschärfe.....	71
5.2.3	Korrelation zwischen den Ratern.....	73
5.2.4	Korrelation zwischen–„Globale Bewertung des Arztes“ (Item 23) mit der Durchschnittsnote (Mittelwert über alle Themenblöcke) .....	73
5.2.5	Varianzanalyse und Tukeys Additivitätstest.....	74
5.2.6	Intraclasskorrelation .....	76
5.3	Unterschiede zwischen den drei Trainingsgruppen .....	77
5.3.1	Unterschiede in der Beurteilung durch die mAGBS-Jugendliche .....	77
5.4	Gesprächsverläufe .....	79
5.4.1	Gesprächsverläufe in den Videogruppen und Fragestil .....	80

5.4.2	Redeanteile, Fragestil, nicht unterbrochene Äußerungen.....	85
5.4.3	Pausen im Gespräch.....	87
5.4.4	Gesprächstempo.....	89
5.4.5	Hörersignale.....	91
6	Diskussion.....	92
6.1	Vergleich mit den Ergebnissen anderer Bewertungsinstrumente.....	92
6.1.1	Einsatz von Videos zur Anwendungstestung von Instrumenten.....	93
6.1.2	Raterhintergrund.....	93
6.1.3	Raterschulung.....	94
6.1.4	Ratingzeiten.....	94
6.1.5	Unterschiede zwischen den Ratern.....	94
6.2	Anwendbarkeit der modifizierten Aufklärungsgesprächsbewertungsskala Jugendliche.....	96
6.3	Geschlechtsspezifische Unterschiede in der Bewertung der Darsteller.....	98
6.4	Einfluss von Rollenspielen.....	99
6.5	Gütekriterien des Bewertungsinstrumentes.....	101
6.5.1	Reliabilität.....	101
6.5.2	Validität.....	105
6.6	Übertragbarkeit des Themas.....	108
6.7	Einfluss von Training und Berufserfahrung.....	109
6.7.1	Trainingseffekt.....	109
6.7.2	Berufserfahrung.....	111
6.8	Gesprächsverläufe: Interaktions-Analyse-Systeme und Interact.....	113
6.8.1	Vergleich zwischen Interactions-Analyse-Systemen und globalen Bewertungsinstrumenten.....	113
6.8.2	Fragestil und Redeanteil.....	113
6.8.3	Die Pause im Gespräch.....	114
6.8.4	Gesprächstempo.....	115
6.8.5	Zusammenfassung.....	115
6.9	Emotionen und nonverbales Verhalten.....	116
6.10	Methodenkritik: Stärken und Schwächen der Arbeit.....	118

## Einleitung

---

6.11	Verbesserungen und Ausblick.....	120
6.12	Zusammenfassung.....	122
7	Literaturverzeichnis .....	125
8	Anhang.....	133
8.1	Die mAGBS-Jugendliche (Bewertungsbogen).....	134
8.2	Ergänzende Tabellen .....	138
8.2.1	Geschlecht der Arztdarsteller .....	138
8.2.2	Gesprächsverläufe .....	139
9	Verzeichnisse.....	142
9.1	Tabellenverzeichnis .....	142
9.2	Abbildungsverzeichnis.....	143
9.3	Formelverzeichnis .....	144
9.4	Abkürzungsverzeichnis .....	144
10	Danksagung .....	145

## 1 Einleitung

Die enorme Bedeutung guter kommunikativer Fähigkeiten für die Ausübung des ärztlichen Berufs ist derzeit überall anerkannt (Langer & Schnell, 2009). Aber in der Praxis des Berufsalltags gestaltet sich gelungene Kommunikation dennoch oft schwierig und immer noch werden Ärzte im Studium und in der Facharztausbildung nicht ausreichend dafür vorbereitet (Buckman, 2002). Sowohl aus der Arzt- als auch aus der Patientenperspektive sind Kommunikationsprobleme eine große Quelle für Missverständnisse und Unzufriedenheit. Patienten beklagen, dass Ärzte ihnen nicht zuhören, sie keine Zeit haben und ihre Kernprobleme nicht erfassen. Ärzte auf der anderen Seite sind immer wieder von der mangelnden Compliance ihrer Patienten irritiert (Maguire & Pitceathly, 2002). Darüber hinaus sind Ärzte mit guten kommunikativen Fähigkeiten viel zufriedener mit ihrem Beruf (Satterfield & Hughes, 2007). Kommunikationsstörungen zwischen Arzt und Patient hat es zu allen Zeiten gegeben, doch in der modernen Medizin, die geprägt ist von großen technischen und wissenschaftlichen Entwicklungen, gewinnen solche Kommunikationsstörungen an Tragweite (Koerfer et al., 2008).

Häufig wird angeführt, dass ein Arzt aus ökonomischen Gesichtspunkten nicht die Zeit habe, sich lange mit der Anamneseerhebung aufzuhalten (Terzioglu, Jonitz, Schwantes & Burger, 2003). Eine gute Gesprächstechnik kann jedoch sogar zeitsparend sein (Geisler, 2002).

Auch Medizinstudenten sind sich bewusst, wie wichtig kommunikative Fertigkeiten für die spätere Berufsausübung sind (Schildmann & Schildmann, 2009). In einer Umfrage, die 2003 an sieben deutschen Universitäten unter Studenten, die vor kurzem ihr Studium beendet hatten, durchgeführt wurde, sahen 85 Prozent der Teilnehmer soziale Kompetenz im Umgang mit Patienten als wichtigste Fähigkeit an und bedauerten, dass der Unterricht in psychosozialen Fächern an den Universitäten einen geringen Stellenwert verglichen mit anderen Fächern hat (Jungbauer, Alfermann, Kamenik & Brahler, 2003).

Je nach Fachrichtung hat die ärztliche Gesprächsführungskunst unterschiedliche Bedeutung. An einen psychotherapeutisch tätigen Arzt werden ganz andere Anforderungen bezüglich seiner kommunikativen Fähigkeiten gestellt als an einen Chirurgen, der überwiegend operativ tätig ist (Dorfmueller, 2001).

Gerade bei psychosozialen und emotionalen Themen zeigen Ärzte häufig Unsicherheit und Überforderung und erleben Patienten dann als schwierig (Satterfield & Hughes, 2007; Schwantes, 2009).

Im englischsprachigen Raum beschäftigt man sich nun seit vielen Jahren in der For-

schung mit der Arzt-Patient-Interaktion. So kam es immer wieder zum Treffen von Experten, die sich bemühten durch das Aufstellen von Richtlinien wie beispielsweise dem Kalmazoo Consensus Statement Defiziten in der Arzt-Patient-Kommunikation entgegenzuwirken (Makoul, 2001; Simpson et al., 1991). In Deutschland wurde dem Thema Arzt–Patient–Kommunikation erst in den letzten Jahren mehr Beachtung geschenkt, so dass hier verglichen mit dem englischsprachigen Raum im universitären Bereich und in der Forschung ein erheblicher Nachholbedarf besteht.

Häufige Fragestellungen dabei sind:

- Was sind wesentliche Elemente eines guten Gespräches?
- Hat Kommunikationstraining eine Verbesserung kommunikativer Fähigkeiten zur Folge?
- Wie kann ein Kommunikationstraining effektiv gestaltet werden?
- Wie kann man den Erfolg eines Kommunikationstrainings angemessen erfassen und evaluieren?
- Was lässt sich aus wissenschaftlichen Erkenntnissen für die Arzt-Patient- Beziehung für die Praxis ableiten?



## **2 Theoretischer Hintergrund: Stand der Forschung**

### **2.1 Arzt-Patient-Kommunikation**

#### **2.1.1 Patientenzentrierte Gesprächsführung**

Lange Zeit herrschte in der Arzt-Patienten-Beziehung das Modell des „ärztlichen Paternalismus“ vor. Der Arzt gab vor, was der Patient zu tun hatte. Seit den achtziger Jahren aber hat dieses Modell starke Änderungen zugunsten der Patientenautonomie erfahren. Ärztliche Diagnosen und Therapieentscheidungen wurden genauer überprüft und an ärztliche Behandlung wurde der Anspruch eines Evidenz-basierten Vorgehens gestellt. Diese Entwicklung stärkte die Position der Patienten.

Man spricht heute oft von einem partnerschaftlichen Modell zwischen Arzt und Patient (Kruse & Tress, 2001). Gerade chronische Erkrankungen erfordern die Mitarbeit des Patienten. Dies setzt jedoch zunächst ein Krankheitsverständnis durch eine gute Erklärung von Seiten des Arztes voraus. Dadurch werden jedoch hohe Erwartungen an die fachliche, aber auch an die kommunikative Kompetenz eines Arztes gestellt (Terzioglu et al., 2003). Ärzte müssen sich auch in ihrer Gesprächsführungstechnik an die Veränderungen in der Arzt-Patient-Beziehung anpassen. Zwangsläufig hat sich dadurch auch die Form der Anamnese von einer rein „arztzentrierten“ Anamnese, bei der der Arzt seinen Fragenkatalog durchgeht, zu einem mehr „patientenzentrierten“ Gesprächsansatz gewandelt. Der Patient soll die Möglichkeit bekommen, am Gespräch aktiv teilzunehmen, denn bewusstes Gesundheitsverhalten setzt eine aktive Mitwirkung voraus. Auch psychosoziale Hintergründe und Beziehungsaspekte bekommen eine stärkere Gewichtung (Doering, Steuernagel, Kape, Lieck & B., 2008; Smith, 2002).

Experten weisen darauf hin, dass folgenden Gesprächstechniken dazu beitragen können, dass ein Gespräch gelingt: eine angemessene Mischung aus offenen und geschlossenen Fragen, aktives Zuhören, Zusammenfassung des Gesagten, -das Zulassen von Pausen, ein angemessenes Gesprächstempo sowie eine verständliche Sprache (Buckman, 1992; Dorf Müller, 2001; Makoul, 2001).

#### **2.1.2 Verbesserung des ärztlichen Kommunikationsverhaltens durch Training?**

Noch in den 80er Jahren des vorigen Jahrhunderts war die Meinung unter Ärzten weit verbreitet, dass gute Kommunikationsfähigkeiten mehr eine angeborene Begabung als eine erlernbare Fertigkeit seien (Buckmann 2002). Jedoch wurde schon Ende der 80er Jahren von Maguire gezeigt, dass Studierende, die eine intensives Kommunikations-

training mit Videofeedback durchliefen, auch nach 5 Jahren noch über bessere kommunikative Fertigkeiten verfügten als Studierende die kein Training bekommen hatten. Insgesamt 36 Teilnehmer aus beiden Gruppen führten 5 Jahre nach Studienabschluss Interviews mit psychiatrischen Patienten. Diese wurden auf Video aufgezeichnet und von unabhängigen Beurteilern mittels eines Bewertungsinstrumentes bewertet. Die Teilnehmer, die inzwischen als Ärzte arbeiteten, hatten sich alle insgesamt verbessert, jedoch stellten die Ärzte aus der Trainingsgruppe mehr offene Fragen, gingen stärker auf nonverbale Signale der Patienten ein und sprachen psychosoziale Themen häufiger an (Maguire, Fairbairn & Fletcher, 1986).

Im Laufe der Jahre wurde immer mehr erkannt, dass Ärzte bei wichtigen Themen wie dem „Überbringen schlechter Nachrichten“ große Probleme haben. Gerade hier wirken sich wegen der Komplexität des Themas und den aufkommenden Emotionen auf Patienten-, aber auch Arztseite, Defizite in der Gesprächsführung besonders drastisch aus. Viele Kurse und Weiterbildungen zur Verbesserung des ärztlichen Gesprächsverhalten widmen sich daher dem Thema „Überbringen schlechter Nachrichten“ siehe hierzu (Back et al., 2007; Fallowfield, 1996; Schildmann, Härlein, N., Schlögl & Vollmann, 2006).

So konnte in einer randomisierten, kontrollierten Studie nachgewiesen werden, dass auch bereits berufserfahrene Onkologen sich nach einem intensiven 3-tägigen Kurs mit Theorie, Rollenspielen und Diskussion bezüglich ihres Gesprächsverhaltens deutlich verbesserten. Gespräche wurden vor und nach dem Training auf Video aufgenommen und mit einem Bewertungsinstrument beurteilt. Die Verbesserungen betrafen vor allem den vermehrten Einsatz offener Fragen, ein empathischeres Verhalten und ein besseres Eingehen auf nonverbale Signale des Patienten gegenüber einer Kontrollgruppe, die kein Training bekommen hatte (Fallowfield et al., 2002).

Doch auch an anderen Themen, vor allem mit psychosozialen Fokus wächst das Interesse. Langewitz et. al zeigten, dass ein verhaltensorientiertes, zeitlich auf 14 Stunden begrenztes Training in patientenzentrierter Kommunikation bei Internisten zu einer signifikanten Verbesserung in der Strukturierung des Gespräches, dem Herausarbeiten der Beschwerden des Patienten sowie im Eingehen auf den Patienten, gegenüber einer Kontrollgruppe, die dieses 14-stündige Training nicht bekommen hatte, führte. Die Verbesserung wurde mittels eines Bewertungsinstrumentes im Vorher-Nachher-Vergleich gemessen. Patientenfragebögen zur Zufriedenheit der Patienten mit dem Gesprächsverlauf, die zusätzlich erfasst wurden, zeigten jedoch keinen Unterschied zwischen den Gruppen (Langewitz, W. A., Eich, Kiss & Wossmer, 1998). Aus dem Bereich der Pädiatrie zeigen Studien, dass ein Kommunikationstraining Kinderärzten hel-

fen kann, besser mit psychosozialen Themen, die in diesem Fachbereich häufig auftauchen, umzugehen (Hoffmann et al., 2007; Kemper, Foy, Wissow & Shore, 2008; van Dulmen & Holl, 2000). Auch an der Universitätsklinik für Kinder- und Jugendmedizin in Heidelberg wurden Assistenzärzte mit einem 8-stündigen Training geschult. Sie gaben in einem Fragebogen zur Selbsteinschätzung an, sich durch das Training im Gespräch mit Eltern und Kindern mehr Zeit zu nehmen, sich besser in die Eltern hineinversetzen zu können und sich insgesamt in schwierigen Situationen kompetenter zu fühlen (Hoffmann et al., 2007).

Im Bereich der Psychosomatik wurde von einer Kölner Arbeitsgruppe nachgewiesen, dass sich eine Fortbildung auch an formalen Gesprächsmerkmalen, wie der Gesprächsdauer, dem Redeanteil des Arztes und des Patienten, sowie der Länge an nicht unterbrochenen Patientenäußerungen bemerkbar macht (Köhle, Kaerger - Sommerfeld, Koerfer, Obliers & Thomas, 2001).

### **2.1.3 Kommunikation-Schlüsselfertigkeit im Medizinstudium?**

Auch Medizinstudenten sehen einen Bedarf an Unterricht in kommunikativen Fertigkeiten (Jungbauer et al., 2003).

In den letzten Jahren haben sich an vielen medizinischen Fakultäten in Deutschland Kurse etabliert, um theoretische Grundlagen in Gesprächsführung zu vermitteln und auch praktische Übung zu bieten (Petersen, Busche, Bergelt & Huse-Kleinstoll, 2005).

Häufig werden dabei auch sogenannte Simulationspatienten eingesetzt. Darunter versteht man Schauspieler, die Patienten mit einem bestimmten Krankheitsbild darstellen und an denen Studenten Anamneseführung und Untersuchung üben können (Simmenroth-Nayda et al., 2007).

Doch es ist nicht immer einfach, traditionelle klinische Lehre und neuere Unterrichtsmethoden miteinander in Einklang zu bringen (Petersen et al., 2005). Für die Lehrenden ist es eine Herausforderung, Themen zur Gesprächsführung in den Kontext der klinischen Fächer einzubauen. Trennt man nämlich den Unterricht in kommunikativen Fertigkeiten und inhaltliche Themen der einzelnen medizinischen Fächer, entsteht der Eindruck bei den Studierenden, dass dies zwei unterschiedliche Herangehensweisen sind. Die inhaltlichen Themen der einzelnen klinischen Fächer werden meist stärker gewichtet als kommunikative Fertigkeiten und Studenten lernen auf diese Weise nicht integrativ zu denken (Kurtz, S., Silverman, Benson & Draper, 2003).

„Die Förderung kommunikativer Kompetenz ist eben nicht in einer bloßen Rhetorikschulung zu erreichen, sondern muss Inhalten und Zwecken ärztlicher Gesprächsführung folgen, (...)“ so bringen es Koerfer et al. auf den Punkt (Koerfer et al.,

2008).

Die Erfahrung an der Universität Witten/Herdecke zeigt, dass sich dieses Problem lösen lässt, wenn man den Unterricht in kommunikativen Fertigkeiten mit dem klinischen Wissen der einzelnen Fächer verbindet. Dadurch kann den Studenten vermittelt werden, dass eine Anamnese nur durch die Verknüpfung von Gesprächsinhalt, Gesprächstechnik und Gesprächsstil gelingt (Schnell & Langer, 2009).

Um zu bewerten, ob ein ärztliches Gespräch gut ist, muss zunächst festgelegt sein, was überhaupt mit „gut“ gemeint ist (Heaven, Maguire & Green, 2003).

Hilfreich kann dabei auch das Aufstellen von fakultätsübergreifenden Lernzielen für soziale und kommunikative Kompetenzen sein, um im gesamten deutschsprachigen Raum eine gemeinsame Basis zu erreichen.

So trafen sich 2008 30 Vertreter verschiedener deutscher, österreichischer und schweizer Fakultäten zu einem Workshop, bei dem sie verschiedene Ausbildungsziele schriftlich als „Positionspapier“ festlegten. Dieses Schreiben ist mit den vorher erwähnten angloamerikanischen Consensus Statements vergleichbar (siehe hierzu 1 Einleitung) (Kiessling et al., 2008).

#### **2.1.4 Einfluss des Geschlechts auf die Arzt-Patient-Kommunikation**

Frauen und Männer unterscheiden sich in ihrer verbalen und nonverbalen Kommunikation. Frauen geben mehr persönliche Informationen, besitzen einen wärmeren nonverbalen Kommunikationsstil, indem sie mehr lachen, mehr Blickkontakt halten und mehr Emotionen zeigen. Diese allgemeine Erfahrung lässt sich auch auf Ärztinnen und Ärzte übertragen (Schmid - Mast & Klöckner - Cronauer, 2009).

So hat sich in einer Metaanalyse von Roter gezeigt, dass Ärztinnen mehr auf psychosoziale Themen eingehen und versuchen, eine partnerschaftliche Beziehung zum Patienten aufzubauen. Ärztinnen zeigen auch ein bestärkenderes nonverbales Verhalten als Ärzte, sie lächeln und nicken im Gespräch mehr (Roter, D. L., Hall & Aoki, 2002).

Doch auch vom Geschlecht des Patienten ist der Verlauf des Arzt-Patient-Gespräches abhängig. Patientinnen werden einfühlsamer behandelt und kommen mehr zu Wort. (Schmid - Mast & Klöckner - Cronauer, 2009).

Seit mehreren Jahren zeichnet sich ein Trend zu einer Verweiblichung der Medizin ab. Inzwischen sind mehr Studienanfänger weiblich, Frauen sind, so zeigt es sich in der späteren Wahl des medizinischen Fachbereichs, auch interessierter an psychosozialen Themen (Hibbeler & Korzilius, 2008).

## **2.2 Evaluation kommunikativer Fertigkeiten**

### **2.2.1 Problemstellung**

Als schwierig in der Umsetzung erweist sich auch eine angemessene Evaluation und Prüfung kommunikativer Fertigkeiten (Koerfer et al., 2008). Um sich verbessern zu können, brauchen Studierende und Ärzte neben einem angemessenen Training eine genaue Rückmeldung über Stärken und Schwächen in ihren aktuellen Gesprächsfertigkeiten (Fallowfield et al., 2002; Maguire & Pitceathly, 2002).

Es hat sich gezeigt, dass auch die Art des Feedbacks Einfluss auf den Lernerfolg hat.

In einer qualitativen Umfrage gaben 83 Teilnehmer eines Weiterbildungskurses zur ärztlichen Gesprächsführung an, folgende Feedback-Methoden als hilfreich zu empfinden: das Schaffen einer angenehmen und respektvollen Atmosphäre, das Mitteilen von eigenen Empfindungen, konstruktive Kritik mit Verbesserungsvorschlägen, die richtige Menge an Feedback (Hewson & Little, 1998).

Auch scheint Feedback mittels Video einen Vorteil gegenüber anderen Methoden zu haben (siehe hierzu 2.2.4 Videotechnik und Beobachtung).

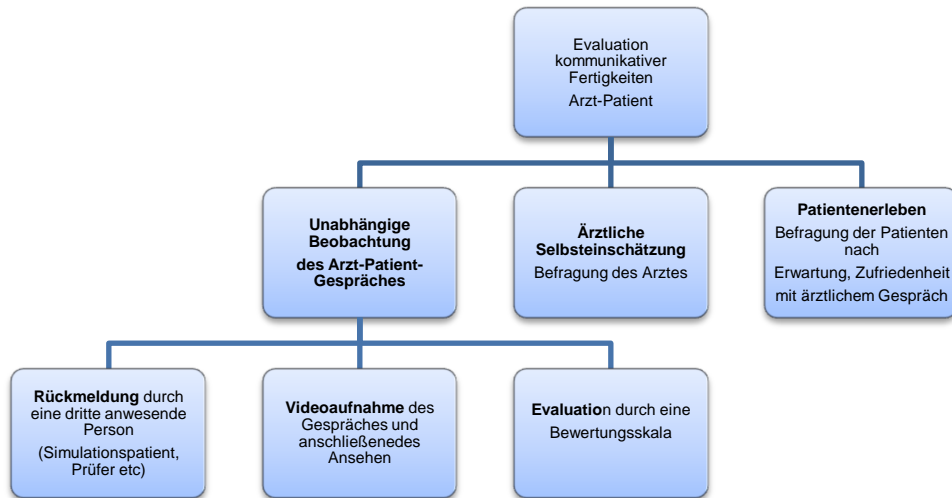
Das Problem an den Universitäten ist jedoch, dass häufig die personellen Kapazitäten fehlen, um ein optimales Feedback zu geben (Puchner-Matzner, Gleiss, Schmidts & Frischenschlager, 2006).

Räumt man kommunikativen Fertigkeiten mehr Gewicht in der Lehre und Weiterbildung ein, so muss ein Lernfortschritt auch durch angemessene Prüfungen kontrolliert werden können. Die üblichen Prüfungsverfahren in Form schriftlicher Tests helfen, wenn es um das Prüfen von Fertigkeiten geht, nur eingeschränkt weiter (Schirmer, J. M. et al., 2005). Die Universitäten stehen vor dem Dilemma, sehr komplexe Fertigkeiten wie Gesprächsführung, die sich je nach Situation, Patientenbedürfnissen und Persönlichkeit des Arztes unterschiedlich zeigen, standardisiert und valide prüfen zu müssen (Zoppi & Epstein, 2002). Eine häufig zur Anwendung kommende Prüfungsform ist das OSCE (Objective structured clinical examination). Doch hier braucht man gut validierte Instrumente (Chenot & Ehrhardt, 2003; Koch & Resch, 2009).

### **2.2.2 Evaluationsmöglichkeiten kommunikativer Fertigkeiten**

Zur Evaluation kommunikativer Fertigkeiten hat man verschiedene Möglichkeiten. Die folgende Abbildung gibt dazu einen Überblick.

Abbildung 1 Übersicht zu Evaluationsmöglichkeiten des Arzt-Patient-Gesprächs



Man kann von den Ärzten eine Selbsteinschätzung bezüglich ihres Kommunikationsverhaltens einholen. Hier zeigt sich jedoch häufig das Problem, dass Ärzte ihre Fähigkeiten über- oder unterschätzen (Davis et al., 2006).

Es gibt darüber hinaus die Möglichkeit, das „Outcome“ von Arzt-Patient-Kommunikation zu messen, indem man Patienten Fragebögen zuschickt und die Zufriedenheit nach einer Behandlung erfragt. Doch auch diese Methode ist kritisch zu sehen, da viele Faktoren, die über die alleinigen kommunikativen Fertigkeiten des Arztes hinausgehen, für die Patienten eine Rolle spielen (Zoppi & Epstein, 2002). Es ist oft nicht möglich eine einzelne Gesprächssituation bewerten zu lassen, sondern die Patienten geben ihren Gesamteindruck während eines Krankenhausaufenthalts wieder (Schirmer, C., Stahl & Günther, 2009).

Eine gerade im Studentenunterricht angewendete Methode ist die unabhängige Verhaltensbeobachtung. So wird bei Übungen zur Gesprächsführung eine direkte Rückmeldung zum Gesprächsverhalten sowohl durch Dozenten als auch durch die übrigen Teilnehmern eingeholt (Kopecky-Wenzel, Maier, Muntau, Reinhardt & Frank, 2009).

Eine besondere Bedeutung kommt auch hier wieder dem Einsatz von Simulationspatienten zu. Denn diese können geschult werden, um den Studenten aus Patientensicht Rückmeldung über das Gespräch zu geben (Simmenroth-Nayda et al., 2007). Auch die Aufnahme des Gesprächs auf Video und anschließendes Ansehen einzelner Sequenzen kann sehr hilfreich sein (siehe hierzu 2.2.4. Videotechnik und Beobachtung). Um jedoch gerade für Prüfungssituationen wie den OSCE eine einheitliches Vorgehen si-

cherezustellen, können Bewertungsskalen sehr nützlich sein (Schirmer, J. M. et al., 2005). In der Psychologie und Psychiatrie ist es eine weit verbreitete Methode, Verhalten mit Hilfe von Skalen zu beurteilen und so subjektive Eindrücke von Verhaltensweisen auf ein objektivierbares Maß zu bringen.

Derartige Bewertungsskalen gibt es auch für die Kommunikations- und Interaktionsforschung in der Arzt–Patient-Kommunikation.

### **2.2.3 Videotechnik und Beobachtung**

#### **2.2.3.1 Theoretische Grundlagen**

Da das Verhalten eines Menschen etwas sehr Komplexes ist, entzieht es sich unserer Wahrnehmung sehr leicht. Auch auf die eigene Wahrnehmung ist nicht immer Verlass, denn sie verändert sich rasch und im Nachhinein findet häufig eine Verzerrung statt.

Durch eine Videoaufnahme hingegen ist es möglich, Verhalten festzuhalten. So ergibt sich die Möglichkeit, das Verhalten im Nachhinein zu analysieren (Thiel, 1997). Die Arbeit mit Videotechnik wird heute in vielen Bereichen eingesetzt. Seit den frühen 70er Jahren wendet man sie in der Psychiatrie und Psychotherapie<sup>1</sup> an, zur Supervision, zur ärztlichen Weiterbildung und auch zur gestützten Selbstkonfrontation von Patienten (Ronge, 1997). Es wird immer wieder diskutiert, ob das Bewusstsein um die statt findende Videoaufnahme das Verhalten der Menschen, die gefilmt werden, verändert. In Studien hat sich jedoch gezeigt, dass nach einer kurzen Gewöhnungsphase ein der Realität entsprechendes Verhalten der Beobachteten auftritt (Pringle & Stewart-Evans, 1990; Ram et al., 1999).

Dennoch ist es wichtig, sich bewusst zu sein, wenn man sich mit Videoaufnahmen beschäftigt, dass sie immer nur Ausschnitte darstellen und nie die ganze Realität widerspiegeln.

Besonders wichtig ist die Beachtung rechtlicher Grundlagen bei der Verwendung von Videoaufnahmen. Die Teilnehmer müssen aus Datenschutzgründen darüber informiert sein und es muss bei Auswertung eine schriftliche Einwilligung vorliegen (Ellgring, 1989).

#### **2.2.3.2 Einsatzmöglichkeiten im Rahmen des medizinischen Unterrichts**

Auch im Studentenunterricht und in ärztlichen Fortbildungsveranstaltungen wird die Arbeit mit Videos eingesetzt, vor allem, um Arzt-Patient-Kommunikation zu üben.

<sup>1</sup>Der Internationale Arbeitskreis Audiovision in Psychiatrie und Psychotherapie wurde 1977 gegründet.

Wie hilfreich dabei der Einsatz von Videotechnik sein kann, zeigt eine Arbeit, in der Studenten beim Erlernen der psychiatrischen Anamnese verschiedene Arten von Feedback bekamen. Die Gruppe, die eine Rückmeldung des Dozenten durch das gemeinsame Ansehen einer gefilmten Übungsgesprächs bekam, konnte sich stärker verbessern, als Studierende aus Kontrollgruppen, die Feedback ohne das Ansehen eines Videos erhalten hatte (Maguire, 1980).

Der Einsatz von Videotechnik erlaubt eine genaue Analyse der Körpersprache, des nonverbalen und emotionalen Verhaltens. Durch die Möglichkeit, einzelne Abschnitte wiederholt anzuschauen, wird der Lerneffekt verstärkt (Kopecky-Wenzel et al., 2009).

Gerade für die Untersuchung von Arzt-Patient-Gesprächen mit einem Bewertungssystem gibt es kein besseres Medium als „Video“ und es wird empfohlen, damit zu arbeiten (Wasserma, Inui 1983).

Ein neuer Ansatz, Videotechnik in der Medizin einzusetzen, ist es auch, sehr kurze Sequenzen aus dem medizinischen Arbeitsumfeld zu filmen, um so einen Zugang zu komplexen Interaktionen und Verhaltensmustern im Gesundheitswesen zu bekommen. In einer Studie wurden kurze Bestandteile von Interaktionen im Operationssaal und ein Arzt-Patient-Gespräch, während der Arzt nebenbei am Computer schreibt, gefilmt und detailliert dargestellt, wie Kommunikation und Interaktion in diesen Situationen ablaufen (Heath, Sanchez 2007).

## **2.3 Überblick über Bewertungsinstrumente**

Der überwiegende Teil der Bewertungssysteme, die es zur Bewertung von Arzt-Patient-Kommunikation derzeit gibt, stammen aus dem englischsprachigen Raum.

Es gibt kaum validierte deutschsprachige Instrumente (Wand, Schildmann, Buchardi & Vollmann, 2008). In den letzten Jahren wurden zwar einige Instrumente ins Deutsche übersetzt und validiert, doch befinden sich die deutschen Universitäten, was Einsatz von Bewertungsinstrumenten betrifft, noch in der Anfangsphase (Scheffer, Muehlinghaus, Froehmel & Ortwein, 2008; Simmenroth-Nayda, Nolte, Fischer & Himmerl, 2008).

### **2.3.1 Anforderungen an Bewertungsinstrumente**

Ein Bewertungssystem sollte der Gesprächstechnik, der Struktur und dem Inhalt des ärztlichen Gesprächs gerecht werden. Gerade vor dem Hintergrund knapper personeller und zeitlicher Kapazitäten an Universitäten sollte es nicht zu kompliziert und zeitintensiv sein (Wand et al., 2008).

Darüber hinaus sollten Bewertungsinstrumente wie jedes andere psychologische



Messinstrument die Kriterien der Testgüte Objektivität, Reliabilität und Validität erfüllen. Häufig verwendet wird im Bereich von Bewertungsskalen der Begriff „Item“. „Item“ bedeutet Variable, Merkmal oder Frage. Im Folgenden sollen anhand der Gütekriterien, Reliabilität, Validität und Objektivität die psychometrischen Qualitäten von Bewertungsinstrumenten diskutiert werden.

Die Reliabilität bezeichnet die Genauigkeit oder Zuverlässigkeit einer Messung. Wenn andere Beurteiler mit gleichem Wissenstand zu einem übereinstimmenden Urteil kommen, kann man von einem reliablen Urteil sprechen. Anders formuliert gibt die Reliabilität den Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte an (Bühner, 2006). Es gibt verschiedene Arten von Reliabilität. Bei Bewertungsinstrumenten wird häufig die Interraterreliabilität verwendet.

Die Interraterreliabilität ist auch als Kriterium der Objektivität zu verstehen. Die Objektivität beantwortet die Frage, ob sich alle Rater an demselben Merkmal bei der Einschätzung orientieren und dieses präzise erfassen.

Die Validität beantwortet die Frage, ob ein Messinstrument das Kriterium, das es messen soll, auch tatsächlich erfasst. Am Beispiel eines Bewertungsinstrumentes wäre Validität dann erfüllt, wenn das Instrument auch tatsächlich kommunikative Fertigkeiten erfasst. Es gibt auch bei der Validität verschiedene Arten: die Inhalts- die Konstrukt- und die Kriteriumsvalidität.

Alle 3 Gütekriterien bauen aufeinander auf und ohne ausreichende Reliabilität gibt es auch keine Validität (Wirtz & Caspar, 2002).

### **2.3.2 Beispiele für Bewertungsinstrumente**

Es gibt zwei unterschiedliche Ansätze, bei Ratingsystemen Arzt-Patient-Beziehungen zu beurteilen. Zum einen in Form eines sehr genauen Interaktions-Analyse-Systems oder einer Checkliste oder aber durch eine globale Ratingskala. Die folgende Tabelle stellt diese beiden Ansätze noch einmal für einen besseren Überblick gegenüber.

Tabelle 1 Gegenüberstellung von Bewertungsinstrumenten

<p><b>Interaktions-Analyse-Systeme/ Checkliste</b></p>	<p><b>Globale Ratingskalen</b></p>
<p>Mikroanalyse des Gesprächs, jede einzelne Äußerung wird in eine Kategorie eingeteilt</p> <p><b>Quantitative Bewertung</b></p>	<p>Globale Bewertung des Gesprächsverlaufs nach verschiedenen Kriterien</p> <p><b>Qualitative Bewertung</b></p>

### 2.3.2.1 Roter Interaction Analysis System (RIAS)

Ein häufig verwendetes System zur Analyse von Interaktionsprozessen im medizinischen Bereich stellt das in den 90er Jahren von Roter entwickelte „Roter Interaction Analysis System“<sup>2</sup> dar (Roter, D. & Larson, 2002). Die Methode basiert auf Arbeiten des Sozialpsychologen Robert Bales (Interaction Process Analysis, Cambridge 1950). Er entwickelte ein Instrument, um das soziale und emotionale Verhalten in kleinen Gruppen zu untersuchen. Die RIAS ist letztendlich eine Modifikation dieses System mit Fokus auf die Arzt–Patient-Kommunikation. Es werden dabei einzelne Äußerungen sowohl des Arztes als auch des Patienten nach verschiedenen Kategorien kodiert. Dabei wird unter Äußerung eine Gesprächseinheit definiert, die einen inhaltlichen und sinngemäßen Zusammenhang hat. Eine Äußerung kann ein einzelnes Wort oder aber auch ein ganzer Satz sein. Es werden folgende Kategorien unterschieden: offene und geschlossene Fragen mit den Bereichen medizinisch, therapeutisch, psychosozial und informativ. Darüber hinaus gibt es globale Items, um den emotionalen Kontext des Gespräches zu bewerten. Es gibt ein ausführliches Manual (Roter, 2006), eine eigene Internetseite [www.rias.org](http://www.rias.org) und eine Software, die das Bewerten eines Gespräches mithilfe des Manuals vereinfacht.

Häufig wird bei der RIAS die Kritik angebracht, sie sei zu allgemein gehalten und würde dem Kontext, der entscheidend ist für die Art, wie der Arzt mit dem Patienten spricht, zu wenig gerecht. Hier wurde von den Autoren der Vorschlag gebracht, parallel zum Kodieren der Äußerungen je nach Fachgebiet Anmerkungen zu machen (siehe Manual Roter, 2006).

So wurde die RIAS beispielsweise weiterentwickelt für onkologische Gesprächssituati-

<sup>2</sup> Im folgenden wird das Roter-Interaktions-Analyse- System abgekürzt mit RIAS

onen, wobei die Grundstrukturen beibehalten wurden, es entstand eine Kombination aus dem RIAS, sowie der Analyse von auf Video aufgezeichneten Arzt-Patient-Gesprächen im onkologischen Setting (Ford, Hall, Ratcliffe & Fallowfield, 2000).

### **2.3.2.2 Maastricht History Taking and Advice Checklist (MAAS-R)**

Ein globales Instrument zur Fremdbeurteilung von Arzt-Patient-Gesprächen, auf das in der Literatur immer wieder Bezug genommen wird, ist die Maastricht-History-Taking and Advice Checklist<sup>3</sup>. Das Instrument wurde von einer holländischen Arbeitsgruppe 1987 entwickelt, mit dem Ziel Lernende beim Erwerb medizinischer Anamnese Fertigkeiten zu unterstützen. Im Laufe der Jahre wurde es von verschiedenen Arbeitsgruppen revidiert und verändert.

Für die Autoren des Instruments bedeutete Gesprächsführungskompetenz mehr als eine reine Datensammlung, empathisches und nonverbales Verhalten wurden als wesentliche Elemente gesehen. Daher kommen psychosoziale Aspekte im Gespräch schwerpunktmäßig vor. Beispielsweise gibt es ein Item, das der Empathie besondere Beachtung schenkt. Um die Handhabung des Instrumentes zu erleichtern, gibt es ein sehr ausführliches Manual, in dem die einzelnen Items ausführlich beschrieben werden. Der Hauptzweck des Ratings wird darin gesehen, objektive Maße für den Einsatz bei der Beurteilung, für Unterricht und Weiterbildung zu erhalten.<sup>4</sup>

Das Instrument besteht aus drei Kategorien – kommunikative Fertigkeiten für einzelne Gesprächsabschnitte, allgemeine kommunikative Fertigkeiten sowie medizinische Aspekte - mit insgesamt 17 Items, die entsprechend einer Skala von 0 (= nicht vorhanden) bis 6 (= hervorragend) zu bewerten sind. Es gibt ein Item „Summarizations“ – Zusammenfassungen. Dies fand auch Eingang in die in dieser Arbeit modifizierte Version der Bewertungsskala (mAGBS-Jugendliche). Es wird gefordert, dass der Arzt während des gesamten Gespräches ausreichend zusammenfasst (Van Dalen, Prince, Scherpbier & Van Der Vleuten, 1998; van Thiel, Ram & van Dalen, 2000).

### **2.3.2.3 Die Calgary-Cambridge-Observation Guides (CCOG)**

Die Calgary-Cambridge-Observation Guides<sup>5</sup> wurden entwickelt, um Lernziele für den Unterricht in kommunikativen Fähigkeiten festzulegen. Darüber hinaus sollen sie eine evidenzbasierte Struktur für Unterricht und Prüfung dieser Fertigkeiten bieten.

<sup>3</sup> Im folgenden abgekürzt mit MAAS-R für revised Maastricht history taking and advice checklist

<sup>4</sup> “The main purpose of rating is to obtain objective measures for use in feedback and judgement, for education and assessment.” (MAAS-Manual, van Thiel)

<sup>5</sup> Die Calgary Cambridge Observation guides werden im Folgenden abgekürzt mit CCOG.

Das Instrument ist sehr ausführlich und deckt alle Teile eines Anamnesegesprächs ab: Gesprächsbeginn, Information bekommen, Strukturierung und Beziehungsaufbau. Die Reaktion des Arztes auf verbale und nonverbale Hinweise des Patienten macht ein eigenes Bewertungskriterium aus.

Ebenso gibt es in den Calgary-Cambridge Observation Guides einen Punkt „Providing structure“, der abfragt, ob der Gesprächsführende eine Struktur ins Gespräch bringt (Kurtz, S. et al., 2003). Inzwischen gibt es eine validierte deutsche Übersetzung der CCOG (Simmenroth-Nayda et al., 2008) sowie Modifizierungen des Systems für spezielle Gesprächssituationen. So wurde beispielsweise der Structured communication adolescent guide“ für Arzt-Jugendlichen-Gespräche modifiziert. Wesentliche Elemente der CCOG wurden darin beibehalten. Modifiziert wurde das Instrument jedoch vor allem inhaltlich im Bezug auf den Lebensstil eines Jugendlichen. Es wurden speziell Themen wie Beziehung, Rauchen, Alkohol, Sexualität und Familie eingefügt (Blake et al., 2005).

#### **2.3.2.4 Die Checkliste „Professionelles ärztliches Kommunikationsverhalten“**

Dieses Instrument wurde für den Einsatz in Unterricht und Evaluation kommunikativer Fertigkeiten an der Universität Wien entwickelt. Es entstand unter Einarbeitung bereits vorhandener Instrumente wie der Aufklärungsgesprächsbewertungsskala (siehe 2.3.2.6 Die Aufklärungsgesprächsbewertungsskala) und einem Kölner Evaluierungsbogen für Kommunikation<sup>6</sup>.

Das Instrument umfasst drei Bereiche, die sich wiederum aus einzelnen Items zusammensetzen. Der Bereich A deckt Gesprächsinhalte ab, Bereich B Gesprächsstil widmet sich formalen Aspekten eines Gespräches und Bereich C dient der Erfassung von Beziehungsaufbau und Kontakt. Die Beurteilung erfolgt für jedes einzelne Item je nachdem, ob das Verhalten beim Gesprächsführenden beobachtet wird oder nicht. Das Instrument wurde von den Autoren anhand von zwei Videobeispielen auf Anwendbarkeit und Testgütekriterien überprüft (Puchner-Matzner et al., 2006).

#### **2.3.2.5 Das „Breaking bad news assessment schedule“ (BAS)**

Da ein englischsprachiges Instrument, die Breaking bad news assessment schedule<sup>7</sup> als Grundlage für die in der vorliegenden Arbeit modifizierte „Aufklärungsgesprächs-

<sup>6</sup> Der Kölner Evaluationsbogen Kommunikation, entwickelt von der Klinik für Psychosomatik und Psychotherapie der Universität Köln hier aus Platzgründen nur kurz angesprochen, ist ebenfalls ein Instrument für Unterricht und Evaluation (Koerfer et al., 2008)

<sup>7</sup> Die Breaking bad news assessment schedule wird im Folgenden mit BAS abgekürzt.

bewertungsskala“<sup>8</sup> (Wand et al., 2008) diente, soll an dieser Stelle das Instrument, das 1999 von Miller et. al. vor dem Kontext zunehmenden Interesses an der Thematik des „schlechte Nachrichten Überbringens“ entwickelt wurde, vorgestellt werden. Die Autoren legten für die Entwicklung der BAS zuerst mit Hilfe einer systematischen Literaturrecherche Inhalte und Schlüsselfertigkeiten, die beim Überbringen einer schlechten Nachricht von einem Arzt gefordert werden, fest. Anschließend fassten sie Themenblöcke zusammen und entschieden sich für die endgültige Version. Die einzelnen Items sind auf einer Likert-Skala von 1 (= überhaupt nicht) bis 5 (= definitiv) zu beurteilen. Das Instrument wurde anhand von mit Simulationspatienten geführten und auf Video aufgenommenen Gesprächen von Studenten und Ärzten mit unterschiedlichem Erfahrungshintergrund auf seine Praktikabilität hin überprüft. Wie das MAAS-R und die CCOG ist die BAS ein globales Bewertungsinstrument, das aus 5 Themenblöcken mit insgesamt 23 Items besteht.

Die Autoren der BAS bedauern, dass sie dem Umgang mit Emotionen nicht mehr Raum geschaffen haben und empfehlen dies für nachfolgende Versionen (Miller, Hope & Talbot, 1999). Sie schlagen zwei Möglichkeiten vor: einmal von Patienten ein Urteil über das Kommunikationsverhalten des Arztes einzuholen und dieses dann mit der Beurteilung durch das Instrumentes zu vergleichen. Die zweite Möglichkeit ist es, Experten Gesprächsleistung bewerten zu lassen und diese Bewertungen mit dem Instrument zu vergleichen (Miller et al., 1999).

### **2.3.2.6 Die Aufklärungsgesprächsbewertungsskala (AGBS)**

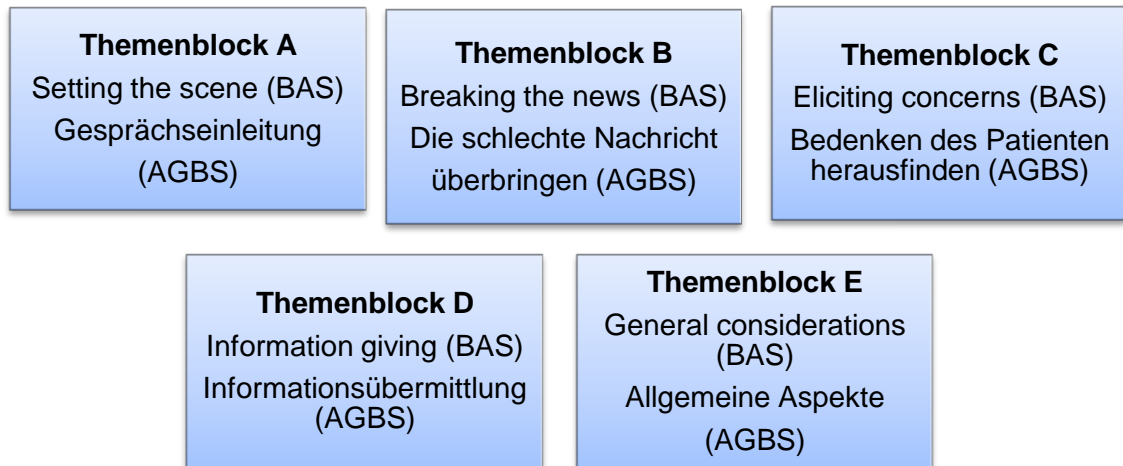
Die AGBS wurde nach Vorlage des BAS (Miller, Hope, et al. 1999) ins Deutsche übersetzt und validiert. Es gibt in Deutschland bei der Ausbildung von Studenten, aber auch bei Ärzten, die im Beruf stehen, großes Interesse an Kursen für Gesprächstraining, aber kaum gut validierte Bewertungsinstrumente. Vor diesem Hintergrund entschieden sich die Autoren zur Übersetzung und anschließenden Validierung eines englischsprachigen Instrumentes. Das Originalinstrument wurde dabei nur leicht bezüglich des Bewertungssystems verändert.

Wie die nachfolgende Abbildung zeigt, besteht sowohl die BAS als auch die AGBS aus fünf Teilbereichen mit 22 separat zu bewertenden Items genannt) die auf einer Skala von 1-5 in ähnlich dem deutschen Schulnotensystem (1 = sehr gut, 5 = sehr schlecht) bewertet werden. Die Items werden in Teilbereichen thematisch zusammengefasst. Für jedes Item gibt es wiederum Unterpunkte, die die Bewertung erklären und so das Be-

<sup>8</sup> Die Aufklärungsgesprächsbewertungsskala wiederum liegt dieser Arbeit zugrunde und wird abgekürzt mit AGBS.

werten für den Rater einfacher machen. Die Unterpunkte geben die einzelnen Kriterien an, die erfüllt sein müssen, damit ein Gesprächsführender z.B. eine sehr gute Note bekommt.

Abbildung 2 Themenblöcke von „Breaking bad news assessment scale“(BAS) und „Aufklärungsgesprächsbeurteilungsskala“(AGBS)



### 2.3.3 Zusammenfassung:

In der nachfolgenden Tabelle werden die vorgestellten Instrumente noch einmal dargestellt anhand von Anzahl der Videos, Aufwand, Gütekriterien und Zielsetzung.

Tabelle 2 Überblick und Gegenüberstellung Bewertungsinstrumente

	Probandenanzahl	Aufwand für Schulung und Kodierung	Gütekriterien	Zielsetzung
<b>RIAS</b>	Das System wird in zahlreichen Studien verwendet	6-8 Wochen Raterschulung Ca. 5-fache Gesprächszeit für Kodierung	Reliabilität Validität Objektivität wurde überprüft	Erforschung der Arzt-Patient-Interaktion auf wissenschaftlicher Basis
<b>MAAS-R</b>	Zahlreiche Studien	Raterschulung notwendig k. Angaben. zu Kodierung	Gute Werte für Reliabilität, Validität	Unterricht in Anamnese-fertigkeiten
<b>Calgary-Cambridge-Observation-Guide</b>	Beispiel: SCAG 42 Videos	Raterschulung notwendig k. Angaben zu Kodierung	4 Rater, ICC 0,93	Lernziele für Prüfung
<b>Checkliste Wien</b>	2 Videos	Schulung wenige Std.; 10-minütige Anwendung	14 Rater Gesamtübereinstimmung 75,6 %	Einsatz in Unterricht und Prüfung
<b>BAS</b>	20 Videos	3 Std. Raterschulung,	3 Rater Gesamtübereinstimmung 62 %	Schlüsselfertigkeiten beim Überbringen schlechter Nachrichten
<b>AGBS</b>	58 Videos	2-tägige Raterschulung, 15-20 minütige Anwendung	3 Rater ICC-Gesamt 0,86	Evaluation von Kommunikationstraining

Was den Zeitaspekt sowohl hinsichtlich Raterschulung als auch Anwendung und die Gütekriterien betrifft, erscheint die BAS/AGBS mit Abstand am angemessensten. Auch die Zielsetzung deckt sich gut mit dem Einsatz als Prüfungsinstrument an einer Universität.

Die AGBS lässt sich auf ein 5-minütiges Arzt-Patient-Gespräch in 15-20 Minuten nach entsprechender Schulung anwenden.



## **2.4 Kommunikation in der Kinder- und Jugendpsychiatrie**

### **2.4.1 Die Besonderheiten der Kinder- und Jugendpsychiatrie**

Deutschlandweit haben psychische Krankheiten bei Jugendlichen eine sehr hohe Prävalenz<sup>9</sup> und nehmen auch weiter zu (Ravens-Sieberer, Erhart, Wille & Bullinger, 2008). In der Pädiatrie hat es in den letzten Jahren einen Wechsel zu eher psychosozialen Problemen und Entwicklungsfragen gegeben, was wiederum den Bedarf an Ärzten, die mit solchen Fragestellungen umgehen können, erhöht (van Dulmen & Holl, 2000).

Gerade auch die interdisziplinäre Stellung zwischen Pädiatrie, Psychiatrie, Psychologie und Pädagogik macht die Kinder- und Jugendpsychiatrie zu einem besonderen Fachgebiet. Die Kinder- und Jugendpsychiatrie ist ein Fach der „sprechenden Medizin“ und stellt hohe Anforderungen an die kommunikativen Fertigkeiten der Ärzte, die in diesem Bereich tätig sind.

Ähnlich wie in der Pädiatrie gibt es auch in der Kinder- und Jugendpsychiatrie eine Dreiecksbeziehung zwischen Jugendlichen, Eltern und Arzt. Je nach Alter und Reife geben Kinder selbst über Anliegen und Beschwerden Auskunft. Sie können und müssen sogar in die Therapie einbezogen werden. Im juristischen Sinne sind Kinder und Jugendliche zwar nicht geschäftsfähig, es wird ihnen aber abhängig von der individuellen Reife Mündigkeit zugesprochen. Rechtlich gesehen benötigen Ärzte bereits bei einer Befunderhebung das Einverständnis der Eltern oder des Sorgeberechtigten (Schmidt, 1998).

Es gibt jedoch Situationen z.B. bei Missbrauchs- oder Vernachlässigungsfällen, in denen der Arzt zum Wohl des Jugendlichen gegen die Eltern handeln muss. Langer schreibt in seinem Artikel „Das Gespräch zwischen Patient, Eltern und Arzt“ von einem Spannungsverhältnis in der Pädiatrie je älter die Kinder werden. Das psychotherapeutische Arbeiten mit Jugendlichen und deren Eltern bzw. Sorgeberechtigten beinhaltet bei Interessenkonflikten sehr komplexe und manchmal problembehaftete Konstellationen (Langer, 2009).

### **2.4.2 Gespräch mit Jugendlichen**

Bei allen Ärzten, die mit Jugendlichen in Kontakt treten wollen, muss ein gewisses Ver-

<sup>9</sup> Erste Ergebnisse der Bella-Studie ,welche großangelegt die psychische Gesundheit von Kindern und Jugendlichen untersucht, zeigen, dass 5 % aller Kinder und Jugendlichen behandlungsbedürftig und 13 % auffällig sind.

ständnis<sup>10</sup> für die Pubertät und Entwicklungsprozesse in diesem Alter vorhanden sein. Auf einen Jugendlichen kommen viele wichtige Entwicklungsaufgaben zu. Die Entwicklung eigener Wertvorstellung und einer eigenen Identität bedingt eine schrittweise Ablösung vom Elternhaus, wobei es zu einer Orientierung an Gleichartigen kommt. Freundschaft und Partnerschaft werden für Jugendliche sehr wichtig. Die körperliche Reifung führt zu einem verstärkten Interesse an Sexualität. Zugleich befinden sich die Jugendlichen auch was ihre Ausbildung betrifft, in einer wichtigen Phase. Durch schulische Leistungen und die Entscheidung für einen Beruf werden die Weichen für das spätere Leben gestellt. Es ist nicht verwunderlich, dass die Adoleszenz gerade vor diesem „Berg“ an zu bewältigenden Aufgaben ein hohes psychisches und soziales Konfliktpotential bietet. Identitätskrisen und Konflikte vor allem mit Eltern, Gesellschaft, Schule können dabei oft als Normvariante gesehen werden, gleichzeitig beinhalten sie aber immer die Gefahr starker psychischer Fehlentwicklungen. Jede Krise kann zu einem Notfall werden und sich ins Erwachsenenalter fortsetzen. Jugendliche sind gefährdet für Süchte, sei es Substanzmissbrauch oder Essstörungen, mögliche Probleme ergeben sich auch in emotionalen Störungen wie Aggressionen, Zwänge, Ängste und Depressionen bis hin zu suizidalem Verhalten (Remschmidt, 1992).

Jede Behandlung von Jugendlichen, die wegen physischen oder psychischen Problemen zu einem Arzt oder Therapeuten kommen, beginnt mit einem Gespräch.

Ein Besuch beim Arzt ist für einen Jugendlichen, der häufig Schwierigkeiten hat, sich einem Erwachsenen gegenüber zu öffnen und von seinen Problemen zu erzählen, etwas wenig Vertrautes. Von einem Arzt werden spezielle Fähigkeiten erwartet, um eine vertrauensvolle Beziehung zu einem Jugendlichen aufbauen zu können (Hardoff, Schonmann 2001). Remschmidt fordert als Grundvoraussetzung bei einem Arzt Offenheit und eine freundliche, aber neutrale Gesprächshaltung (Remschmidt, 1992). Dabei gilt zu beachten, dass gerade Jugendliche, die mit Problemen konfrontiert werden ein besondere Verhaltensmuster zeigen. Ein Teil nimmt Probleme hypochondrisch wahr, der Großteil Jugendlicher neigt aber dazu, Probleme abzuwiegeln oder gar zu verleugnen. Hilfsangebote werden oft aus Scham und Abneigung gegenüber Ärzten nicht in Anspruch genommen.

Jugendliche kommen selten freiwillig, sondern werden von den Eltern, Schule oder Jugendamt geschickt. Vor diesem Hintergrund kommt dem Erstkontakt eine große Bedeutung zu. Dem Untersucher muss es gelingen, einen Bezug zum Jugendlichen her-

<sup>10</sup> Remschmidt bedauert, dass die Probleme der Adoleszenz in verschiedenen klinischen Disziplinen vernachlässigt werden. (Remschmidt, 1992)

zustellen, indem er Interesse an seiner Person und seinem Lebensumfeld zeigt und sich ihm vorsichtig annähert. Doch der Balanceakt zwischen den Bedürfnissen der Jugendlichen, den Interessen der Eltern und auch anderen an der Lebenswelt beteiligter Strukturen wie Schule, Arbeitsplatz oder Jugendamt ist schwierig und bedarf einer hohen Kompetenz des Arztes. Fragt man Jugendliche, was ihnen bei einem Arzt wichtig ist, steht ganz oben der Wunsch, dass ihre Probleme vertraulich behandelt werden. Außerdem erwarten sie respektvolle Behandlung, Zeit und offene Gesprächsmöglichkeit über Gesundheitsthemen (Cottrell, Nield & Perkins, 2006). Es hat sich bewährt, Gespräche getrennt nur mit dem Jugendlichen, den Eltern, aber auch in gemeinsamen Familiengesprächen zu führen (Remschmidt, 1992).

Nachdem im englischsprachigen Raum schon vor einigen Jahren erkannt wurde, wie wichtig spezielle Gesprächsfertigkeiten im Umgang mit Jugendlichen für angehende Ärzte sind (Blake et al., 2005; Brown, Doonan & Shellenberger, 2005; Hardoff & Schonmann, 2001), wird dem Gespräch mit Kindern und Jugendlichen auch in Deutschland mehr Aufmerksamkeit geschenkt (Koch & Resch, 2009).

Zusammengefasst kann man sagen, dass ein Arzt in einem Gespräch mit einem Jugendlichen all das einbringen soll, was auch in jedem anderen Arzt-Patient-Gespräch wichtig ist: eine angemessene Mischung aus offenen und geschlossenen Fragen, Vermeidung von Suggestivfragen, aktives Zuhören, verständliche Sprache und Empathie. Gerade deswegen eignet sich diese Gesprächssituation hervorragend für angehende Ärzte, um Gesprächsfertigkeiten zu üben.

### **2.4.3 Notfallsituation Suizidalität**

Suizide gehören im Jugendalter mit zu den häufigsten Todesursachen. Suizidversuche kommen noch häufiger<sup>11</sup> vor. Auch Suizidgedanken und –phantasien sind bei Jugendlichen häufig (Braun-Scharm & Poustka, 2003). Folgende Ursachen für einen Suizidversuch sind in der Altersgruppe der 15-25-jährigen von Bedeutung: Familiäre Konflikte, schulische oder berufliche Probleme und Partnerkonflikt oder –verlust, sowie Suchtprobleme und Depressionen.

Auch bei psychischen Krankheiten und Persönlichkeitsstörungen, wie dem Borderline-Syndrom kommen Suizide und Suizidversuche gehäuft vor (Remschmidt, 1992). Von zentraler Bedeutung im Umgang mit einem Jugendlichen nach einem Suizidversuch ist neben der Sicherung der vitalen Funktionen die Herstellung eines Gesprächskontakts.

<sup>11</sup> Hier gibt es einen Geschlechtsunterschied, vollendete Suizide sind bei Jungen häufiger, bei Mädchen ist die Zahl der Suizidversuche höher. Als Angabe für die Häufigkeit von Versuchen finden sich Zahlen von 127/100000 bei Jungen und 367/100000 bei Mädchen. Die Dunkelziffer wird aber um einiges höher geschätzt. (Braun-Scharm & Poustka, 2003)

Dabei gibt es einige grundlegende Dinge zu beachten. Der Arzt sollte sich dem Jugendlichen vorstellen, den Jugendlichen mit Namen anreden, wenn möglich einen geeigneten Raum für das Gespräch suchen und sich hinsetzen. Er muss neben allgemeinen Fragen zur Person und Lebenssituation den Suizidversuch thematisieren. Ziel dieses ersten Gespräches ist es, eine Einschätzung über eine weitere bestehende Gefährdung des Jugendlichen zu bekommen. Es ist zu klären, inwieweit der Jugendliche sich vom Suizidversuch distanzieren kann. Dem Jugendlichen müssen Hilfe und ein Plan für weiteres Vorgehen angeboten werden (Frank, 2005). Mögliche Fehlerquellen im Umgang mit einem Patienten sind dabei ein Unterschätzen des Suizidrisikos und eine rein somatische Versorgung ohne Ansprechen des Suizidversuchs (Schmidt, 1998).

### **2.4.4 Umgang von Medizinstudenten und Ärzten mit Suizidalität**

Die meisten Patienten kommen nach einem Suizidversuch zunächst in die Notaufnahme eines Krankenhauses. Dort beschränkt man sich häufig auf die somatische Behandlung und gibt die psychische Versorgung an einen psychiatrischen Kollegen ab (Geisler, 2002).

„Selbstmörder zählen bei Ärzten und Schwester zu den unbeliebtesten Patienten“, so bringt es Linus Geisler in seinem Buch „Arzt und Patient-Begegnung im Gespräch“ auf den Punkt. Es fällt vielen Ärzten einerseits schwer, den Suizidversuch ernst zu nehmen, andererseits ist auch sehr viel Unsicherheit und Scheu vorhanden. Diese Tabuisierung und Abschiebung wiederum verstärkt die emotionale Abkapslung der Patienten. Offene Kommunikation wird blockiert. Geisler sieht eine wesentliche Ursache dieser Missstände darin, dass der Umgang mit Suizidpatienten im Medizinstudium kaum gelehrt wird (Geisler, 2002).

## **2.5 Gesprächstraining für Medizinstudenten an der Ludwigs-Maximilians-Universität-München**

### **2.5.1 Pflichtveranstaltungen im Lehrplan**

An der Ludwigs–Maximilians–Universität in München wurden in den letzten Jahren vermehrt Kurse zur ärztlichen Gesprächsführung eingeführt.

Für die Studenten des 10. Semesters wird von den Mitarbeitern der Kinder- und Jugendpsychiatrie und der Pädiatrie eine Pflichtlehrveranstaltung zum Thema „Überbringen schlechter Nachrichten“ angeboten. Hierbei erhalten die Studenten die Aufgabe, in videogestützten Rollenspielen entweder die Rolle des Arztes oder der Eltern bei der Mitteilung einer schwerwiegenden Erkrankung des Kindes zu übernehmen (Kopecky-Wenzel et al., 2009).

### **2.5.2 Das Wahlfach Kinder- und Jugendpsychiatrie an der Ludwigs-Maximilians-Universität**

Unter Medizinstudenten ist Kinder- und Jugendpsychiatrie, unter anderem weil es kein Pflichtfach im Lehrplan ist, relativ unbekannt (Warnke & Lehmkuhl, 2003).

An der Ludwigs-Maximilians-Universität wird Kinder- und Jugendpsychiatrie im Rahmen des Modul 5 „Gezeiten und Lebensabschnitte“ als viertägiges Wahlfach „Kinder- und Jugendpsychiatrie und Psychotherapie“ angeboten. Die Teilnehmer, überwiegend Studenten des 10. Semesters, erhalten hierdurch einen Einblick in das Fachgebiet. Die Veranstaltung spricht vor allem Studenten an, die sich für psychosoziale Themen interessieren. Die Teilnehmerzahl ist auf maximal 12 Teilnehmern zweimal pro Semester begrenzt. So ist es sehr gut möglich, viele praktische Übungen mit einzubauen, und auch kommunikative Fähigkeiten im geschützten Rahmen der Kleingruppe einzuüben. Dies wird durch den gezielten Einsatz von Videomaterial und Kamera unterstützt. Anhand von Videoaufnahmen von Patienten bekommen die Studenten dabei zunächst einen Zugang zu den „typischen Krankheitsbildern“ der Kinderpsychiatrie wie hyperkinetisches Syndrom, Störung des Sozialverhaltens, aber auch Ängste, Sucht und Depressionen bis hin zu suizidalem Verhalten werden thematisiert. Dabei lernen die Studenten durch genaues Beobachten der Videosequenzen, einen psychopathologischen Befund zu erstellen. Nach der Vermittlung des theoretischen Hintergrunds sollen sich die Studenten unter Anleitung der Dozenten damit auseinandersetzen, wie sie selbst ein ärztliches Gespräch mit einem Kind, einem Jugendlichen oder auch den Eltern führen würden.

Hierbei dienen die davor gesehenen Aufnahmen als eine Art „Drehbuch“. Gemeinsame

Grundlage und Bezugspunkt für die folgenden Gespräche sind die zuvor gesehenen Videoaufnahmen.

In einer Stoffsammlung werden die Themen, die im Gespräch mit einem Kind, später dann im Gespräch mit einem Jugendlichen, in der nächsten Stufe mit Eltern eines entwicklungsverzögerten Kindes und schließlich am Ende in der Notfallsituation mit einem Jugendlichen nach einem Suizidversuch von Belang sind, zusammengetragen.

Der Schwierigkeitsgrad steigt im Laufe des Kurses an. Nachdem die Studenten inhaltlich auf die Gesprächsführung vorbereitet wurden, übernehmen sie für 5-minütige Rollenspiele entweder die Rolle des Arztes oder des Patienten. Die Studenten, die die Rolle der Patienten übernehmen, bekommen den Auftrag, sich so gut wie möglich in die Situation eines Kindes, Jugendlichen oder Elternteils hineinzusetzen. Danach sollen sich die Studenten selbst einschätzen und bekommen auch eine Rückmeldung der Beobachter. Sie werden vom Kursleiter gebeten, zu schildern, wie sie sich im Gespräch gefühlt haben, was sie gut gemacht haben und was sie schwierig fanden. Die Beobachter geben ebenfalls eine Rückmeldung. Die Hauptaufgabe an die Gesprächsführenden lautet dabei stets „ins Gespräch kommen“, die Studenten sollen vor allem Basisfertigkeiten der Kommunikation lernen und ihre Scheu auch bei schwierigen Themen wie Drogen, Alkohol, Depression sowie Suizidgedanken verlieren. Als Basisfertigkeiten der Gesprächsführung werden unter anderem Begrüßung, aktives Zuhören und das Stellen offener Fragen vermittelt. Die folgende Tabelle gibt noch einmal einen Überblick über die Gesprächssituation und die Ziele, die erreicht werden sollen (Frank, Gegenfurtner, Steininger, Kopecky-Wenzel & Noterdaeme, 2009).

Tabelle 3 Lernziele Wahlfach "Kinder- und Jugendpsychiatrie"

<b>Situation</b>	<b>Gesprächsziel</b>	<b>Themen</b>
<b>Kind (6-7 Jahre)</b>	Mit dem Kind in Kontakt und ins Gespräch kommen	Interessen, Freundschaften, Familie, Schulsituation, Schwierigkeiten
<b>Jugendlicher (14-Jahre)</b>	<b>Ins Gespräch kommen, auch bei schwierigen Themen im Gespräch bleiben</b>	<b>Interessen, Situation zu Hause, Schule, Freundschaften, Nikotin, Alkohol, Perspektiven</b>
<b>Eltern eines Kindes mit Entwicklungsrückstand</b>	Umgang mit einer Mehrpersonensituation, Übereinstimmung finden, Basis schaffen	Eltern den psychischen Befund ihres Kindes vermitteln, Erfahrungen und Beobachtungen beider Eltern erfragen
<b>Jugendlicher (14-Jahre) nach Suizidversuch</b>	<b>Ins Gespräch kommen, Beziehungsaufbau</b>	<b>Erfragen der Lebenssituation, Belastungsfaktoren (Sucht, Depression) und Ressourcen erarbeiten, konkrete Perspektive</b>

## **2.6 Gesprächstraining für Ärzte mit Berufserfahrung**

Der „Kompaktkurs Pädiatrie“ richtet sich an approbierte Ärzte, die sich in ihrer Ausbildung zum Allgemeinmediziner befinden. Das für die Facharztqualifikation vorgeschriebene Jahr Weiterbildung in der Kinderheilkunde kann durch eine entsprechende Weiterbildung in einem Gebiet mit Patientenbezug ersetzt werden. Zusätzlich dazu müssen die Ärzte an einem von der Ärztekammer anerkannten 80-stündigen Kurs mit kinderärztlichen Inhalten teilnehmen. Schwerpunktthemen des Kompaktkurses sind die kindliche Entwicklung mit besonderem Bezug auf die Vorsorgeuntersuchungen, Allgemeinsymptome und Notfälle, sowie Impfungen und Allergien.

Ein weiterer Baustein des Kurses ist eine dreistündige Einheit zur Gesprächsführung. Die Teilnehmer erarbeiten sich zunächst die Themen, die sie mit einem Kind oder einem Jugendlichen bei den jeweiligen Vorsorgeuntersuchungen besprechen sollten. Danach übernehmen sie in Rollenspielen in Gruppen von ca. 20 Personen, die Arzt- oder Patientenrolle.

Ziel ist auch hier, genau wie im Wahlfach Kinder- und Jugendpsychiatrie, mit dem Patienten ins Gespräch zu kommen. Die Rollenspiele werden gefilmt. Anschließend erhalten die Spielenden Rückmeldung durch die Zuschauer. Die Videoaufnahmen ermöglichen eine unmittelbare Überprüfung. Exemplarisch soll dies am Beispiel der Jugendgesundheitsuntersuchung J1 gezeigt werden. Die Jugendgesundheits-Untersuchung wurde 1998 mit dem Ziel einer besseren Prävention vor allem von Sucht, Essstörungen und anderen psychischen Auffälligkeiten eingeführt. 25 Prozent der Untersuchungen werden von Allgemeinärzten durchgeführt. (Weissenrieder N. & Stier, 2006).



Tabelle 4 Lernziele im Kompaktkurs Pädiatrie

<b>Situation</b>	<b>Gesprächsziel</b>	<b>Themen</b>
<b>Jugendlicher zur Jugendgesundheitsuntersuchung</b> <b>J1 (12-15 Jahre)</b>	<b>Ins Gespräch kommen,</b> <b>offene Fragen stellen</b> Beratung, Vermittlung Gesundheits-spezifischer Themen	Gesundheitsverhalten Interessen, Sport, Hobbys, Schule, Ausbildung, Freundschaften, Partnerschaft, Sexualität, Alkohol, Drogen

### **3 Zielsetzung**

Die vorliegende Arbeit hat zum Ziel, anhand von Videoaufnahmen von Gesprächsübungen eine modifizierte globale Bewertungsskala auf ihre Anwendbarkeit für Unterricht und Prüfung im Kontext psychosozialer Fragestellungen zu überprüfen.

Ergänzend werden im Anschluss die verschiedenen Gesprächsverläufe der Videoaufnahmen mittels Interact dargestellt. Dabei werden einzelne wichtige Sequenzen aus den Aufnahmen herausgegriffen und in Bezug zur mAGBS-Jugendliche gesetzt.

Im Fokus des Interesses steht die Frage, welche Folgen Aussagen und Verhaltensweisen wie Gesprächspausen, das Stellen offener und geschlossener Fragen, Lücken und Füllwörter für den Gesamtverlauf des Gespräches haben.

Die Arbeit besteht aus zwei Teilen. Der erste Teil umfasst die Weiterentwicklung und Reliabilitätstestung einer bereits validierten Bewertungsskala zur Evaluation der Arzt-Patient-Kommunikation.

Der zweite Teil befasste sich ausgehend von den Ergebnissen der Reliabilitätstestung der modifizierten Bewertungsskala mit Fragen zu Kommunikationsstrukturen in der Arzt-Patient-Beziehung. So wurde der Frage nachgegangen, welchen Einfluss Redezeiten haben, welche Wirkung Gesprächspausen und Gesprächstempo haben und durch welchen Fragestil ein Patient zum Reden angeregt werden kann.

Die Autorin dieser Arbeit entschied sich, Videos von Trainings zur Gesprächsführung mit Jugendlichen, zur Notfallsituation und Videos zur Gesundheitsuntersuchung bei Jugendlichen zur Bewertung und Reliabilitätstestung der Bewertungsskala und auch für die Analyse der Gesprächsverläufe zu verwenden.

Tabelle 5 Studiendesign

<b>Trainingssituation</b>	Gespräch mit Jugendlichen	Gespräch in Notfallsituation (Suizidversuch)	Allgemeinärzte bei der Jugendgesundheitsuntersuchung
<b>Ausbildungshintergrund</b>	Medizin-Studenten	Medizin-Studenten	Angehende Allgemeinärzte
	<b>Effekt des Trainings</b>		<b>Berufserfahrung</b>
<b>Zahl der Videos</b>	11	11	10

Der Vorteil in der Auswahl der Videos lag darin, dass es auf der einen Seite immer um das Gespräch mit einem Jugendlichen ging. Dadurch war die Gemeinsamkeit des Themas gegeben. Auf der anderen Seite kam aber hinzu, dass gleichzeitig verschiedene Schwierigkeitsstufen der Gesprächssituation sowie verschiedene Ausbildungshintergründe der Gesprächsführenden vorlagen. So konnte neben einer Reliabilitätstestung auch überprüft werden, inwiefern das Bewertungsinstrument verschiedenen Situationen und Ausbildungsstufen gerecht wird.

**Der Schwerpunkt liegt bei der Fragestellung, wie und ob man gute bzw. schlechte Gesprächsführung objektivierbar beurteilen kann und ob sich Muster in der Gesprächsführung erkennen lassen.**

In diesem Zusammenhang stellen sich folgende Fragen:

- Welche Ergebnisse erzielt das Instrument hinsichtlich der Reliabilität?
- Lässt sich das Instrument auf die Gruppen mit verschiedenen Rahmenbedingungen anwenden?
- Bildet es einen Unterschied zwischen den Gruppen ab?

Hierbei sollen die verschiedenen Einflussfaktoren für eine unterschiedliche Bewertung diskutiert werden

- Wie realistisch sind Rollenspiele?
- Wie ist der Einfluss des Geschlechtes auf die Gesprächsführung zu bewerten?
- Welchen Effekt hat ein Training auf Gesprächsfertigkeiten?
- Was bringt Berufserfahrung?

## 4 Material und Methoden

### 4.1 Auswahl der Videos

Die Videoaufnahmen stammen aus dem Wahlfach „Kinder- und Jugendpsychiatrie“ sowie aus dem Kompaktkurs „Pädiatrie für Allgemeinmediziner“ (siehe hierzu 2.4 Gesprächstraining an der Ludwigs-Maximilians-Universität und 2.5 Kompaktkurs „Pädiatrie für Allgemeinmediziner“).

Da aus dem Kompaktkurs „Pädiatrie“ 10 Videos nach Anwendung der Ausschlusskriterien (siehe 4.1.3 Ausschlusskriterien) zur Verfügung standen, wurden um eine vergleichbare Anzahl auch aus den anderen Gruppen zu haben, jeweils 11 Videos zum Gespräch mit Jugendlichen und 11 zur Notfallsituation ausgewählt.

Um eine bessere Übersicht zwischen den verschiedenen Videogruppen zu bewahren, werden im Folgenden die 11 Videos aus dem Gespräch mit Jugendlichen als „**Jugendlichenvideos**“ bezeichnet, die 11 Videos zur Notfallsituation als „**Notfallvideos**“, sowie die 10 Videos aus dem Kompaktkurs „Pädiatrie“ als „**Allgemeinärztevideos**“.

Es lagen insgesamt 32 Videos vor, die für die Anwendungstestung der Bewertungsskala (m-AGBS-Jugendliche) genutzt werden konnten.

Die Videos wurden chronologisch rückwärts ausgewählt – es wurde mit den neusten Videos begonnen, da sich sowohl das Wahlfach als auch der Kompaktkurs durch Rückmeldungen der Teilnehmer verändert hatten. In den neueren Kursen waren die Zielvorgaben für die Gesprächsführung inhaltlich klarer vorgegeben. Parallel dazu konzentrierte sich ein weiterer Doktorand auf die Auswertung von Elterngesprächen, die ebenfalls im Wahlfach Kinder- und Jugendpsychiatrie entstanden waren.

Das Wahlfach Kinder- und Jugendpsychiatrie wird an der Ludwigs-Maximilians-Universität seit dem Wintersemester 2005/2006 angeboten. Da der Kurs zweimal pro Semester stattfindet und von Anfang an alle Gespräche gefilmt wurden, lagen bis zum Beginn des Ratings im Januar 2008 insgesamt 48 Videos von Gesprächen mit Jugendlichen vor. In 8 Fällen lag keine Einwilligung vor, die Videos für Lehr- oder Forschungszwecke weiter zu verwenden.

Vom Kompaktkurs „Pädiatrie für Allgemeinmediziner“ lagen 20 Videos im Zeitraum von 2004 - 2008 vor. Die Einwilligung war in 4 Fällen nicht gegeben. Bei zwei weiteren war kein Ton vorhanden. 4 Videos waren zu Trainingszwecken eingesetzt worden und wurden deswegen ausgeschlossen.

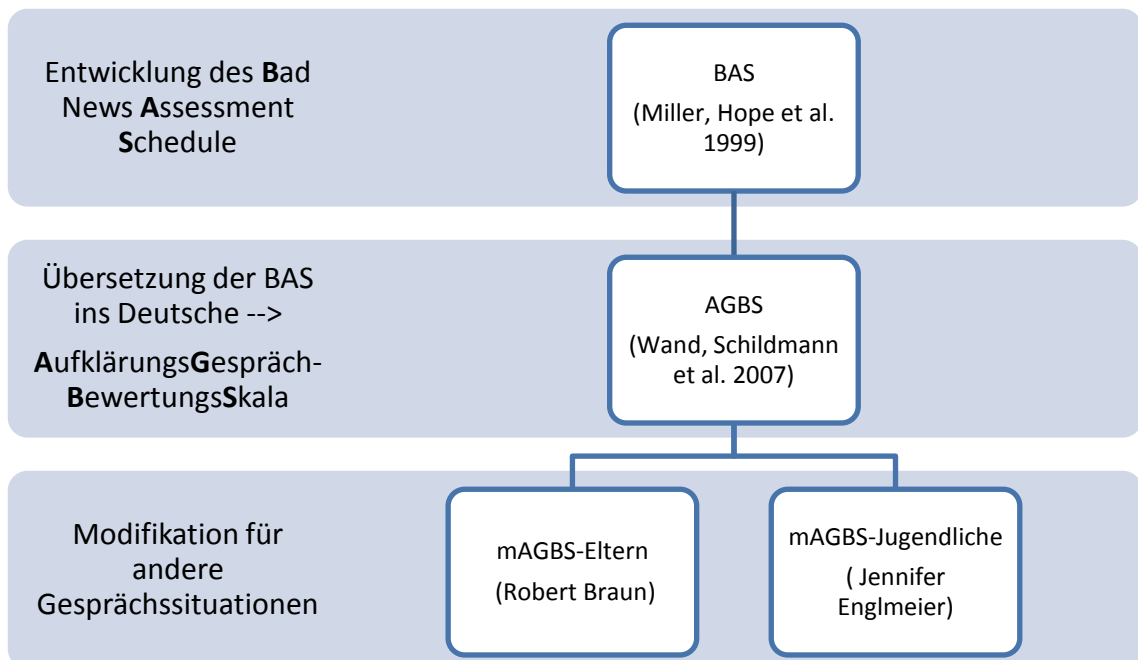
## 4.2 Beurteilungsinstrumente

### 4.2.1 Modifizierte Aufklärungsgesprächsbewertungsskala – mAGBS-Jugendliche

#### 4.2.1.1 Auswahl und Entwicklung der mAGBS-Jugendliche

Als Grundlage für die Entwicklung des in dieser Studie verwendeten Instruments, diente eine 2008 publizierte „Aufklärungsgesprächsbewertungsskala“ (AGBS) (Wand et al., 2008). Die Aufklärungsgesprächsbewertungsskala wiederum wurde aus dem Bad News Assessment Schedule (Miller et al., 1999) entwickelt. Siehe dazu die folgende Abbildung.

Abbildung 3 Entwicklung der Bewertungsskala



Um die AGBS im Kontext des Arzt-Jugendlichen-Gesprächs anwenden zu können, mussten vor allem inhaltliche Änderungen vorgenommen werden, da es in den Gesprächen mehr um das Erhalten von Information - also im weitesten Sinne ein Anamnesegespräch – als um das Überbringen einer schlechten Nachricht geht.

Mithilfe einer Literaturrecherche und Orientierung am Wahlfach wurden Inhalte, die in einem Arzt-Jugendlichen-Gespräch auf jeden Fall vorkommen sollten, festgelegt (siehe hierzu 2.3 Kommunikation in der Kinder- und Jugendpsychiatrie sowie 2.4 Gesprächstraining an der Ludwigs-Maximilians-Universität).

#### 4.2.1.2 Allgemeine Modifikationen

Alle Items der AGBS wurden so umformuliert, dass nicht mehr nach dem Kontakt des Arztes zum Patienten oder der Reaktion des Patienten gefragt wurde, sondern der Jugendliche als Bezugsperson des Arztes benannt wurde. Die folgende Tabelle zeigt Beispiele hierfür.

Tabelle 6 Allgemeine Modifikationen in den Items

AGBS	mAGBS - Jugendliche
<p>Item 2:</p> <p>Hat der Arzt <b>den Patienten</b> angemessen begrüßt und sich vorgestellt?</p>	<p>Item 2:</p> <p>Hat der Arzt <b>den Jugendlichen</b> angemessen begrüßt und sich vorgestellt?</p>
<p>Item 3:</p> <p>Hat sich der Arzt zu Beginn des Gesprächs nach dem momentanen Befinden und den Lebensumständen <b>des Patienten</b> erkundigt?</p>	<p>Item 7:</p> <p>Hat sich der Arzt zu Beginn des Gesprächs nach dem momentanen Befinden <b>des Jugendlichen</b> erkundigt?</p>

Teil B „Die schlechte Nachricht überbringen“ musste vollständig geändert werden, da es in der Gesprächssituation zwischen Arzt und Jugendlichem keine schlechte Nachricht zu überbringen gab.

Deshalb wurde Teil B in „Information bekommen“, angelehnt an die Calgary-Cambridge-Observation.-Guides „Gathering information“, (Kurtz, S. M. & Silverman, 1996) umgewandelt. Einzelne Items wurden jedoch aus den AGBS beibehalten. Teil D „Informationsübermittlung“ wurde umgewandelt in „Perspektive vermitteln“, da die Vermittlung von gesundheitsrelevantem Verhalten zwar Bestandteil des Gespräches war, mehr aber noch die Vermittlung eines weiteren Vorgehens im Vordergrund stand.

#### 4.2.1.3 Inhaltlich geänderte Items

Im Folgenden soll genauer auf die einzelnen inhaltlichen Änderungen eingegangen werden.

Dabei wird zunächst kurz das Originalitem aufgeführt, dann werden die konkreten Änderungen aufgezeigt.

**AGBS – Abschnitt A Gesprächseinleitung – Item 1:** *Wie gut ist es dem Arzt gelungen, eine geeignete Umgebung für das Gespräch zu schaffen?*

Die AGBS bewertet hier nach den drei Kriterien:

- a) Der Arzt hat die Stühle so platziert, dass Arzt und Patient über Eck am Tisch sitzen können.
- b) Der Arzt hat die Kaffeetassen und Unterlagen auf dem Tisch geordnet/beiseite geräumt.
- c) Der Arzt hat die Taschentücher in Reichweite hingestellt.

Für die m-AGBS-Jugendliche erschien der explizite Verweis auf Kaffeetassen und Taschentücher nicht adäquat, da sie auch der Rollenspielsituation nicht gerecht werden, deshalb wurden diese drei Kriterien ersetzt durch:

- a) Der Arzt hat für ausreichend Sitzgelegenheit für den Jugendlichen gesorgt
- b) Der Arzt hat den Tisch geordnet bzw. leergeräumt;
- c) Der Arzt wirkt gepflegt, ordentlich und angemessen gekleidet.

**AGBS – Abschnitt A– Item 3:** *Hat sich der Arzt zu Beginn des Gespräches nach den momentanen Befinden und den Lebensumständen des Patienten erkundigt?*

Die AGBS fragen hier zusätzlich nach den Lebensumständen.

In den m-AGBS-Jugendliche wurden die Einzelkriterien wie in den AGBS beibehalten, es wurde lediglich das Stichwort Lebensumstände gestrichen, da es gerade darum im weiteren Gespräch zwischen Arzt und Jugendlichen gehen soll.

**AGBS - Abschnitt B – Item 4:** *Hat der Arzt erfragt, wie viel der Patient bereits weiß, bevor er die Nachricht überbracht hat?*

Dieses Item wurde wie der ganze Block B, in dem sich dieses Item befindet, inhaltlich abgewandelt.

Es wurde aber versucht, den Sinn von Item 4 beizubehalten, nämlich in wieweit es dem Gesprächsführenden gelingt, die Perspektive des Patienten vor dem vertieften Einstieg ins Gespräch zu erfassen. Es wird vom Arzt gefordert, sich auf die individuelle Situation des Patienten einzulassen und ihn dort abzuholen, wo er steht.

Das Item wurde umgeformt in die Frage „Ist der Arzt mit dem Jugendlichen ins Gespräch gekommen?“ Dabei wurde aber nur der Unterpunkt

- a) der Arzt hat den Patienten gefragt, wie viel er über seinen Krankheitszustand/ die Diagnose weiß, ersetzt durch: „Ist der Arzt mit dem Jugendlichen ins Gespräch gekommen?“

Die anderen beiden Kriterien wurden übernommen.

### **AGBS – Abschnitt B**

**Item 5:** *Hat der Arzt die schlechte Nachricht vorsichtig eingeleitet („Warnschuss“), bevor er sie ausgesprochen hat und sich anschließend an den Bedürfnissen des Patienten orientiert?*

Dieses Item wurde umgeformt in die Frage *„Hat sich der Arzt dem Jugendlichen langsam angenähert, bevor er zu persönlichen Themen übergegangen ist.“*

Es wurde als wichtiges Kriterium erachtet, dass der Arzt nicht übergangslos intime und persönliche Fragen wie z.B. nach Beziehungen, Sexualität und Suchtverhalten stellt, sondern sich dem Patienten langsam annähert und sich sensibel dafür zeigt, dass ein Jugendlicher sich erst langsam öffnet.

Entsprechend wurden die Unterpunkte geändert.

Der Arzt hat

- a) allgemeine Fragen zu alltagsrelevanten Themen gestellt
- b) dem Jugendlichen Raum gegeben, von sich zu erzählen
- c) sich anschließend an den Bedürfnissen des Jugendlichen orientiert

**Item 7:** *Hat der Arzt die Nachricht in einem angemessenen Tempo und an die Aufnahmefähigkeit des Patienten angepasst überbracht?*

**Item 8:** *Hat der Arzt nach dem Überbringen der Nachricht eine angemessene Pause gelassen?*

Diese beiden Items der AGBS wurden in den MAGBS-Jugendlichen inhaltlich zu einem zusammengefasst, da es wiederum nicht um das Überbringen der „schlechte Nachricht“ ging.

### **AGBS - Abschnitt C**

**Item 10:** *Hat der Arzt versucht, die Sorgen/Bedenken des Patienten zu erfassen?*

Auch bei diesem Punkt wurde in den m-AGBS-Jugendliche versucht, die Kernaussage, ob es dem Arzt gelingt, die wichtigsten Anliegen des Patienten zu erfassen, zu erhalten.

So lautet die entsprechende Frage im m-AGBS-Jugendliche-Abschnitt-C- Item 8: *Hat der Arzt genauer nachgefragt, wenn der Jugendliche Problempunkte angesprochen hat?*



### **AGBS- Abschnitt D**

**Item 12:** *Waren die vom Arzt übermittelten Informationen den vom Patienten genannten Fragen bzw. Bedenken angepasst?*

Auch dieses Item wurde leicht modifiziert, da ja in dem Gespräch zwischen Arzt und Jugendlichen, nicht die Informationsvermittlung vordergründig war. Das abgeänderte Item bestand darin, zu bewerten, ob die Äußerungen des Arztes situationsgemäß waren.

### **AGBS- Abschnitt E**

**Item 14:** *Hat der Arzt es geschafft, das Augenmerk auf positive Aspekte zu lenken?*

Das Item wurde stehen gelassen, aber die Einzelkriterien wurden abgeändert und zwar wurde Kriterium

a) die möglichen „Therapieoptionen auf positive Art dargestellt“ umgewandelt in:

„dem Jugendlichen eine positive Rückmeldung gegeben“

b) „eine gute Balance zwischen der Darstellung von Vorteilen und unerwünschten Wirkungen eingehalten wurde“ umgewandelt in:

„eine gute Balance zwischen der Darstellung von Dingen, die gut laufen und Dingen, die geändert werden müssen, gegeben“

c) „es geschafft, korrekte Informationen über die Prognose zu geben, ohne die Hoffnung zu nehmen“, wurde belassen.

**Item 16:** *Wie viele der folgenden Schlüsselbereiche wurden besprochen?*

Auch dieses Item wurde vom Grundgedanken beibehalten, wobei die Schlüsselbereiche an die Themen Jugendlicher angepasst wurden. Schlüsselbereiche:

Behandlung - Prognose – Gefühle und Emotionen – Aspekte von Familie- und Partnerschaft - Auswirkung auf die soziale Stellung

wurde durch folgendes ersetzt:

Schule – soziales Umfeld - familiäre Situation – Risikoverhalten – psychisches Befinden

### **AGBS - Abschnitt C**

**Item 17:** *Wurden psychosoziale Aspekte, die der Patient während des Gesprächs verbal signalisierte, erörtert?*

Dieses Item, das den Schlüsselbereichen direkt angeordnet ist, wurde dahingehend umformuliert, inwiefern das Kriterium der weiterführenden Besprechung eines der Schlüsselbereiche hinreichend erfüllt ist. Da sich das ganze Gespräch zwischen Arzt

um Jugendlichen um psychosoziale Aspekte dreht, schien es überflüssig, die einzelnen Aspekte zu beleuchten, das sollte Item 16 erfüllen, aber es wurde als wichtiges Kriterium erachtet, inwiefern sich der Gesprächsführende tiefer auf einzelne Schlüsselbereiche einlässt.

### 4.2.1.4 Entfernte Items

Folgende Items schienen trotz Umformulierung und Veränderung nicht auf die Gesprächssituation Arzt-Jugendlicher übertragbar.

#### **AGBS - Abschnitt C**

**Item 9:** *Hat der Arzt den Patienten eingeladen, Fragen zu stellen?*

Dieser Punkt wurde als eigenes Item gestrichen und in den m-AGBS-Jugendlichen dem Item 10-Abschnitt C als ein Unterpunkt zugeordnet.

**Item 11:** *Hat der Arzt erfasst welche Sorgen/Bedenken im Zusammenhang mit der Erkrankung und Behandlung für den Patienten am wichtigsten sind?*

Dieses Item wurde ersatzlos gestrichen, da das Kriterium nach Ansicht der Autorin von anderen Items hinreichend abgedeckt wird.

**Item 13:** *Hat der Arzt die Information strukturiert vermittelt und sich vergewissert, dass der Patient die Information verstanden hat?*

Dieses Item wurde weggelassen, da es in dem Arzt-Jugendlichen-Gespräch selten so viel Information zu überbringen gab, dass bewertet werden konnte, ob diese auch strukturiert vermittelt wurde. Stattdessen wurde ein erweitertes Item, dass die gesamte Gesprächsstruktur bewerten soll, (siehe 4.2.1.5 Hinzugefügte Items) hinzugefügt.

### 4.2.1.5 Hinzugefügte Items

Bei der Betrachtung erster Videos fiel auf, dass einige Aspekte, die essentiell für eine gute Gesprächsführung sind, durch die AGBS nicht hinreichend abgedeckt werden. Beispielsweise bekommt der emotionale Aspekt, der gerade in der psychotherapeutischen Medizin eine gewichtige Rolle spielt, nicht genug Beachtung.

Unter Einbeziehung des Calgary –Cambridge-Observation- Guide to the medical interview<sup>12</sup> (Kurtz, Silvermann 1996), der MAAS-Global Rating List for Consultation Skills of Doctors (Van Dalen, Prince et al. 1998; van Thiel J 2000) und von The Roter Method of Interaction Process Analysis (Roter and Larson 2002), vergleiche hierzu auch 2.3 „Überblick über Bewertungsinstrumente“, wurden für die vorliegende Gesprächssituati-

<sup>12</sup> Die genannten Bewertungsskalen werden in der Einleitung näher erläutert.

on wichtig erscheinende Bereiche umformuliert.

So wurden insgesamt 4 Items neu eingesetzt.

„Hat der Arzt auch nonverbale Äußerungen mit einbezogen?“ in die mAGBS-Jugendliche als neues Item (Item 9) aufgenommen.

Auch Strukturierung des Gespräches wurde in der mAGBS-Jugendliche als eigenes Item „Ist es dem Arzt gelungen, dem Gespräch eine Struktur zu geben?“ (Item 11) erfasst.

### **mAGBS – Jugendliche– Abschnitt D**

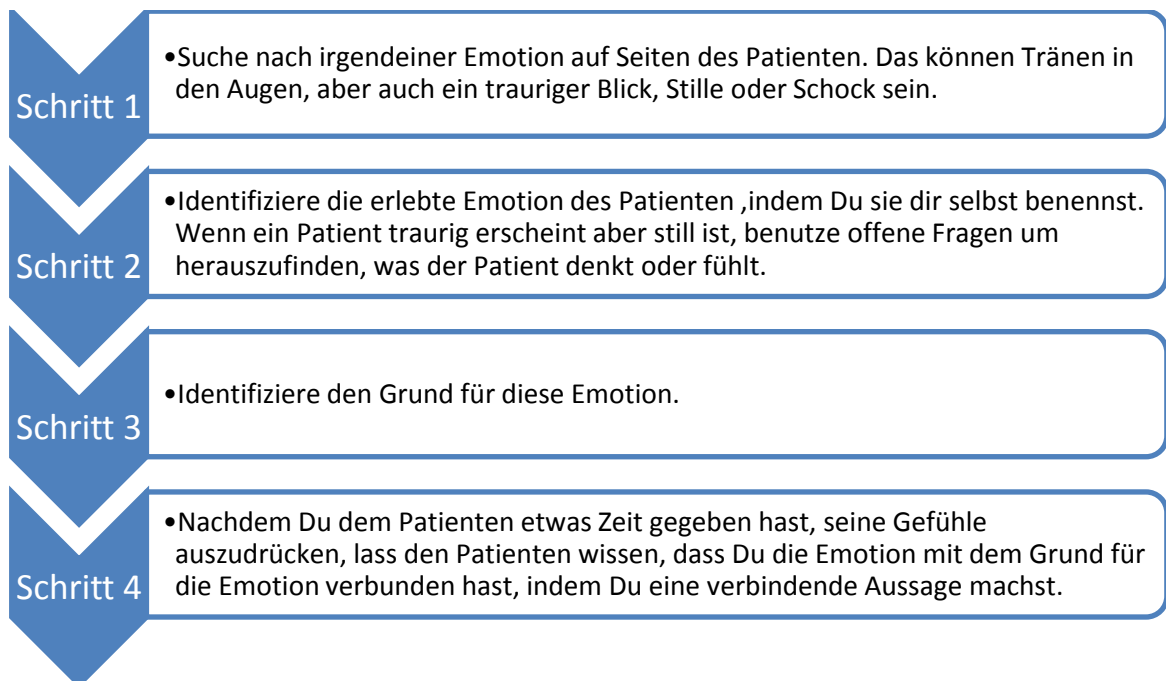
**Item 12:** *Benutzt der Arzt Wiederholungen und Zusammenfassungen?*

**Item 14:** *Hat der Arzt Empathie gezeigt?* Zur Definition siehe (Baile et al., 2000).

Um eine genaue Definition der Empathie unabhängig vom Verständnis des Raters zu haben, wird in der Fragestellung auf die von (Baile et al., 2000) und (Ptacek & TL, 1996) verwendete Beschreibung hingewiesen:

Eine empathische Äußerung besteht aus vier Schritten (Ptacek & Tara, 1996):

**Abbildung 4** Definition von Empathie



### **mAGBS – Jugendliche – Abschnitt F**

**Item 23:** *Globaler Eindruck des Raters – Hat der Arzt das Gespräch gut geführt?*

Dieses Item dient der klinisch-globalen Einschätzung des Gesprächsverhaltens des Arztes, es wird nicht in die mAGBS-Jugendliche Wertung einfließen. Ziel dieses Items ist, herauszufinden, ob die Gesamtwertung der mAGBS-Jugendliche auch dem Ein-

druck der Rater beim Beobachten des Gespraches entspricht und ob eventuell ein subjektiver Eindruck die Bewertung beeinflusst.

In hnlicher Form findet sich das Kriterium auch in den Calgary-Cambridge-Observation Guides.

**Item 24:** Globaler Eindruck des Raters – Wurde der Patient gut dargestellt?

Hier wird ein neues Item eingefuhrt, welches ohne festgelegte Kriterien bewertet werden soll. Ziel ist es zu beurteilen, ob die unterschiedliche Darstellung der Eltern im Rollenspiel zu einer schlechteren oder auch besseren Gesprachsleistung der rzte fuhrt. Vor allem in einer Rollenspielsituation in der sich die Darsteller vor dem Spiel bereits personlich kennen, wie es moglicherweise in einem Studentenkurs der Fall ist, ist anzunehmen, dass die Jugendlichen-Darsteller Einfluss auf das Gesprachsverhalten des Arzt-Darstellers, also eines Kommilitonen, nehmen.

#### 4.2.1.6 berblick ber die mAGBS-Jugendliche

Die folgende Abbildung zeigt noch einmal die Einzelschritte, die zur endgultigen Version der mAGBS-Jugendliche fuhrte.

Abbildung 5 Entwicklungsschritte zur mAGBS-Jugendliche



Damit besteht die mAGBS-Jugendliche aus sechs Teilbereichen mit 22 + 2 separat zu bewertenden Fragen, die auf einer Skala von 1-5 (1 = beste Bewertung, 5 = schlechteste Bewertung) bewertet werden mussen. Das Originalinstrument, die AGBS, besteht ebenfalls aus 22 Fragen, es fehlen allerdings die beiden globalen Items.

#### 4.2.1.7 Kurze Beschreibung des Manuals

Nach Vorbild der AGBS (Schildmann & Wand, 2008) ist das Instrument zusatzlich von einem Manual begleitet. Dieses Manual gibt jeweils eine kurze Einfuhrung in die Themenblocke, erklart einzelne Items genauer (siehe Item 14 „Empathie“), und legt die Notenvergabe fest. Es soll die einheitliche Bewertung der Rater unterstutzen und in Zukunft auch ungeschulten Ratern eine Moglichkeit geben, sich mit der Bewertungsskala (mAGBS-Jugendliche) rasch vertraut zu machen.

#### 4.2.2 Verlaufscharakteristik von Gesprachen

In einem Gesprach sind sowohl der Inhalt als auch formale Gesprachstechniken von Bedeutung. Die mAGBS-Jugendliche berucksichtigt zwar in einzelnen Items formale

Gesprächskriterien: wie den Einsatz von offenen Fragen und geschlossenen Fragen (Item 3 „Befinden“ und Item 8 „Genauer Nachfragen“), das Pausen setzen (Item 7 „Pausen“) Gesprächstempo (Item 21 „Gesprächstempo“) und Gesprächsstruktur (Item 11 „Struktur“). Jedoch zeigte sich bei einer Sichtung der Ergebnisse, dass die auf den ersten Blick formal erscheinenden – daher gut zu operationalisierenden Kriterien - beim Bewerten Schwierigkeiten bereiten. In der vorliegenden Arbeit wurde versucht, einen Zugang zu den Mustern der Kommunikation neben der Anwendungstestung der mAGBS–Jugendliche über den Einsatz von der Software „Interact“ zu bekommen. Mittels „Interact“ soll nachvollzogen werden, welche Äußerung worauf folgt und wie der Arzt ein Gespräch gestaltet. Hier zeigen sich Parallelen zu den Interaktions-Analyse-Systemen (siehe 2.3 Überblick über Bewertungsinstrumente).

### **4.2.2.1 Beschreibung der Software „Interact“ (Firma Mangold International)**

„Interact“ ist eine Software zur Erfassung und Auswertung von Daten in Beobachtungsstudien. So kann man ein Video über eine beliebige Situation sei es eine Gesprächssituation oder eine andere Interaktion mittels „Interact“ auswerten, indem man vorhandene Kodierungsmöglichkeiten des Programms nutzt oder selbst neue Kodierungen festlegt. Man überlegt sich, welche Interaktionen und Verhaltensweisen im Gespräch wesentlich für die jeweilige Fragestellung sind. Für jedes Ereignis wird ein Buchstabe festgelegt. Diese Kodierung wird gespeichert. Sobald das Ereignis auftritt, reicht das einfache Drücken der jeweiligen Taste. Auf diese Weise kann die Länge und Häufigkeit verschiedener Ereignisse gemessen werden.

### **4.2.2.2 Formale Gesprächskriterien**

Mithilfe von Interact wurde bei allen 32 Videos gemessen, wie viel Redeanteil der Arzt und wie viel Redeanteil der Jugendliche am Gesamtgespräch hat. Darüber hinaus wurden die Gespräche in 3 thematische Blöcke zerlegt: Begrüßung, Hauptteil, Abschluss. Zusätzlich wurde der Fragestil des Arztes untersucht, indem jede Frage entweder als offene Frage oder geschlossene Frage kodiert wurde.

Außerdem wurden sogenannte Hörsignale des Arztes gezählt. Unter Hörsignale versteht man Äußerungen wie „hm, ja, aha, ach ja, ach so...“ die den Gesprächsfluss in Gang halten und dem Gegenüber signalisieren, das aktiv zugehört wird.

Bei „Interact“ gibt es die Möglichkeit, sich die kodierten Ereignisse, - wenn der Arzt spricht bzw. wenn der Jugendliche spricht – in Form einer Zeitleiste grafisch darstellen zu lassen. Solch ein Interaktionsgraph wurde für jedes Video erstellt und es wurde über diese Methode versucht, auf Muster der Kommunikation zu stoßen. Zunächst wurde

wieder jede Videogruppe für sich untersucht. Anschließend sollen Gesprächsverläufe auf ihre jeweiligen Besonderheiten dargestellt und in Bezug zur mAGBS-Jugendlichen gesetzt werden. Die Software „Interact“ wurde verwendet, um Redeanteil, Fragestil, Hörsignale und Länge der nicht unterbrochenen Patientenäußerungen zu messen und auch daraus Muster der Gesprächsführung herzuleiten.

Ein besondere Schwerpunkt wird dabei auf das Thema „Pausen aushalten“ und „Gesprächstempo“ gelegt, darüber hinaus soll ein möglicher Bezug zwischen dem Fragestil des Arztes und den Gesprächsanteilen des Jugendlichen hergestellt werden.

### **4.3 Rater<sup>13</sup> und Raterschulung**

Im Folgenden wird das Wort Rater (aus dem Englischen to rate = bewerten) anstelle von Beobachter verwendet. Dies schien der Autorin passender, da auch von Rating-skala gesprochen wird und in der Originalarbeit ebenfalls der Begriff „Rater“ verwendet wird.

In Anlehnung an die Publikation von Wand (2008) wurde für diese Studie eine Anzahl von drei Ratern als ausreichend erachtet, um die Videos zu bewerten.

Die Rater brachten dabei einen unterschiedlichen Ausbildungshintergrund mit. Zwei von ihnen, die Autorin und ein weiterer Doktorand, befanden sich zum Zeitpunkt der Videobewertung am Ende des klinischen Abschnittes des Humanmedizinstudiums in München. Rater 3 (Frau Dipl. Soz.-Päd. Sabine Klemz) hat zusätzlich zu ihrem sozialpädagogischen Hintergrund eine psychotherapeutische Ausbildung und hatte sich auch mit anderen Videoprojekten zum Thema Kommunikation befasst. Sie hatte bereits im Rahmen einer vorhergehenden Doktorarbeit 40 Videos beurteilt und viel Übung im Ansehen und Beurteilen von Videos mit verschiedenen Situationen.

Das Ziel der Schulung bestand darin, einen gleichbleibenden Beurteilerstandard zu erreichen. Beide Studenten (Rater 1 und Rater 2) hatten den MeCum-Kurs Kinder- und Jugendpsychiatrie (siehe 2.5.2 Das Wahlfach Kinder- und Jugendpsychiatrie) besucht.

Auf Grund der Erfahrung von Rater 3 wurden die beiden anderen Rater größtenteils von ihr in der Bewertung von Videomaterial unterwiesen. Einen weiteren Teil der Unterweisung unternahm Prof. Dr. Reiner Frank.

Am Anfang wurde den beiden Studenten anhand verschiedener Videoaufnahmen ein Gespür für Videoaufnahmen in schwierigen Gesprächssituationen vermittelt, indem in

der Gruppe über gesehene Beispiele diskutiert und Meinungen ausgetauscht wurden. Diese Phase dauerte etwa 8 Stunden, verteilt auf 4 Tage.

Auf diese Weise wurden zu ersten Trainingszwecken insgesamt 16 Videos (Überbringen schlechter Nachrichten, Übermitteln eines psychischen Befundes im Arzt-Eltern Gespräch, Gespräch zwischen Arzt und suizidalem Jugendlichen) gesehen und diskutiert. Alle diese Videos hatten eine Länge von etwa fünf Minuten.

Drei Videos wurden mit der Aufklärungsgesprächsbewertungsskala<sup>14</sup> (Wand et al., 2008) bewertet, um mögliche Verständnisfragen zu klären und Unklarheiten zwischen den Ratern zu beseitigen.

Die beiden Studenten, die Autorin (Rater 1) und Herr Robert Braun (Rater 2), entwickelten parallel zu dieser Schulung ihre modifizierte Version der Aufklärungsgesprächsbewertungsskala. Robert Braun modifizierte die Bewertungsskala für Elterngespräche.

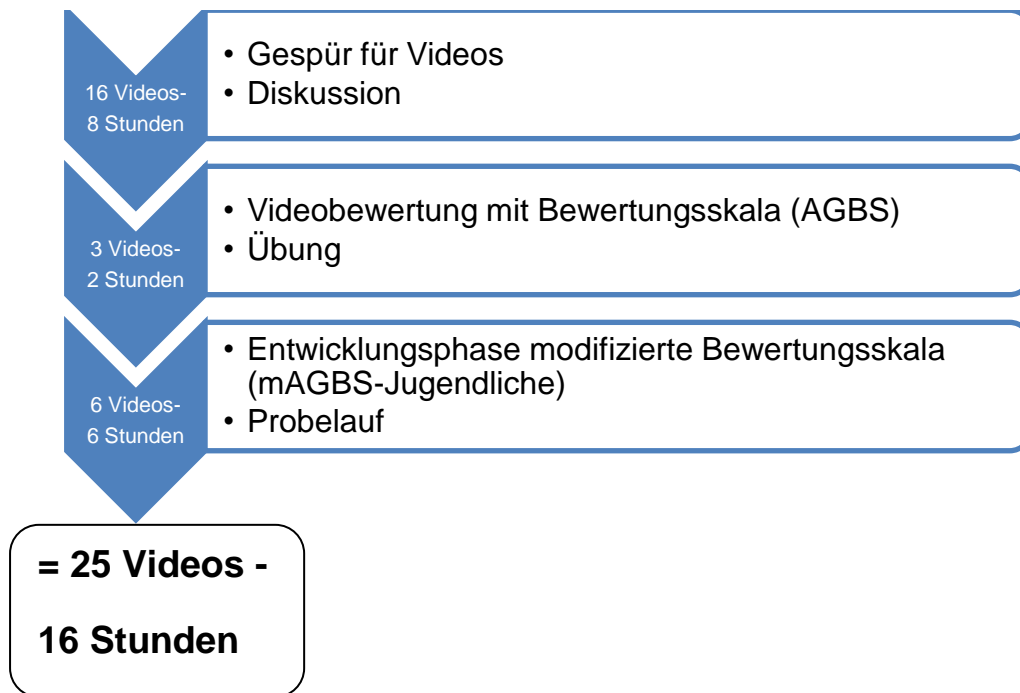
Die modifizierte Aufklärungsgesprächsbewertungsskala für Jugendliche wurde anschließend von allen drei Ratern gemeinsam an zwei Videos getestet und besprochen. Dabei wurden die einzelnen Items klar definiert. Die Dauer zur Bewertung der beiden Videos mithilfe der Bewertungsskala betrug etwa zwei Stunden.

Den Abschluss der Trainingsphase bildete das unabhängige Bewerten von vier Videos aus Arzt-Jugendlichen-Gesprächen. Diese Phase nahm je nach Beobachter zwischen drei und vier Stunden in Anspruch.

Abschließend wurde noch einmal einheitliches Vorgehen für die Bewertung festgelegt. Außerdem wurde abgesprochen, zwischen dem Ansehen der einzelnen Videos angemessene Pausen zu machen, um einer Ermüdung und damit nachlassender Motivation vorzubeugen. Der Raum, in dem die Videos gesehen wurden, bot eine reizarme Umgebung. Die Rater waren angewiesen, ihre Ratingzeiten, sowie eventuell unterbrochene Bewertungseinheiten zu dokumentieren.

<sup>14</sup> Die Aufklärungsgesprächsbewertungsskala, im folgenden abgekürzt als AGBS, entwickelt 2008 (Wand et al., 2008) ist Grundlage für die Entwicklung der modifizierten Skala, die in der vorliegenden Arbeit verwendet wurde.

Abbildung 6 Umfang der Raterschulung



## 4.4 Statistische Auswertung

Die Berechnungen wurden zunächst für alle Einzelitems und alle drei Rater durchgeführt. In SPSS kann unter dem Menüpunkt „Analysieren“ die Verknüpfung Reliabilitätsanalyse aufgerufen werden und es können alle für eine vollständige Reliabilitätsanalyse notwendigen Berechnungen durchgeführt werden.

Die Trennschärfen der Rater und die Korrelationen zwischen den Ratern<sup>15</sup> wurden berechnet, um zu überprüfen, ob das Urteil des jeweiligen Rater in einem Zusammenhang mit dem Urteil der übrigen Rater steht und wie homogen die Rater urteilen. Es wurde eine Varianzanalyse durchgeführt und Tukeys Additivitätstest durchgeführt. Zuletzt wurde der zweifaktorielle, unjustierte ICC berechnet.

### 4.4.1 Güte der Daten

Vor der statistischen Auswertung der mAGBS-Jugendliche wurde die Eingabe der Daten in SPSS stichprobenartig überprüft. Es wurde die Eingabe von 20 der insgesamt 96 Bögen kontrolliert<sup>16</sup>. Ziel dieser Überprüfung war, eine Fehlerrate unter 5 % zu erreichen. Es wurde bei einer Kontrolle der Dateneingabe auf 940 eingegebene Parameter

<sup>15</sup> Dadurch wird zusätzlich zur Trennschärfeanalyse durch die Betrachtung der paarweisen Korrelationen eine differenzierte Analyse von Uneinstimmigkeiten zwischen Ratern möglich.

<sup>16</sup> Pro Bogen wurden 47 Parameter erfasst, es wurden also mit 20 Bögen 940 Parameter überprüft.



4 Fehler gefunden, was eine Fehlerrate von < 0,5 % bedeutet.

#### 4.4.2 Mittelwerte und Standardabweichungen

Ausgehend von der Annahme, dass es sich bei der vorliegenden Ratingskala um eine Intervallskalierung handelt<sup>17</sup>, wurden zunächst die Mittelwerte und Standardabweichungen für die Einzelratings und die Gesamtwertung berechnet. Alle Berechnungen wurden mittels SPSS 16 für Windows durchgeführt.

#### 4.4.3 Reliabilität, Varianzanalysen

##### 4.4.3.1 Interne Konsistenz

Als Maß der Reliabilität<sup>18</sup> wurde zum einen die interne Konsistenz als Cronbachs Alpha berechnet. Die interne Konsistenz beruht auf einer durchschnittlichen Korrelation zwischen den einzelnen Items. Sie gibt an, welcher Anteil der Varianz der wahren Merkmalsausprägung eines Objekts aus dem Testergebnis fehlerfrei vorhersagbar ist. (Wirtz, Caspar 2002)

Cronbachs Alpha hat folgende Formel:

**Formel 1 Cronbachs Alpha**

$$\alpha = \frac{c}{c-1} \times \left[ 1 - \frac{\sum_j^c s_j^2}{s_x^2} \right]$$

Dabei ist:

$s_j^2$  die Varianz der Testitems

c die Anzahl der Testitems und

$s_x^2$  Varianz der Gesamtrohwerte

<sup>17</sup> Als Mindestkriterium für die Definition als Intervallskala ist zu fordern, dass die Gleichabständigkeit der Skalenpunkte nicht augenscheinlich verletzt sein darf. Wenn die Gleichabständigkeit der Skalenpunkte moderat verletzt, ist kann ebenfalls das Verfahren für intervallskalierte Daten verwendet werden. Diese Entscheidung wurde nach dem Vorbild von Wand getroffen. Der Autorin ist bewusst, dass diese Annahme nicht unumstritten ist.

<sup>18</sup> Es gibt verschiedene Methoden, um die Reliabilität nachzuweisen. Die Retest-Reliabilität bedeutet, dass der Test zu einem weiteren Zeitpunkt wiederholt wird, bei der Split-Half-Reliabilität wird der Test in 2 Hälften geteilt und korreliert. Auf diese beiden Methoden wurde jedoch in der vorliegenden Arbeit verzichtet.

#### 4.4.3.2 Interraterreliabilität

Analog zum Vorgehen von Wand<sup>19</sup> wurde zur Berechnung der Interraterreliabilität das Modell der zweifaktoriellen, unjustierten Intraklassenkorrelation gewählt. Das dazugehörige Maß, der Intraklassen-Korrelationskoeffizient (ICC) setzt intervallskalierte Daten voraus. Um den ICC zu berechnen, wird die Varianz zwischen verschiedenen Ratings in Bezug auf dasselbe Messobjekt mit der über alle Ratings und Messmethoden entstandenen Varianz verglichen. Damit der ICC groß, die Interraterreliabilität also hoch ist, muss der Unterschied zwischen den Meßobjekten relativ groß, der Unterschied zwischen den Ratern in Bezug auf das jeweilige Messobjekt aber gering sein. Weil in diesem Fall alle Rater jeweils alle Videos beurteilt hatten, wurde der zweifaktorielle ICC berechnet. Unjustiert bedeutet, dass Mittelwertunterschiede zwischen den Ratern als Fehlerquelle verrechnet werden. Sie wird dann benutzt, wenn Entscheidungen anhand von absoluten Werten auf der Ratingskala getroffen werden, wie dies in der vorliegenden Skala der Fall war. Das unjustierte Modell entspricht der strengeren Prüfung.

Dies ist die Formel zur Berechnung des ICC:

**Formel 2 Intraclasscorrelationskoeffizient**

$$ICC = \frac{MS_{zw} - MS_{err}}{MS_{zw} + \frac{MS_{zwrat} - MS_{err}}{n}}$$

Wobei  $MS_{zw}$  die Varianz zwischen den Fällen,  $MS_{zwrat}$  die Varianz zwischen den Ratern und  $MS_{err}$  die Restvarianz ist,  $n$  ist die Anzahl der Messobjekte.

#### 4.4.3.3 Weitere Varianzanalysen

##### 4.4.3.3.1 Tukeys Additivitätstest

Unterschiede in den Ratings werden zum einen von der Varianz des Merkmals, dessen Ausprägung anhand der Skala gemessen wird, von der Varianz zwischen den Ratern und von der Varianz eines unsystematischen Fehlers erzeugt. Im vorliegenden zweifaktoriellen Modell, wenn also alle Videos von denselben Ratern beurteilt worden sind, lässt sich diese Varianz, die nicht auf Unterschieden zwischen den Ratern beruht, weiter aufspalten.

So fließen als Varianzanteile neben den Unterschieden zwischen den beurteilten Videos, Unterschiede zwischen den Ratern<sup>20</sup> und Interaktionseffekte<sup>21</sup> zwischen Ratern

<sup>19</sup> Die Auswertung erfolgte analog zu der Vorgehensweise im Originalinstrument, um eine spätere Vergleichbarkeit zu gewährleisten.

<sup>20</sup> Dies wird als Inhomogenität zwischen den Ratermittelwerten bezeichnet.

und Videos ein. Dieser Interaktionseffekt lässt sich mittels Tukeys Additivitätstest ermitteln und differenzieren.

### 4.4.3.3.2 Mehrfaktorielle univariate Varianzanalyse

Die mehrfaktorielle univariate Varianzanalyse wird eingesetzt, um den Einfluss verschiedener Faktoren auf eine Variable abschätzen zu können (Bortz, 2004).

In der vorliegenden Arbeit wurde der Einfluss von Geschlecht, Training und Ausbildungshintergrund auf das Gesprächsverhalten mittels einer Varianzanalyse näher untersucht.

### 4.4.3.4 Pearsonscher und Spearmans Korrelationskoeffizient

Die Korrelation ist ein Maß für den Zusammenhang zwischen zwei Variablen. Das Maß der Korrelation wird als Korrelationskoeffizient angegeben.

Er kann Werte zwischen  $-1$  und  $+1$  annehmen. Bei einem Wert von  $+1$  (bzw.  $-1$ ) besteht ein vollständig positiver (bzw. negativer) linearer Zusammenhang zwischen den betrachteten Merkmalen. Dabei wird der Pearsonsche Koeffizient verwendet, um den Zusammenhang zwischen mindestens intervallskalierten Merkmalen anzugeben, der Spearmans Koeffizient, um ordinalskalierte Merkmale anzugeben (Bortz, 2004).

### 4.4.3.5 Effektstärke

Um zu erkennen, wie bedeutsam ein Mittelwertunterschied zwischen den verschiedenen Videogruppen, nämlich den Jugendlichenvideos, den Notfallvideos und den Allgemeinärztevideos ist, wurde die Effektstärke berechnet.

Die Effektstärke stellt eine an der Merkmalsstreuung relativierte Differenz zwischen den Mittelwerten zweier Gruppen hinsichtlich eines Merkmals dar.

Sie wird berechnet aus dem Quotient der Differenz der Gruppenmittelwerte und den gemittelten Standardabweichungen (Bortz, 2004).

---

<sup>21</sup> Eine Interaktion würde bestehen, wenn die Rater je nach beurteilten Personen unterschiedlichen Einschätzungstendenzen unterliegen würden.

### Formel 3 Effektstärke

$$\frac{MW1 - MW2}{(STD1 + STD2) / 2}$$

#### 4.4.3.6 Validität

Über verschiedene Darstellungen wurde versucht, eine Annäherung an die Validität zu erreichen. Die Inhaltsvalidität wurde in einer ausführlichen Literaturrecherche, im Vergleich mit dem Originalinstrument sowie anderen modifizierten Instrumenten erfasst.

## **5 Ergebnisse**

Zunächst wird die Videoqualität, die Ratingdauer, die Geschlechterverteilung und die Rollenidentifikation der Darsteller beschrieben werden. Dann soll die Notenverteilung über die einzelnen Items dargestellt werden und davon ausgehend die Analyse zur Interraterreliabilität.

### **5.1 Deskriptive Darstellung**

#### **5.1.1 Qualität der Videoaufnahmen**

Insgesamt 32 Videos wurden von allen 3 Ratern gesehen.

Die Videos dauerten durchschnittlich 5 Minuten und 30 Sekunden, wobei das kürzeste 2 Min 30 Sekunden dauerte und das längste 6 Minuten 50 Sekunden.

Die Tonqualität wurde bei 2 Videos als mittelmäßig, bei einem als schlecht bewertet. Dabei war das Kriterium für eine mittelmäßige Bewertung, wenn einzelne Worte unverständlich waren, das Kriterium für eine schlechte Bewertung wenn ganze Sätze nicht zu verstehen waren.

Die Bildqualität wurde bei 2 Videos als mittelmäßig angegeben, was bedeutet, dass zwar alle Beteiligten im Bild waren, aber die Mimik nicht zu erkennen war.

8 Videos wurden als nicht vollständig aufgenommen bewertet, wobei bei 4 Videos der Anfang abgeschnitten war, bei 3 Videos das Ende, mehr als 15 Sekunden fehlten allerdings nur bei einem Video.

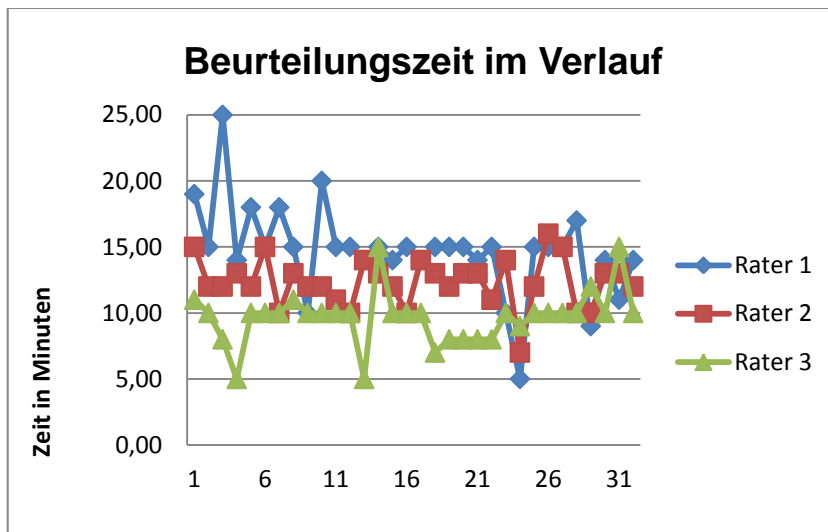
Es gab keine sprachlichen Probleme und keine Störfaktoren auf den ausgewählten Videos.

Rater 1 unterbrach insgesamt 4-mal während eines Ratings, Rater 2 und Rater 3 je einmal.

#### **5.1.2 Ratingzeiten**

Die 3 Rater benötigten im Durchschnitt 12 Minuten 20 Sekunden zur Bewertung eines Videos. Rater 1 benötigte durchschnittlich 14 Minuten, Rater 2 12 Minuten 30 Sekunden und Rater 3 10 Minuten. Die Rater wurden im Verlauf des Ansehens der 32 Videos schneller.

Abbildung 7 Beurteilungszeit im Verlauf



Zusammengefasst kann man sagen, dass für Videos von durchschnittlich 5 Minuten und 30 Sekunden Dauer, 12 Minuten und 20 Sekunden für die Beurteilung mit der mAGBS-Jugendliche nötig waren.

**Es muss also in etwa die zweieinhalbfache Zeit der Gesprächsdauer für die Anwendung der mAGBS-Jugendliche verwendet werden.**

### 5.1.3 Beurteilungsskala – Einzelitems der mAGBS

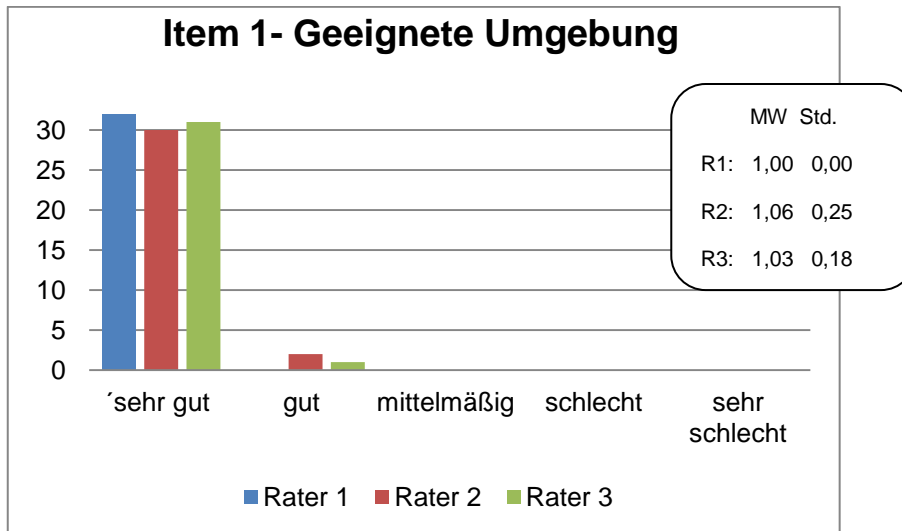
An dieser Stelle wird die Häufigkeit der Bewertungen der Rater für die einzelnen Items dargestellt. Zur besseren Übersicht wird kurz beschrieben, aus welchen Items die Themenblöcke bestehen und jedes Item mit einem Stichwort charakterisiert.

Bei den Histogrammen zeigte sich im Wesentlichen eine breite Ausschöpfung des Notenspektrums. Sehr schlechte Bewertungen finden sich jedoch insgesamt nur vereinzelt. Es gibt jedoch Items, bei denen die sehr guten bis guten Bewertungen dominieren.

### 5.1.3.1 Themenblock A – Einleitung des Gesprächs

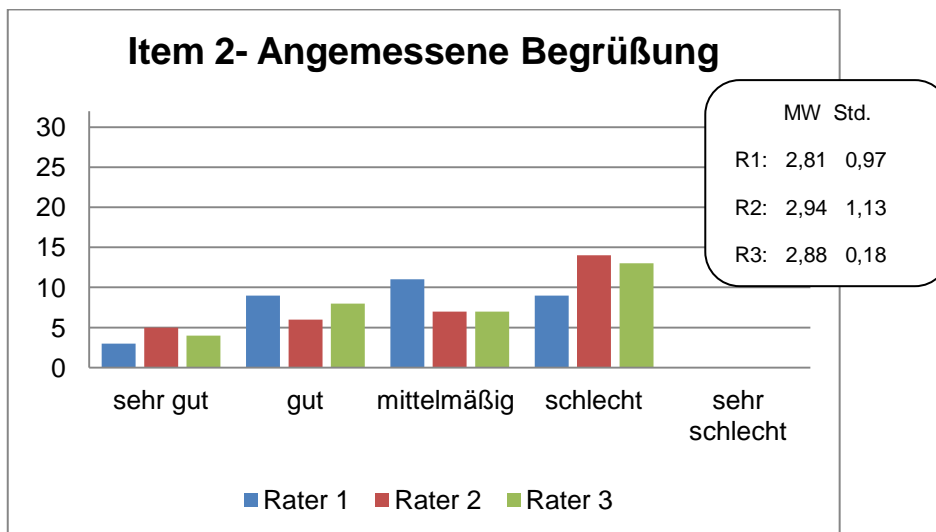
Dieser Block besteht aus 3 Einzelitems: Item 1 „Geeignete Umgebung“, Item 2 „Angemessene Begrüßung“, Item 3 „Befinden“.

Abbildung 8 Item 1



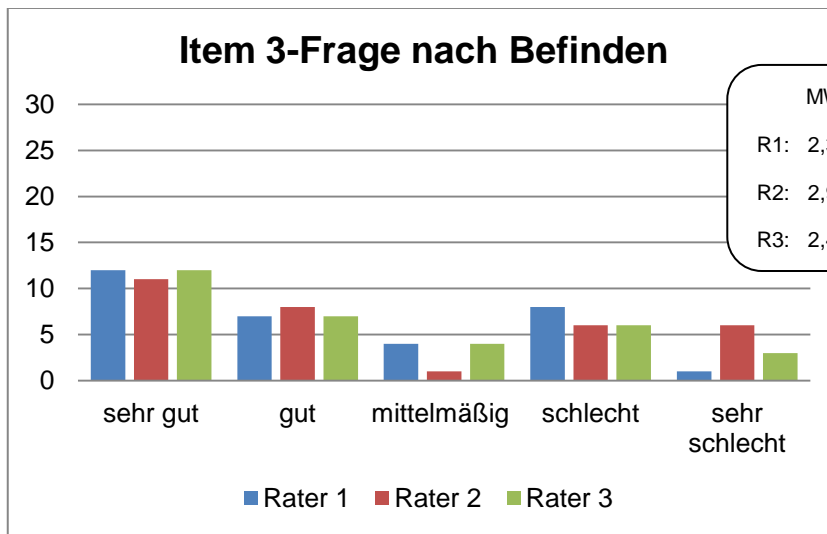
Bei Item 1 vergaben alle 3 Rater fast nur die Note 1 = sehr gut.

Abbildung 9 Item 2



Bei Item 2 finden sich nahezu alle Noten bis auf sehr schlechte Noten. Die Rater beurteilen relativ homogen.

Abbildung 10 Item 3



Bei Item 3 sind alle Noten vertreten. Rater 2 bewertet im Mittel kritischer als die anderen beiden Rater.

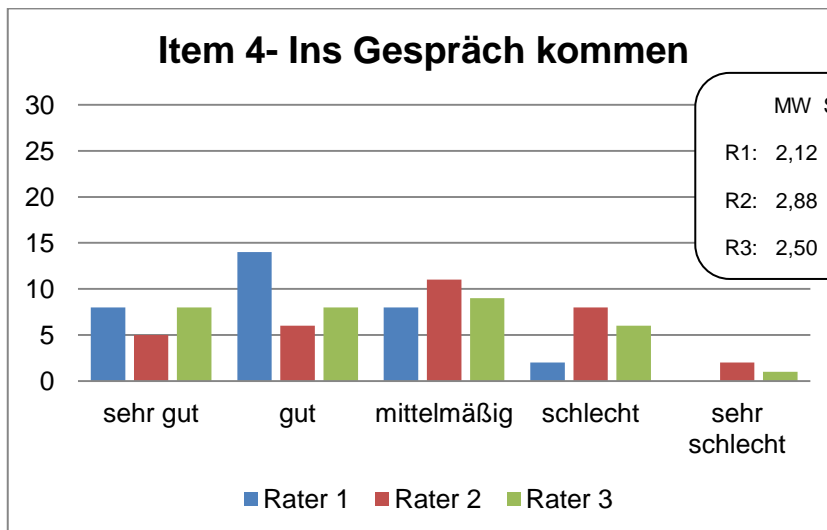
### 5.1.3.2 Themenblock B – Information bekommen

Dieser Block besteht aus 4 Items:

Item 4 „Ins Gespräch kommen“, Item 5 „Langsame Annäherung“

Item 6 „Verständliche Sprache“, Item 7 „Angemessene Pausen“

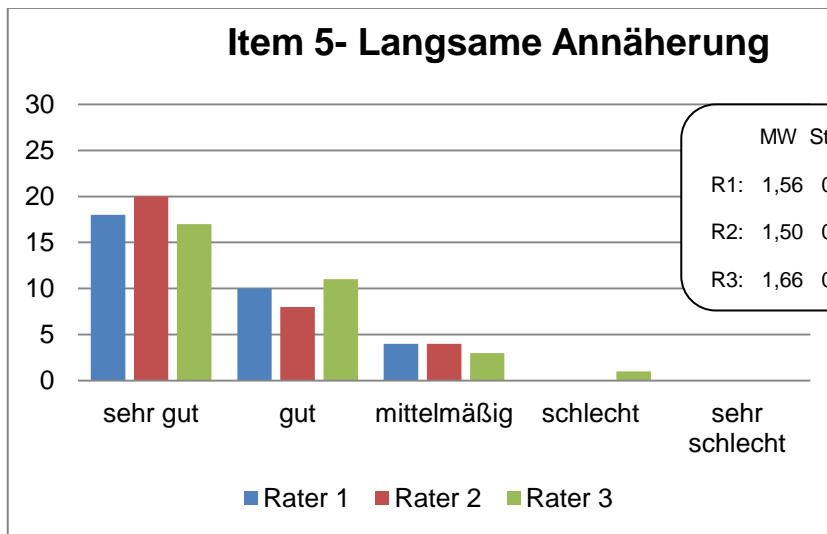
Abbildung 11 Item 4



Rater 1 bewertet viel besser als die anderen beiden Rater.

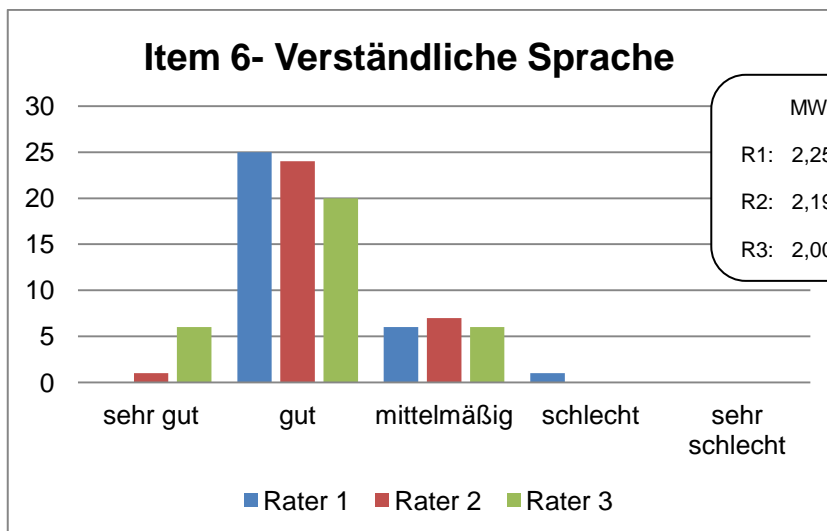


Abbildung 12 Item 5



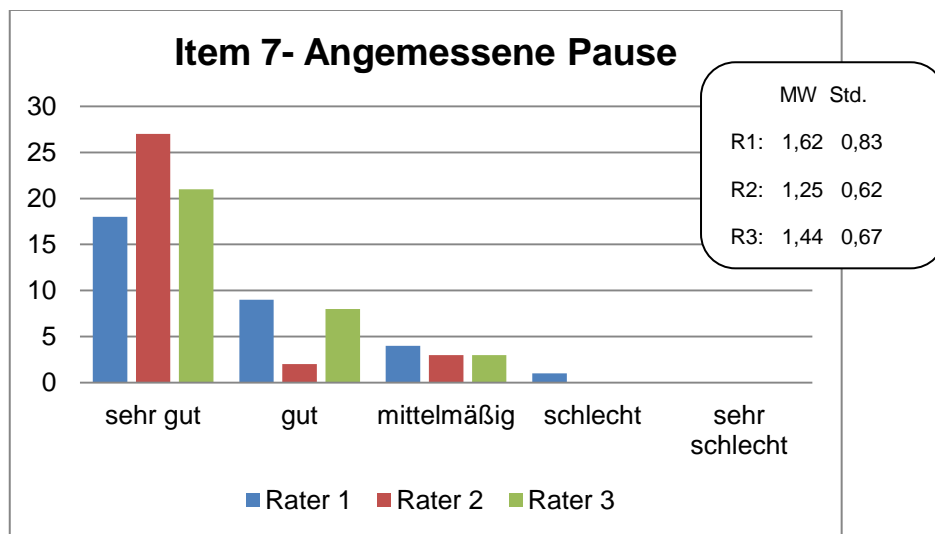
Bei Item 5 wird nicht das gesamte Notenspektrum ausgeschöpft.

Abbildung 13 Item 6



Bei allen 3 Ratern gibt es eine Notenhäufung im Bereich guter Noten.

Abbildung 14 Item 7

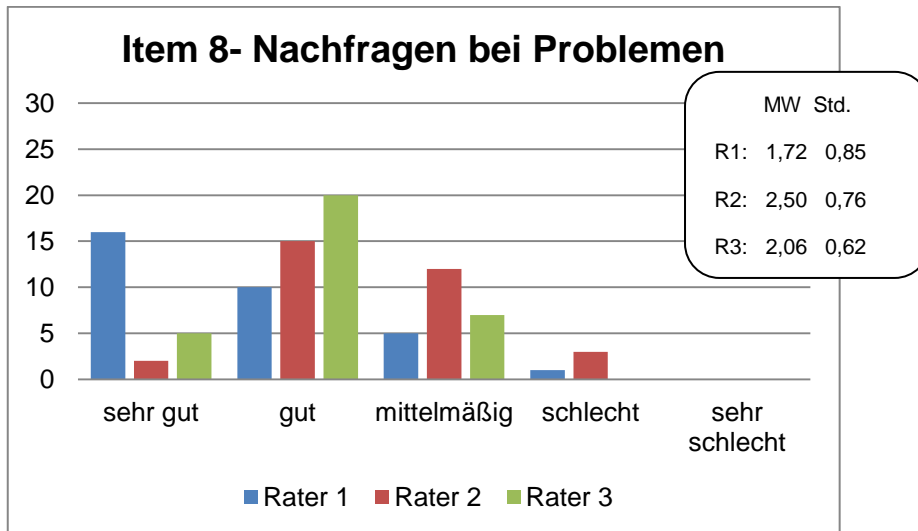


Auch bei diesem Item häufen sich die Noten im sehr guten Bereich.

### 5.1.3.3 Themenblock C - Bedenken des Jugendlichen herausfinden

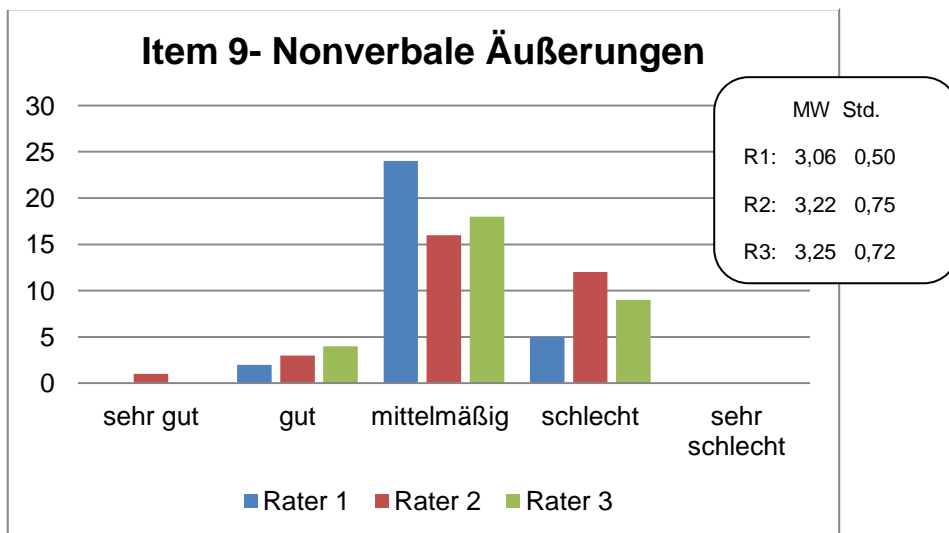
Dieser Block setzt sich aus den beiden Einzelitems Item 8 „Nachfragen bei Problemen“ und Item 9 „Nonverbale Äußerungen“ zusammen.

Abbildung 15 Item 8



Bei Item 8 bewertet Rater 1 um einiges besser als die anderen beiden Rater.

Abbildung 16 Item 9

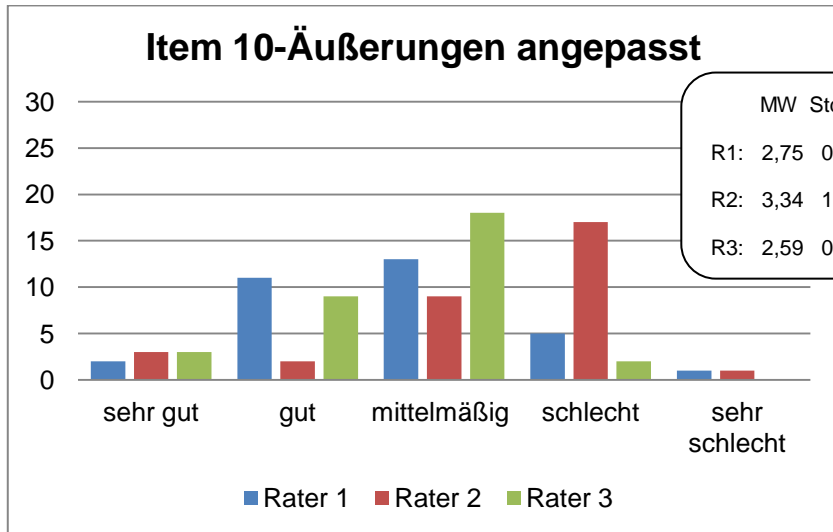


Bei diesem Item ist die starke Tendenz aller drei Rater zu mittelmäßigen Bewertungen auffällig.

### 5.1.3.4 Themenblock D- Perspektive vermitteln

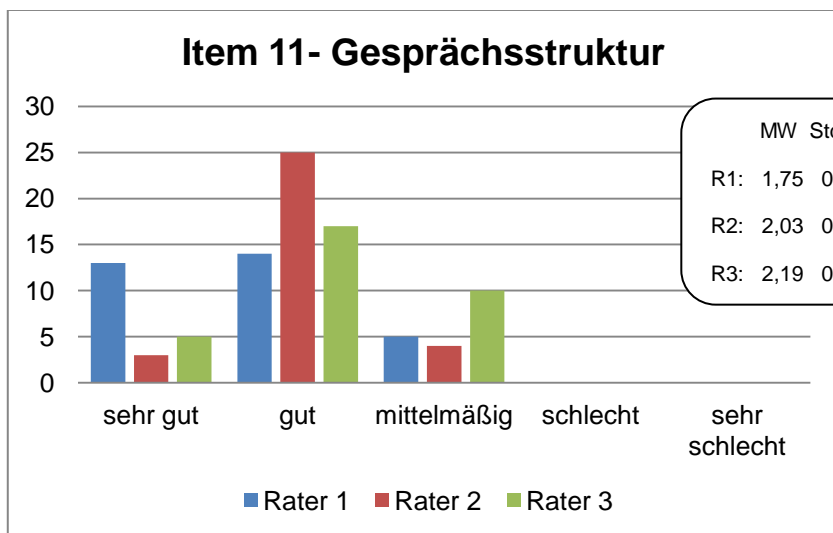
Dieser Block besteht aus 6 Einzelitems: Item 10 „Äußerungen angepasst“, Item 11 „Gesprächsstruktur“ Item 12 „Wiederholungen und Zusammenfassungen“ Item 13 „Positive Aspekte“ und Item 14 „Empathie“ und Item 15 „Fachlich korrekt“

Abbildung 17 Item 10



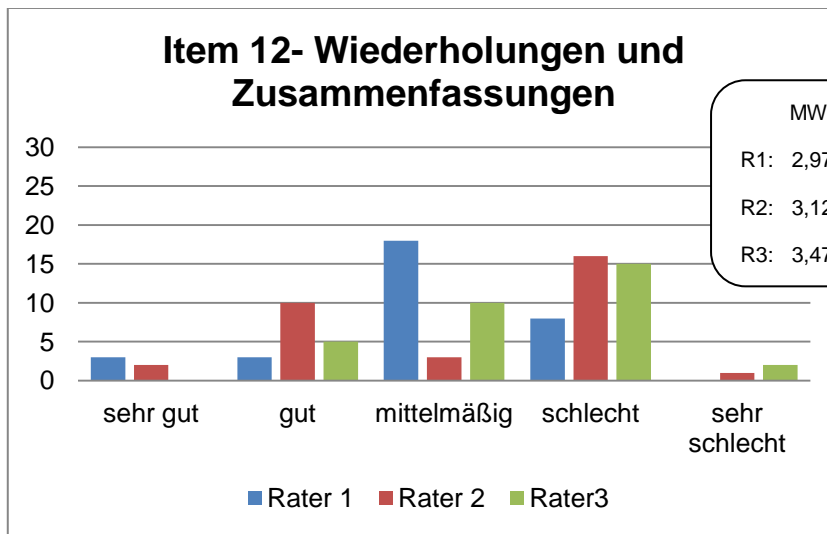
Bei diesem Item sind alle Noten gleichmäßig vertreten.

Abbildung 18 Item 11



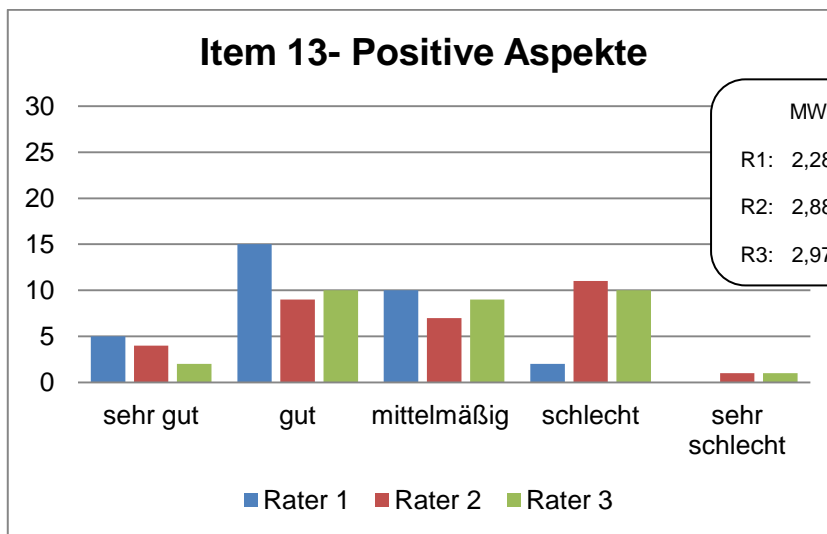
Bei diesem Item finden sich keine schlechten Noten.

Abbildung 19 Item 12



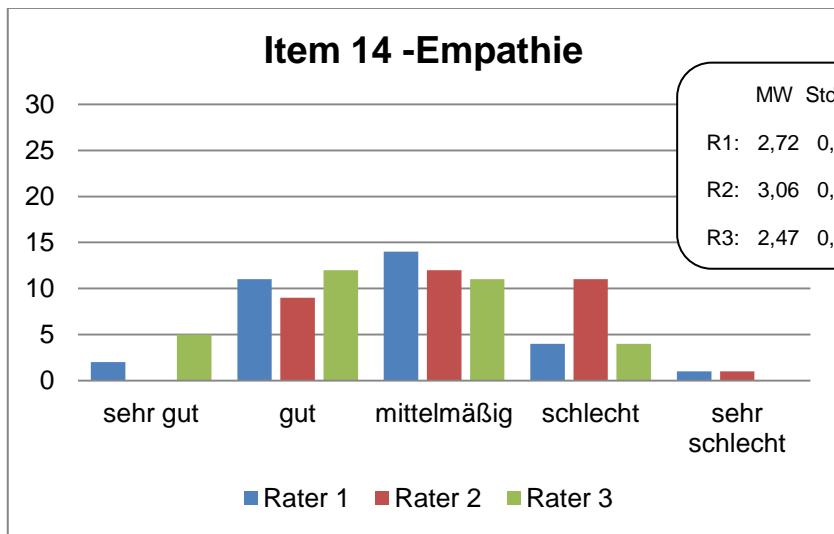
Hier verteilen sich die Noten sehr gleichmäßig.

Abbildung 20 Item 13



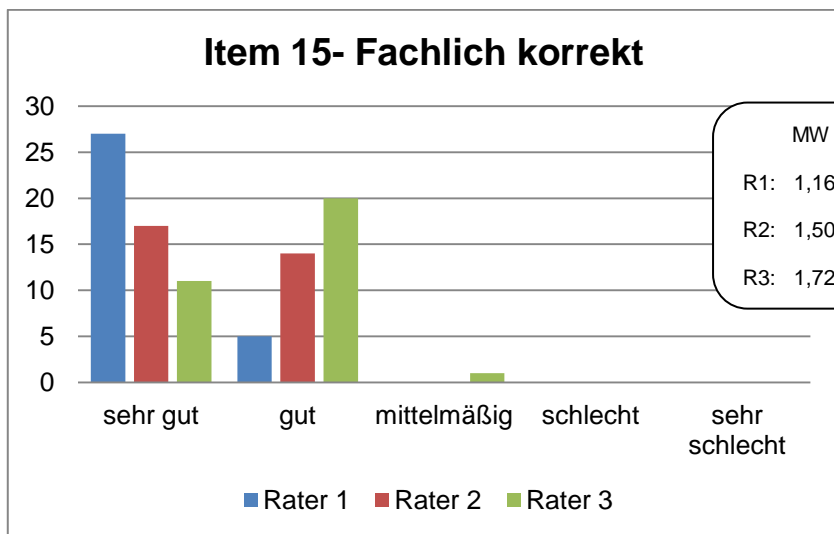
Hier findet sich ebenfalls eine gleichmäßige Verteilung.

Abbildung 21 Item 14



Hier findet sich eine starke Häufung mittelmäßiger Noten.

Abbildung 22 Item 15



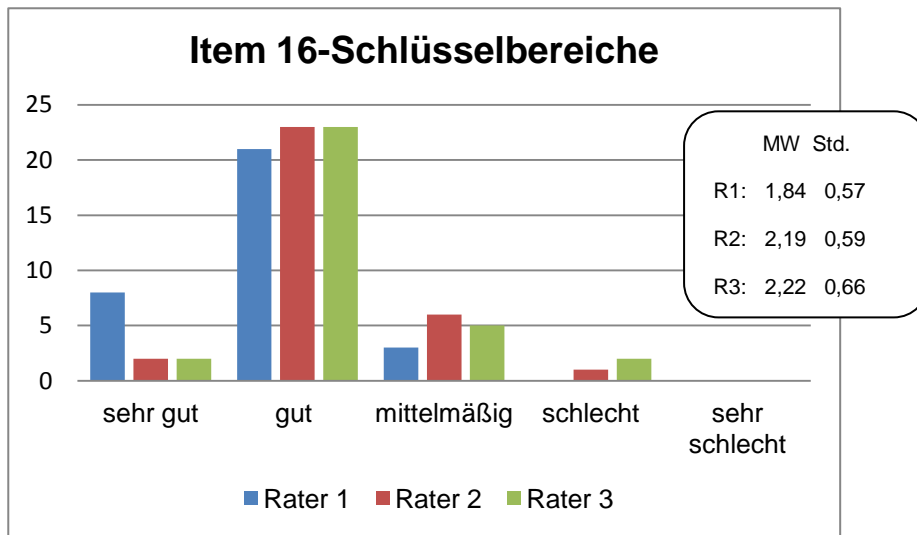
Hier kommen ausschließlich sehr gute und gute Bewertungen vor.

### 5.1.3.5 Themenblock E- Allgemeine Überlegungen

Der letzte Themenblock besteht aus 7 Einzelitems:

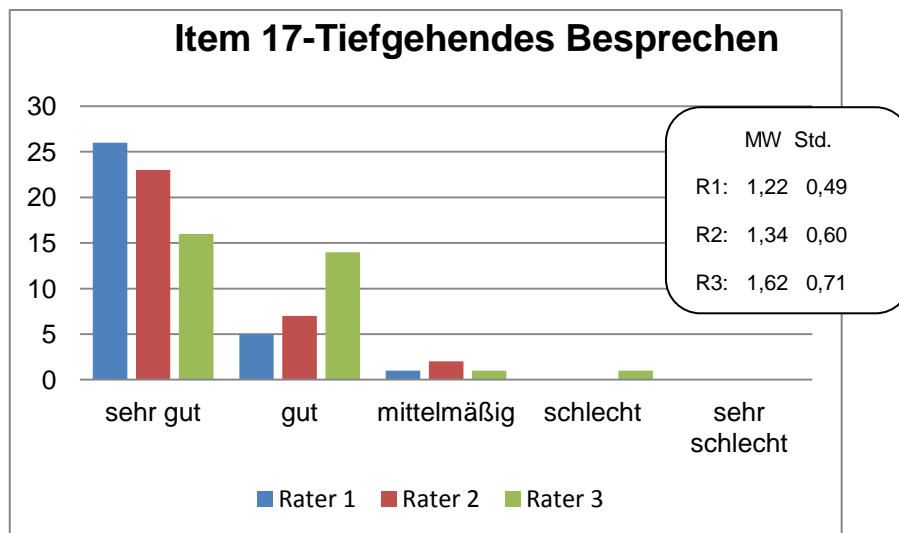
Item 16 „Schlüsselbereiche“, Item 17 „Tiefgehendes Besprechen“, Item 18 „Hilfreich zur Seite stehen“, Item 19 „Körpersprache“, Item 20 „Unbeholfenheit“, Item 21 „Gesprächstempo“, Item 22 „Beenden“

Abbildung 23 Item 16



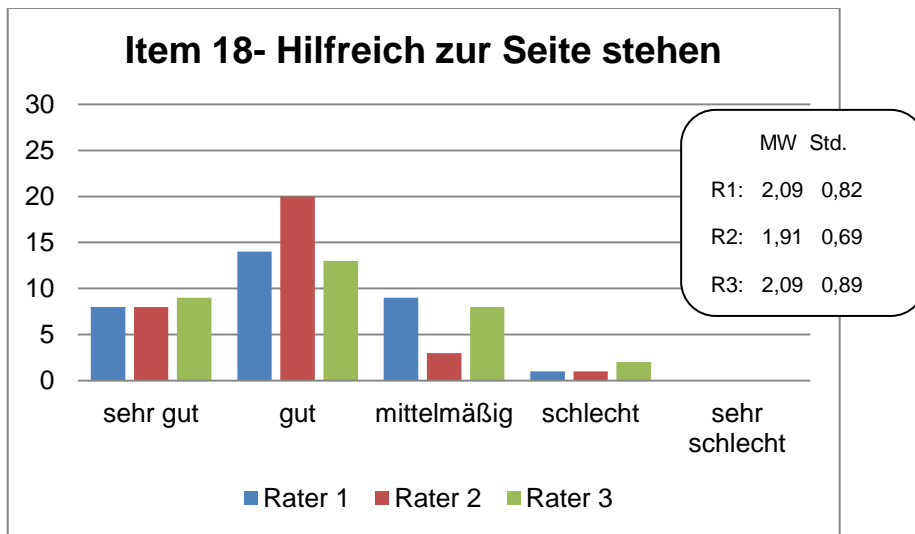
Bei diesem Item findet sich eine hohe Zahl guter Noten und wenige schlechte.

Abbildung 24 Item 17



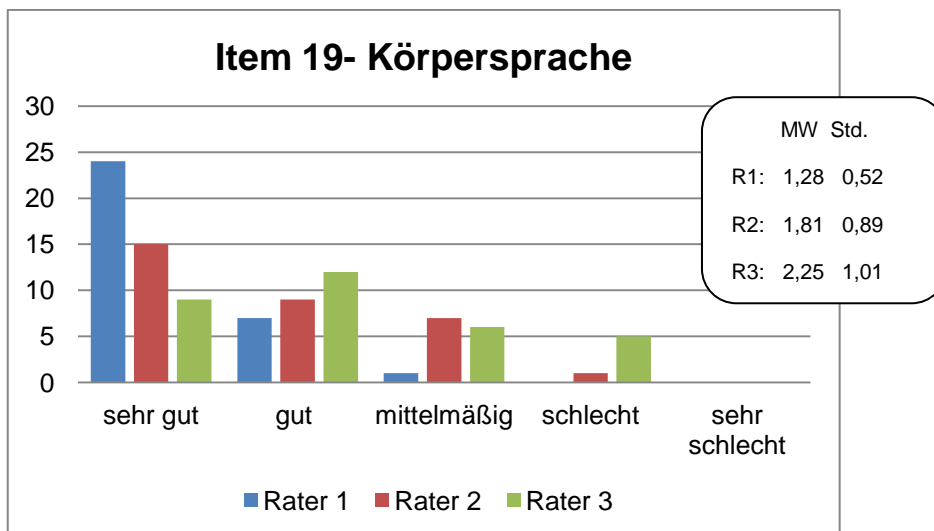
Auch hier finden sich fast ausschließlich sehr gute bis gute Bewertungen

Abbildung 25 Item 18



Bei diesem Item urteilen die Rater sehr homogen.

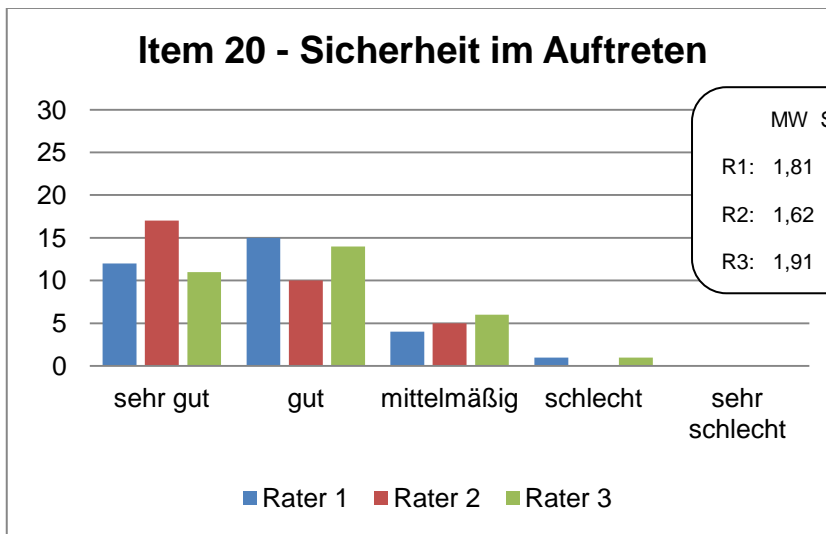
Abbildung 26 Item 19



Bei diesem Item finden sich wieder vor allem sehr gute und gute Bewertungen.

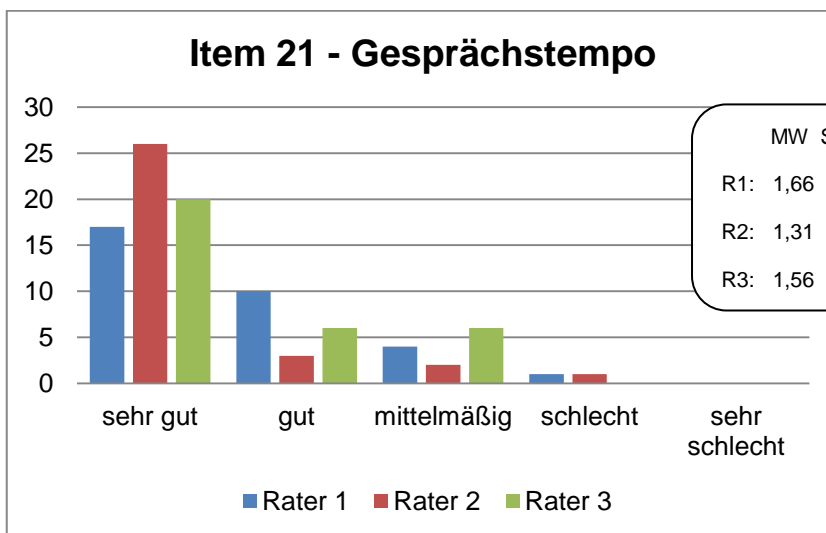


Abbildung 27 Item 20



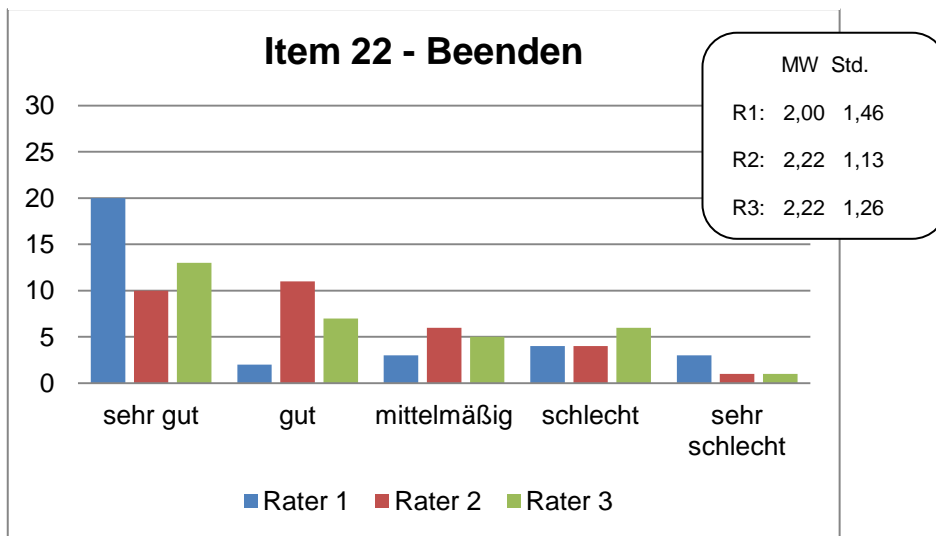
Auch hier finden sich kaum schlechte Bewertungen.

Abbildung 28 Item 21



Auch bei diesem Item fällt die Häufung sehr guter Noten auf.

Abbildung 29 Item 22

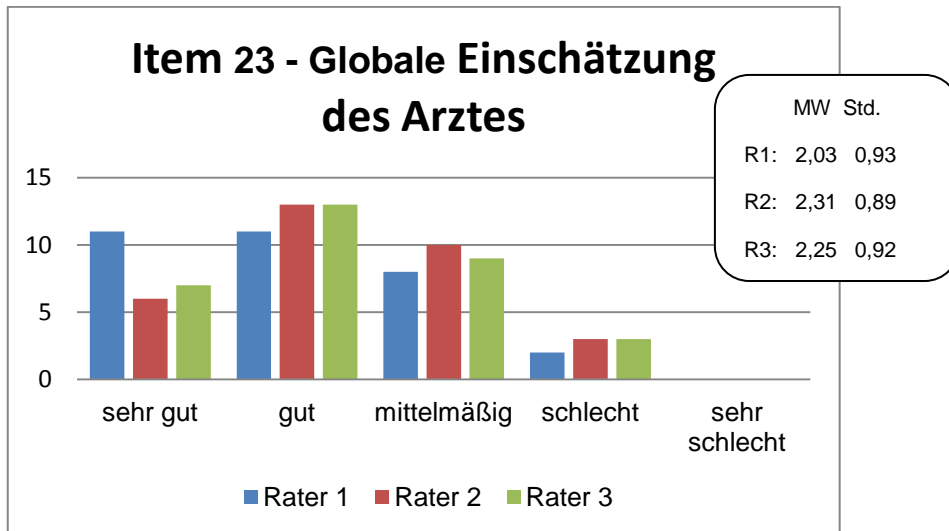


Hier kommen  
alle Noten  
relativ  
gleichmäßig  
vor.

### 5.1.3.6 Themenblock F – Globale Einschätzung der Rater

Item 23 – „Globale Einschätzung des Arztes“ und Item 24 – Darstellung der Patientenrolle“

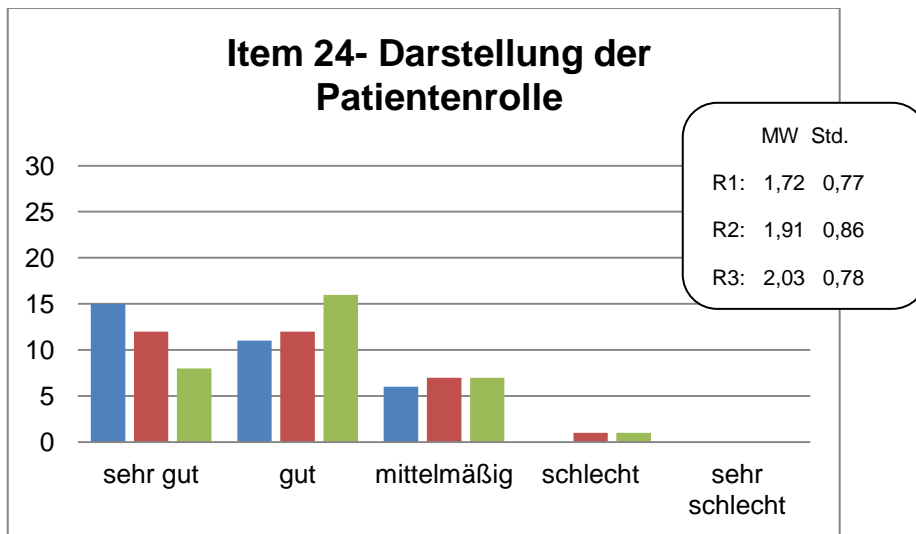
Abbildung 30 Item 23



### 5.1.4 Rollenidentifikation als Patient (Item 24)

Die Rollenidentifikation der Studenten auf den Videos wurde von den Ratern unterschiedlich wahrgenommen. Es gab kein Video, in dem die Studenten für komplett aus der Rolle gefallen angesehen wurden. Rater 1 gab für 5 Videos eine annähernd komplette Rollenidentifikation an, Rater 2 für 9 Videos, Rater 3 bei 3 Videos. Bei den restlichen Videos gab es eine komplette Rollenidentifikation.

Das folgende Histogramm gibt einen Überblick wie die einzelnen Rater die Leistung der Darsteller bewerteten.



Item 23 „Globale Bewertung des Arztes“ und Item 24 „Darstellung der Patientenrolle“ wurden miteinander korreliert, um herauszufinden, ob die Bewertung der Darstellung des Jugendlichen mit der Bewertung der Gesprächsleistung des Arztes zusammenhängt.

Tabelle 7 Korrelation Item 23 und Item 24

Rater 1	Rater 2	Rater 3
0,28	0,46	0,39

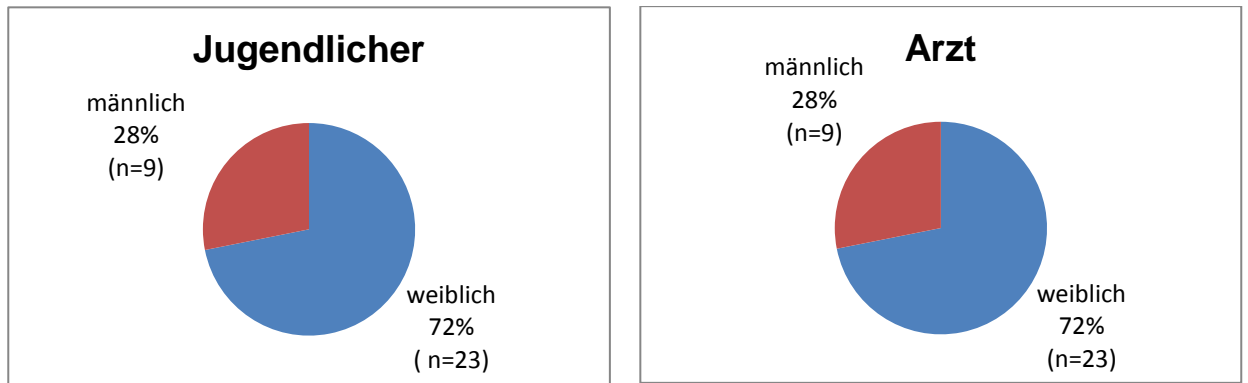
Der Pearsonscher Korrelationskoeffizient zeigt eine geringe bis mittlere Korrelation zwischen diesen beiden Items.

### 5.1.5 Geschlechtsverteilung der Darsteller

Von insgesamt 166 Teilnehmern im Wahlfach Kinder- und Jugendpsychiatrie und dem Kompaktkurs „Pädiatrie“ waren 134 weiblich, 32 männlich.

Die Rollenverteilung gibt dieses Verhältnis wieder.

Die Darstellung zeigt die Geschlechtsverteilung der Darsteller auf den 32 ausgewählten Videos. Die Prozentzahlen der Darsteller spiegeln in etwa die Geschlechtsverteilung der Teilnehmer insgesamt wieder.



Die Kombination Arztrolle weiblich, Jugendlicher weiblich fand sich 17-mal. In 5 Fällen waren beide Darsteller männlich. Die Kombination männliche Arztrolle, Jugendlicher weiblich fand sich 4-mal und die Kombination Arztrolle weiblich besetzt, Jugendlichenrolle männlich besetzt fand sich 6-mal.

Eine Betrachtung der einzelnen Items getrennt nach Geschlecht zeigt in der Benotung für die meisten Items keinen Unterschied.

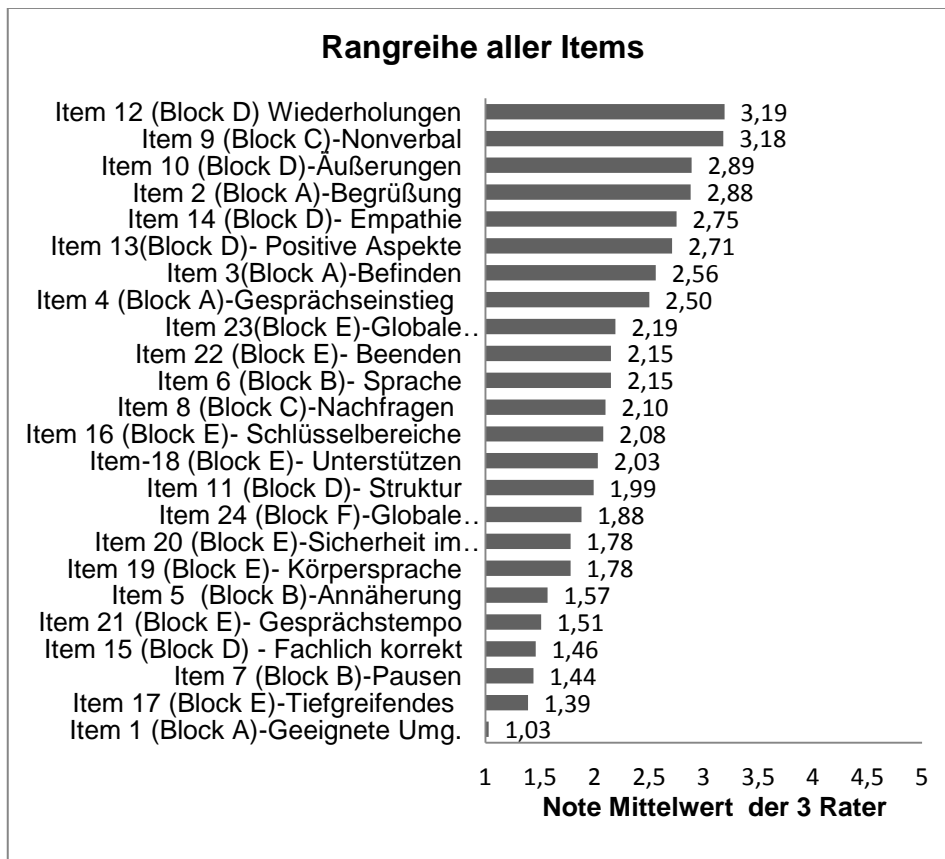
Bei Items, die psychosoziale Aspekte bewerten, Item 8 „Nachfragen bei Problemen“, Item 9 „Nonverbale Äußerungen“, Item 14 „Empathie“ sowie Items, die die Körpersprache und das Auftreten bewerten, Item 18 „Hilfreich zur Seite stehen“ und Item 19 „Körpersprache“, werden männliche Gesprächsführende tendenziell von allen drei Ratern etwas schlechter bewertet.

### 5.1.6 Zusammenfassung: Rangreihe

In der folgenden Tabelle wird die durchschnittliche Bewertung der drei Rater für jedes Item in Form einer Rangreihe (von den niedrigsten bis zu den höchsten Mittelwerten) dargestellt. 1 ist dabei eine sehr gute, 2 eine gute, 3 eine mittelmäßige Bewertung.

Dabei fällt auf, dass vor allem die Items aus Themenblock D – Perspektive vermittelt am schlechtesten bewertet werden. Items aus Themenblock E – Allgemeine Überlegungen bekommen im Schnitt bessere Bewertungen.

Abbildung 31 Rangreihe über einzelne Items



### 5.1.7 Themenblöcke des Gespräches

Der Mittelwert für die Gesamtbewertung durch die einzelnen Rater liegt zwischen 2,12 und 2,35. Die geringste Differenz zwischen den mittleren Bewertungen der Rater findet sich in Themenblock B „Information bekommen“ mit Werten zwischen 1,89 und 1,95. Die größte Differenz findet sich in Themenblock C „Bedenken herausfinden“ mit Werten zwischen 2,40 und 2,86. Die folgende Tabelle gibt dazu einen Überblick.

## Ergebnisse

Tabelle 8 Mittelwert Themenblöcke über alle Videoaufnahmen

	Mittelwert	Standardabweichung
<b>Themenblock A</b>	<b>Gesprächsbeginn</b>	
Rater 1	2,05	0,52
Rater 2	2,21	0,64
Rater 3	2,12	0,63
Durchschnitt	2,13	0,60
<b>Themenblock B</b>	<b>Information bekommen</b>	
Rater 1	1,90	0,54
Rater 2	1,95	0,42
Rater 3	1,90	0,58
Durchschnitt	1,92	0,51
<b>Themenblock C</b>	<b>Bedenken herausfinden</b>	
Rater 1	2,40	0,59
Rater 2	2,86	0,57
Rater 3	2,64	0,54
Durchschnitt	2,63	0,56
<b>Themenblock D</b>	<b>Perspektive vermitteln</b>	
Rater 1	2,27	0,51
Rater 2	2,66	0,49
Rater 3	2,54	0,59
Durchschnitt	2,49	0,53
<b>Themenblock E</b>	<b>Allgemeine Überlegungen</b>	
Rater 1	1,98	0,51
Rater 2	2,07	0,49
Rater 3	2,30	0,71
Durchschnitt	2,12	0,57
<b>Gesamtwertung</b>		
Rater 1	2,12	0,53
Rater 2	2,35	0,52
Rater 3	2,30	0,61
Durchschnitt	2,26	0,55

## 5.2 Reliabilitätsanalyse

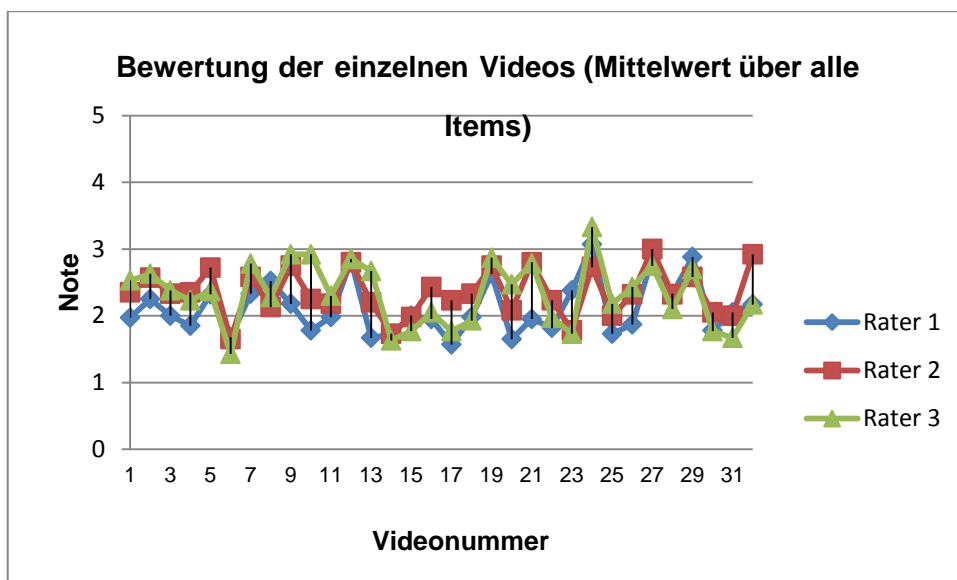
Es wurde mit den Daten aller Rater über alle Items eine Reliabilitätsanalyse durchgeführt. Da in der Originalarbeit von Wand anstelle der Einzelitems mit den Themenblöcken gearbeitet wurde, werden im Folgenden alle Ergebnisse sowohl für die Einzelitems als auch für die Themenblöcke dargestellt.

Die Reliabilitätstestung wurde an allen 32 Videos durchgeführt. Cronbachs Alpha als Wert für die interne Konsistenz über alle Items betrug 0,85.

### 5.2.1 Bewertungstendenz der Rater

Neben den unter der Reliabilitätsanalyse aufgeführten Zusammenhängen zwischen den Ratern soll im Folgenden noch die Bewertungstendenz der Rater stärker hervorgehoben werden.

Abbildung 32 Bewertung der einzelnen Videos



Die vorausgehende Abbildung zeigt die Bewertung gemittelt über alle Items, die die Rater für jedes Video vergaben.

Um zusätzlich einen guten Überblick über die Bewertungstendenz der Rater zu bekommen, wurde eine Boxplot-Darstellung gewählt<sup>22</sup>.

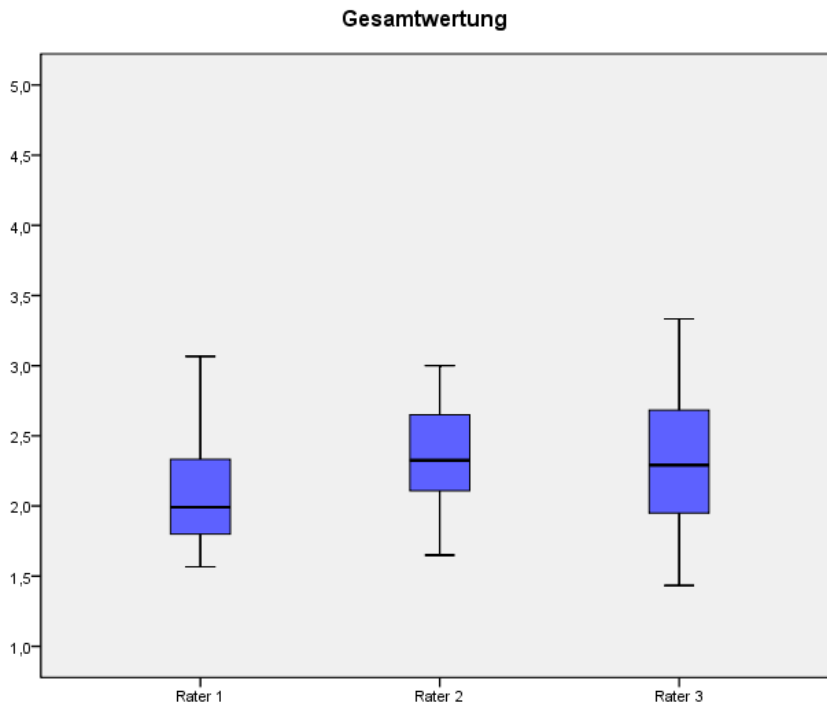
Es zeigt sich für Themenblock A „Gesprächsbeginn“, B „Information bekommen“, Themenblock C „Die Bedenken des Jugendlichen herausfinden“, Themenblock D „Per-

<sup>22</sup> Boxplot-Diagramme haben den Vorteil, dass sie Lage und Verteilung der Werte in komprimierter Form darstellen und man sich so einen raschen Überblick verschaffen kann. Der dicke schwarze Balken kennzeichnet die Lage des Median.



spektive vermitteln“ und Themenblock E „Allgemeine Überlegungen“ sowohl in der Spannweite der Notenvergabe als auch in der Tendenz eine sehr ähnliches Bewertungsmuster. Themenblock C „Die Bedenken des Jugendlichen herausfinden“ zeigt eine größere Abweichung. Rater 2 und Rater 3 urteilen homogener, Rater 1 weicht davon ab.

Abbildung 33 Notendurchschnitt Gesamtwertung



In der Gesamtwertung vergab Rater 1 Noten zwischen 1,57 und 3,07, die Spannweite beträgt 1,5. Der Median liegt bei 1,99.

Rater 2 vergab als beste Note 1,65 als schlechteste 3,0, die Spannweite beträgt 1,35. Der Median liegt bei 2,33.

Rater 3 vergab 1,43 als beste Note und 3,33 als schlechteste. Die Spannweite beträgt 1,9. Der Median liegt bei 2,29.

Hierbei fällt die Tendenz von Rater 1 zur durchschnittlich besseren Bewertung auf.

Rater 3 schöpfte das Notenspektrum am besten aus.

### 5.2.2 Trennschärfe

Ein Wert von 1 steht für eine optimale Korrelation der Bewertung jedes Raters mit der mittleren Bewertung durch die übrigen Rater, ein Wert von 0 für eine fehlende Korrelation. Die korrigierte Item-Skala-Korrelation als Maß für die Trennschärfe liegt für die Gesamtbewertung zwischen 0,64 und 0,78. Die höchste Trennschärfe zeigt sich in

## Ergebnisse

---

Themenblock A „Gesprächsbeginn“ mit Werten zwischen 0,72 und 0,92. Die geringste Trennschärfe zeigt sich für Themenblock C „Bedenken herausfinden“ mit Werten zwischen 0,45 und 0,62.

Betrachtet man die Trennschärfe auf Einzelitemebene so zeigt sich die geringste Trennschärfe für Item 15 „Fachlich korrekt“ mit Werten zwischen 0,13 und 0,22. Die höchste Trennschärfe zeigt sich bei Item 2 „Begrüßung“ mit Werten zwischen 0,90 und 0,95.

**Tabelle 9 Übersicht über Item-Skala-Korrelation**

<b>Korrigierte Item-Skala-Korrelation</b>	<b>Rater 1</b>	<b>Rater 2</b>	<b>Rater 3</b>
<b>A Gesprächsbeginn</b>	0,72	0,92	0,89
<b>B Information bekommen</b>	0,49	0,62	0,72
<b>C Bedenken herausfinden</b>	0,52	0,45	0,62
<b>D Perspektive vermitteln</b>	0,63	0,67	0,71
<b>E Allgemeine Überlegungen</b>	0,56	0,84	0,77
<b>Gesamtbewertung mAGBS-Jugendliche</b>	0,64	0,77	0,76

### 5.2.3 Korrelation zwischen den Ratern

Zusätzlich zur Trennschärfeanalyse wurde eine paarweise Korrelation zwischen den einzelnen Ratern durchgeführt, um eine differenzierte Analyse von möglichen Unstimmigkeiten zwischen den Ratern festzustellen und klar zu differenzieren, welcher Rater am stärksten abweicht. Die Korrelation bezogen auf die Einzelitems findet sich im Anhang. Die folgende Tabelle zeigt, dass Rater 2 und Rater 3 homogener urteilen als Rater 1.

Tabelle 10 Korrelation zwischen den Ratern in den Themenblöcken

Themenblock A	Rater 1	Rater 2	Rater 3	Themenblock B	Rater 1	Rater 2	Rater 3
<b>Rater 1</b>	1,00	0,73	0,69	Rater 1	1,00	0,37	0,54
<b>Rater 2</b>	0,73	1,00	0,94	Rater 2	0,37	1,00	0,71
<b>Rater 3</b>	0,69	0,94	1,00	Rater 3	0,54	0,71	1,00
Themenblock C	Rater 1	Rater 2	Rater 3	Themenblock D	Rater 1	Rater 2	Rater 3
<b>Rater 1</b>	1,00	0,34	0,55	Rater 1	1,00	0,54	0,60
<b>Rater 2</b>	0,34	1,00	0,46	Rater 2	0,54	1,00	0,65
<b>Rater 3</b>	0,55	0,46	1,00	Rater 3	0,60	0,65	1,00
Themenblock E	Rater 1	Rater 2	Rater 3	Gesamtwertung	Rater 1	Rater 2	Rater 3
<b>Rater 1</b>	1,00	0,56	0,52	Rater 1	1,00	0,61	0,61
<b>Rater 2</b>	0,56	1,00	0,86	Rater 2	0,61	1,00	0,79
<b>Rater 3</b>	0,52	0,86	1,00	Rater 3	0,60	0,79	1,00

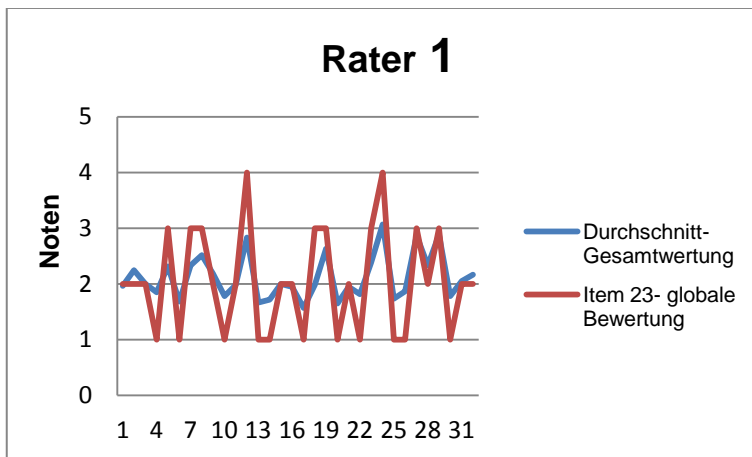
### 5.2.4 Korrelation zwischen–„Globale Bewertung des Arztes“ (Item 23) mit der Durchschnittsnote (Mittelwert über alle Themenblöcke)

Tabelle 11 Pearsonscher Korrelationskoeffizient

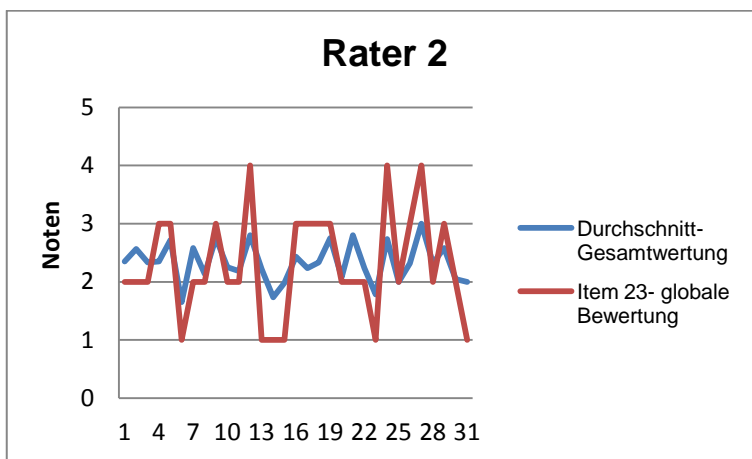
<b>Rater 1: 0,90</b>	<b>Rater 2: 0,48</b>	<b>Rater 3: 0,84</b>
----------------------	----------------------	----------------------

Diese globale Bewertung jedes Raters wurde mit seiner Durchschnittsnote über alle Einzelitems verglichen. Die Korrelation mittels Pearsons Korrelationskoeffizienten zeigt vor allem für Rater 1 und 3 sehr gute Werte.

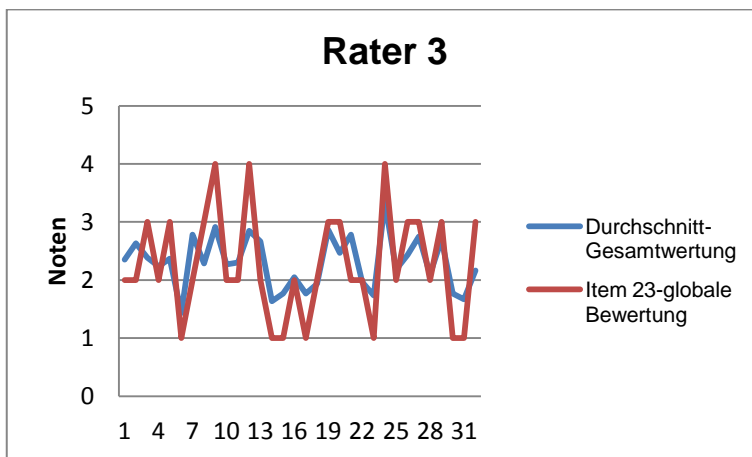
Abbildung 34 Korrelation Durchschnitt Gesamt und globale Bewertung



Die folgenden Abbildungen zeigen den Zusammenhang zwischen der Durchschnittsnote und der globalen Bewertung.



Die Bewertungstendenz zwischen den beiden Noten ist immer die gleiche, jedoch wird bei der globalen Bewertung das Notenspektrum besser ausgeschöpft.



### 5.2.5 Varianzanalyse und Tukeys Additivitätstest

Zur Ermittlung und Differenzierung von systematischen (Inhomogenität der Ratermittelwerte) und unsystematischen Fehlern (Zufallsfehlern bzw. Interaktion von Rater und Proband) wurde eine Varianzanalyse durchgeführt. Mittels p-Wert wurde die Signifikanz der Varianzanalyse angegeben.

## Ergebnisse

---

Für Item 1 war aufgrund zu geringer Varianz (siehe hierzu die Notenverteilung im Histogramm) keine Reliabilitätsanalyse möglich.

Tabelle 12 Varianzanalyse für die Themenblöcke

<b>Varianzanalyse</b>					
	Leistung des Arztes	Inhomogenität der Ratermittelwerte	p-Wert	Interaktion Rater/Arzt	p-Wert
<b>Gesamt</b>	72%	6%	,001	22%	,912
<b>Block A - Gesprächsbeginn</b>	85%	1%	,008	14%	,020
<b>Block B – Information bekommen</b>	65%	3%	,726	32%	,100
<b>Block C - Bedenken herausfinden</b>	57%	10 %	,000	33%	,685
<b>Block D – Perspektive vermitteln</b>	66%	8%	,000	26%	,079
<b>Block E - Allgemeine Überlegungen</b>	71%	12%	,002	15%	,091

Die Varianz der Gesamtbewertung ist zu 72 % auf die Leistung der Gesprächsführenden in der Arztrolle zurückzuführen. In den einzelnen Themenblöcken bewegt sich die Varianz, die auf die Leistung des Arztes geht, zwischen 85 % (Themenblock A) und 57 % (Themenblock C). Die größte Inhomogenität der Ratermittelwerte findet sich im Themenblock E. Der auf einen unsystematischen Fehler zurückzuführende Varianzanteil (Interaktion Rater/Arzt) liegt zwischen 14 und 33 %.

## 5.2.6 Intraclasskorrelation

Tabelle 13 ICC über Einzelitems

<b>Unjustierte Intraclasskorrelation</b>	<b>zufällige ICC</b>
	ICC
<b>Block A - Gesprächsbeginn</b>	<b>0,92</b>
Item 1 - Umgebung	xxx
Item 2 - Begrüßung	0,96
Item 3 - Befinden	0,90
<b>Block B – Information bekommen</b>	<b>0,78</b>
Item 4 - Gesprächseinstieg	0,78
Item 5 - Annäherung	0,77
Item 6 - Sprache	0,71
Item 7 - Pausen	0,54
<b>Block C – Bedenken herausfinden</b>	<b>0,65</b>
Item 8 - Nachfragen	0,53
Item 9 - Nonverbal	0,70
<b>Block D – Perspektive vermitteln</b>	<b>0,77</b>
Item 10 - Äußerungen	0,54
Item 11 - Struktur	0,77
Item 12 - Wiederholungen	0,79
Item 13 - Positive Aspekte	0,76
Item 14 - Empathie	0,66
Item 15 - Fachlich korrekt	0,26
<b>Block E – Allgemeine Überlegungen</b>	<b>0,82</b>
Item 16 Schlüsselbereiche	0,72
Item 17 - Tiefgreifendes	0,77
Item 18 - Unterstützen	0,73
Item 19 - Körpersprache	0,70
Item 20 - Sicherheit	0,78
Item 21 - Gesprächstempo	0,46
Item 22 - Beenden	0,91
<b>Gesamtwertung</b>	<b>0,82</b>

Der ICC für die einzelnen Themenblöcke und die Gesamtwertung liegt in einem guten Bereich und zeigt damit ein zufriedenstellendes Ergebnis. Lediglich Themenblock C „Bedenken des Jugendlichen herausfinden“ liegt mit 0,66 etwas niedriger als die anderen Themenblöcke. Der ICC der einzelnen Items zeigt einen differenzierteren Blick. Bei Item 1 „geeignete Umgebung“ war keine Berechnung des ICCs möglich, da es fast ausschließlich sehr gute Bewertungen gab und damit keine Varianz gezeigt wurde. Item 2 „Begrüßung“, Item 3 „Befinden“, Item 22 „Beenden“ liefern sehr hohe ICCs. Dies deckt sich auch mit der persönlichen Einschätzung der Autorin, dass diese Items gut zu operationalisieren sind. Item 7 „angemessene Pausen“, Item 8 „Nachfragen bei Problemen“, Item 10 „Äußerungen angepasst“, Item 15 „Fachliche Korrektheit“ und Item 21 „Gesprächstempo“ zeigen mit ICCs zwischen 0,26 und 0,54 geringe Übereinstimmungen.

Item 15 „Fachlich korrekt“ zeigt aufgrund der geringen Varianz zwischen den Videos einen niedrigen ICC.

Keine hohe Übereinstimmung zeigt der ICC von Item 7 „Angemessene Pausen“ und Item 21 „Gesprächstempo“.

Diese Items beurteilen Gesprächsfertigkeiten, die auf den ersten Blick sehr objektiv scheinen. Doch die Ansicht, ob Pausen oder das Gesprächstempo angemessen sind, scheint bei den 3 Ratern unterschiedlich zu sein.

### **5.3 Unterschiede zwischen den drei Trainingsgruppen**

Es wurden Videos mit unterschiedlichen Situationen zur Bewertung verwendet. Es waren 11 Videos zu einem Erstgespräch zwischen Arzt und Jugendlichen und 11 Videos zur Notfallsituation „Suizid“. Diese Videos waren mit Medizinstudenten jeweils am Anfang und Ende eines 4-tägigen Trainings aufgenommen worden. 10 Videos waren von angehenden Allgemeinärzten zur Jugendgesundheitsuntersuchung, die nur einmalig ein 2-stündiges Training bekommen hatten.

#### **5.3.1 Unterschiede in der Beurteilung durch die mAGBS-Jugendliche**

Um die Unterschiede zwischen den Gruppen herauszuarbeiten, wurden der Mittelwert und die Standardabweichung über alle drei Rater getrennt nach Videogruppen berechnet. Es zeigten sich Unterschiede zwischen den Gruppen je nach Einzelitems.

So war bei manchen Items die Jugendlichengruppe, bei manchen die Notfallgruppe, und auch die Allgemeinärztegruppe etwas besser in der Benotung.

Unterschiede in der Notenstufe  $\leq 0,1$  wurden nicht als Unterschied gewertet. Hingewiesen sei noch einmal darauf, dass die Skala von 1(= sehr gut) bis 5 (= schlecht) geht.

Die Effektstärke wurde zwischen den Jugendlichenvideos und den Notfallvideos berechnet, um die Bedeutsamkeit des Unterschiedes auszudrücken. Anschließend wird die Effektstärke zwischen den Jugendlichenvideos und Allgemeinärztevideos verglichen, da dies in diesem Fall sehr gut möglich ist. Allerdings bringen die Allgemeinärzte Berufserfahrung mit.

##### **5.3.1.1 Einfluss von Training**

Die Jugendlichenvideos entstanden zu Beginn des 4-tägigen Trainings, die Notfallvideos am Ende.

Es zeigte sich für die Items 12 „Wiederholungen“, Item 15 „Fachlich korrekt“, Item 17 „Tiefgreifendes Besprechen“, Item 18 „Unterstützung“, Item 19 „Körpersprache“, Item 21 „Gesprächstempo“ kein Unterschied zwischen den beiden Gruppen „Jugendliche“ und „Notfall“.

Bei den Items 2 „Begrüßung“, Item 3 „Befinden“, Item 4 „Gesprächseinstieg“, Item 7 „Pausen“, Item 8 „Nachfragen“ und Item 9 „Nonverbal“ schnitt die Gruppe der Notfallvideos, also die trainierte Gruppe, besser ab. Die Effektstärke ist vor allem bei Item 2 „Begrüßung“ und Item 4 „Gesprächseinstieg“ mit 0,54 bzw. 0,51 relativ hoch. Bei Item 8 „Nachfragen“ beträgt die Effektstärke 0,43.

## Ergebnisse

Bei Item 5 „Annäherung“, Item 6 „Sprache“, Item 10 „Äußerungen“, Item 11 „Struktur“ und Item 16 „Schlüsselbereiche“ schnitt die Jugendlichengruppe besser ab.

Die folgende Tabelle gibt dazu einen Überblick.

**Tabelle 14 Überblick Mittelwerte und Effektstärke bei Jugendlichen- und Notfallvideos**

	Jugendliche (Medizinstudenten)		Effektstärke	Notfall (Medizinstudenten)	
	MW	STD		MW	STD
Item 1- Umgebung	1,00	0,00	k.U.	1,03	0,27
Item 2-Begrüßung	3,03	0,88	<b>0,54</b>	2,45	1,28
Item 3-Befinden	2,33	1,56	<b>0,26</b>	1,97	1,26
Item 4-Gesprächseinstieg	2,85	1,15	<b>0,51</b>	2,30	0,98
Item 5-Annäherung	1,21	0,42	1,30	1,97	0,68
Item 6-Sprache	1,97	0,31	0,55	2,21	0,55
Item 7- Pausen	1,39	0,66	<b>0,16</b>	1,30	0,53
Item 8-Nachfragen	2,21	0,82	<b>0,43</b>	1,88	0,70
Item 9-Nonverbal	3,24	0,61	<b>0,33</b>	3,00	0,83
Item 10-Äußerungen	2,82	1,07	0,25	3,06	0,86
Item 11-Struktur	1,88	0,49	0,31	2,06	0,66
Item 12-Wiederholungen	3,33	0,92	k.U.	3,24	0,79
Item 13-Positive Aspekte	3,15	0,83	0,36	2,82	0,98
Item 14-Empathie	2,88	0,74	<b>0,23</b>	2,67	1,11
Item 15-Fachlich korrekt	1,39	0,50	k.U.	1,45	0,51
Item 16-Schlüsselbereich	1,82	0,39	0,86	2,36	0,86
Item 17- Tiefgreifendes	1,24	0,44	k.U.	1,36	0,55
Item 18-Unterstützung	2,03	0,64	k.U.	2,09	0,95
Item 19-Körpersprache	1,85	0,83	k.U.	1,85	1,03
Item 20-Sicherheit	1,82	0,73	k.U.	1,73	0,76
Item 21-Gesprächstempo	1,42	0,79	k.U.	1,36	0,65
Item 22-Beenden	2,03	1,31	0,18	2,27	1,28

### 5.3.1.2 Einfluss von Berufserfahrung

Vergleicht man die Jugendlichengruppe und die Allgemeinarztgruppe, so zeigt sich, dass zumeist die Jugendlichengruppe etwas besser bewertet wurde oder es keinen Unterschied zwischen den Gruppen gibt.

Bessere Bewertung erhielten die Allgemeinärzte bei Item 4 „Gesprächseinstieg“, Item 12 „Wiederholungen“, Item 13 „Positive Aspekte“, Item 14 „Empathie“ sowie Item 19 „Körpersprache“.



Tabelle 15 Überblick Mittelwert und Effektstärke bei Jugendlichen- und Allgemeinärztevideos

	Jugendliche Medizinstudenten)		Effekt- stärke	Allgemeinärzte	
	MW	STD		MW	STD
Item 1- Umgebung	1,00	0,00	k.u.	1,07	0,25
Item 2-Begrüßung	3,03	0,88	k.u.	3,20	0,81
Item 3-Befinden	2,33	1,56	<b>0,57</b>	3,13	1,22
Item 4-Gesprächseinstieg	2,85	1,15	<b>0,51</b>	2,33	1,12
Item 5-Annäherung	1,21	0,42	<b>0,45</b>	1,50	0,86
Item 6-Sprache	1,97	0,31	<b>0,60</b>	2,27	0,69
Item 7- Pausen	1,39	0,66	0,30	1,63	0,93
Item 8-Nachfragen	2,21	0,82	k.u.	2,20	0,89
Item 9-Nonverbal	3,24	0,61	k.u.	3,27	0,45
Item 10-Äußerungen	2,82	1,07	k.u.	2,80	0,89
Item 11-Struktur	1,88	0,49	0,19	2,00	0,79
Item 12-Wiederholungen	3,33	0,92	<b>0,47</b>	2,93	1,12
Item 13-Positive Aspekte	3,15	0,83	1,19	2,07	0,98
Item 14-Empathie	2,88	0,74	0,22	2,67	0,81
Item 15-Fachlich korrekt	1,39	0,50	0,18	1,50	0,63
Item 16-Schlüsselbereich	1,82	0,39	<b>0,67</b>	2,07	0,36
Item 17- Tiefgreifendes	1,24	0,44	<b>0,54</b>	1,57	0,82
Item 18-Unterstützung	2,03	0,64	k.u.	1,97	0,81
Item 19-Körpersprache	1,85	0,83	0,29	1,60	0,89
Item 20-Sicherheit	1,82	0,73	k.u.	1,80	0,88
Item 21-Gesprächstempo	1,42	0,79	0,41	1,77	0,90
Item 22-Beenden	2,03	1,31	k.u.	2,13	1,28

Die beiden Videogruppen Jugendliche und Notfall - also die Medizinstudenten - wurden gegenüber der Gruppe der Allgemeinärzte insgesamt besser bewertet. Mit Ausnahme von Item 12 „Wiederholungen“, Item 13 „Positive Aspekte“, sowie Item 19 „Körpersprache“, dort erhielten die berufserfahrenen Allgemeinärzte eine bessere Bewertung.

## 5.4 Gesprächsverläufe

Es zeigt sich, dass Items wie Item 7 „ Angemessene Pausen“, Item 8 „Nachfragen bei Problemen“, Item 10 „Äußerungen angepasst“, Item 15 „Fachlich korrekt“ sowie Item 21 „Gesprächstempo“ – alles Items, die formale Gesprächskriterien abdecken, die niedrigsten Übereinstimmungen zwischen den Rater hervorbrachten (siehe hierzu das vorangehende Kapitel 5.2.6 Intraclasskorrelation).

Um hier einen weiteren Zugang zu bekommen, wurden deswegen alle Videos mit Interact auf Gesprächsverläufe, Gesprächsanteile, Fragestil, nicht unterbrochene Pati-

entenäußerungen, Gesprächspausen und –tempo sowie Hörersignale untersucht. Dabei wurde explizit darauf geachtet, ob und zu welchem Zeitpunkt es dem Arzt gelungen war, den Jugendlichen zum Reden zu bringen.

Diese Stellen können als „Schlüsselstellen“ im Gespräch gesehen werden (Frank et al., 2009). Anhand der danach folgenden Einzelbeispiele wurde untersucht, welche Bedeutung das Stellen offener Fragen, Pausen und Gesprächstempo auf die Entwicklung des Gespräches haben können.

### **5.4.1 Gesprächsverläufe in den Videogruppen und Fragestil**

Bei allen Jugendlichenvideos fällt fast ausschließlich folgendes Kommunikationsmuster auf: zu Beginn und am Schluss redet der Arzt viel, um eine Einführung und einen Abschluss des Gesprächs zu geben. Es folgt eine Phase in der der Arzt Fragen stellt und der Jugendliche darauf Antwort gibt, häufig sind das geschlossene Fragen.

Fast in jedem Video gibt es eine Phase, in der der Jugendliche 15 Sekunden am Stück spricht, dies sind je nach Sprechgeschwindigkeit drei bis vier Sätze. Dies ist meist dann der Fall, wenn der Arzt eine offene Frage stellt oder eine Pause entstehen lässt, oder bei einem Thema bleibt und sich es näher erläutern lässt.

Diese Stelle, ab der der Jugendliche spricht, kann als Schlüsselstelle des Gespräches betrachtet werden.

Es folgt ein Beispiel für 15 Sekunden Redezeit. Eine Jugendliche hat zuvor bei der Frage nach „Schule“ gemeint, es sei schwierig mit Lehrern. Es scheint, dass der Arzt hier ein Thema gefunden hat, das der Jugendlichen wichtig ist und sie hier erstmals mehr als einen Satz spricht.

*Arzt: „Und wie ist es mit Mitschülern“? (offene Frage)*

*Jugendliche: „Ja, also mit denen komme ich gut aus, da gibt's eigentlich keine Probleme. Ich hab zwei gute Freunde, neben denen sitze ich auch. Wir treffen uns auch nachmittags.“*

Bei der Suchtanamnese greifen alle Ärzte wieder auf die geschlossene Fragestellung zurück, dies stellt jedoch auch ein Trainingsziel des Wahlfaches Kinder- und Jugendpsychiatrie dar.

*Arzt: „Rauchst du?“*

*Jugendliche: „Ja“*

*Arzt: „Wie viel rauchst du denn am Tag?“*

*Jugendliche: „Hmm, so ungefähr 10 Zigaretten.“*

*Arzt: „Rauchst du auch schon morgens nach dem Aufstehen?“*

*Jugendliche: „Ja meistens“.*

*Arzt: „Würdest du sagen du bist süchtig?“*

*Jugendliche: „Weiß ich nicht“.*

Die Studenten sollen lernen, dass offene Fragen nicht immer sinnvoll sind, sondern an manchen Stellen wie im vorangehenden Beispiel gezielte Information benötigt wird. Der Arzt muss eine realistische Einschätzung vom Suchtverhalten der Jugendlichen bekommen.

Der Gesprächseinstieg jedoch gelingt am besten über eine offene Frage. An einigen Videos wird auch deutlich, dass die Gesprächsführenden sehr bemüht sind, alle relevanten Themen anzusprechen. Diese Gespräche ähneln dann eher einem „Verhör“. Es dauert länger bis der Arzt zum Thema des Jugendlichen stößt. Die Herausforderung für die Gesprächsführenden ist es, letztendlich ein Thema zu finden, das den Jugendlichen beschäftigt und für ihn Bedeutung hat.

Bei den Notfallvideos dauert es meist einige Zeit bis der Jugendliche spricht, aber die Redepassagen der Jugendlichen sind dann länger. Der Arzt hat ein Thema, den Suizidversuch des Jugendlichen, das sich ihm als Problem anbietet.

Bei den Allgemeinärzten ist als Kommunikationsmuster auffällig, dass die Gespräche meist stark gelenkt werden und die Ärzte im Gespräch dominieren.

Die folgende Tabelle zeigt für jedes Video stichwortartig wesentliche Auffälligkeiten und die durchschnittliche Bewertung als Mittelwert der 3 Rater, die das Video mittels der mAGBS-Jugendliche erzielte.

## Ergebnisse

Tabelle 16 Gesprächsstil des Jugendlichen in den einzelnen Videos

Jugendlichenvideos	Bewertung (MW 3 Rater)	Bemerkung
1.Video	2,22	Ärztin fasst viel zusammen;  Es gibt einen Punkt ab dem Jugendlicher spricht  bei 3:10 nach einer <b>offenen Frage</b> der Ärztin: „Wie ist es mit deinen Eltern?“ kurze Pause, dann Antwort mit Länge 20 Sekunden; Schlüsselmoment: Ärztin hat Thema gefunden, bleibt beim Thema, wechselt nicht so schnell; aber <b>geschlossene Fragen</b> bei Suchthemen
2.Video	2,48	Anfangs viele <b>geschlossene Fragen</b> , dann mehrere Pausen. Bei 1:36 spricht Jugendliche 18 Sekunden am Stück auf Frage des Arztes: „Streitest du mit deiner Mutter?“ Arzt am Ende sehr ausführlich, gibt Zukunftsperspektive
3.Video	2,24	Viele Suggestivfragen, an diesem Video wird klar: wenn es darum geht <b>Antworten</b> zu bekommen, werden <b>offene Fragen</b> sehr wichtig; Jugendliche spricht auf Frage „Wie sind deine Eltern dazu eingestellt?“ 10 Sekunden am Stück
4.Video	2,14	Ärztin spricht schnell, wirkt hektisch, <b>viele geschlossene Fragen</b>  Ärztin spricht vor allem zu Beginn und am Schluss sehr viel, Jugendliche kommt nur mit 1 Sekunden Antworten zu Wort,erste <b>längere Antwort auf offene Frage</b> : „Wie ist es mit Mitschülern?“, Länge 11 Sekunden
5.Video	2,47	Gespräch ähnelt einem „Verhör“, viele <b>geschlossene Fragen</b> , Ärztin spricht viel zu Beginn und am Schluss, gibt Informationen zur Gesundheit, Jugendlicher spricht dreimal 15 Sekunden am Stück, Ärztin findet ein Thema (Sport) das den Jugendlichen begeistert
6.Video	1,59	Ärztin spricht zu Beginn sehr viel, dann folgt eine <b>offene Frage</b> „Hast du einen Grund, warum du hier bist?“ Sie

## Ergebnisse

		findet „Kernproblem“ der Jugendlichen, lässt es sich genau erklären, viele Hörsignale, dann Themenwechsel, <b>geschlossene Fragen</b> am Ende, Abschluss/Zusammenfassung
7.Video	2,57	Ärztin spricht zu Beginn und am Ende viel, stellt Fragen, die Jugendliche eher knapp beantwortet, „hakt“ Themen ab: Familie, Freunde, Sucht; Jugendliche spricht länger auf die Frage: „Wie ist es mit Freunden?“
8.Video	2,31	Jugendliche kommt oft zu Wort, Fragen werden von der Ärztin näher erläutert, aber <b>viele geschlossen Fragen</b> und Suggestivfragen, Ärztin sendet viele Hörsignale Ärztin deckt Probleme der Jugendlichen auf: „Wie ist es mit Freunden?“ „Ich bin nicht beliebt in der Schule“, „Ich hätte auch gern mal nen Freund“
9.Video	2,62	Viele geschlossene Fragen und Suggestivfragen, Jugendliche antwortet eher kurz, dann findet Arzt ein Thema „strenge Eltern“, Lob: „Freut mich, dass du so ehrlich warst“
10.Video	2,10	Viele offene Fragen, gute Einleitung, guter Führungsstil, Jugendliche spricht gleich zu Beginn 15 Sekunden auf die Frage „ <b>Warum bist du hier?</b> “
11.Video	2,16	Viele Fragen aneinander gereiht, bei Videozeit 40 Sekunden spricht Jugendliche 15 Sekunden auf die Frage „Was beschäftigt dich?“ Am Ende bei Suchtanamnese <b>gezielte, geschlossene Frage</b>
<b>Notfallvideos</b>		
1.Video	2,83	Anfangs kurze, geschlossene Fragen, auf die eine kurze Antwort folgt, dann <b>offene Frage</b> : „Macht es dich traurig?“ Kaum Hörsignale
2.Video	2,18	Viele Pausen, bei Minute 4 des Gesprächs spricht Jugendliche <b>1 Minute am Stück</b> auf die Frage: „Wovor wolltest du dich absichern?“
3.Video	1,69	Sehr ruhige Gesprächsführung, oft Hörsignale, offene Frage, schneller Einstieg ins Thema, Frage: „Gelingen dir

## Ergebnisse

		oft Sachen nicht?“
4.Video	1,92	Ärztin wiederholt Antworten der Jugendlichen sehr oft, viele kurze Antworten der Jugendlichen bei 3:30 des Videos „Was war letzte Nacht?“ folgt längere Antwort der Jugendlichen
5.Video	2,14	Offene Fragen, Jugendliche spricht relativ viel, kein langer Einstieg nötig, Schlüsselfrage: „Warum geht es dir schlecht, warum willst du nicht mehr leben?“ Ab der Hälfte des Gesprächs werden Antworten insgesamt kürzer
6.Video	1,87	Jugendliche spricht sehr lange gar nicht, es gibt einen Punkt, ab dem sie spricht. Frage des Arztes: „Wenn wir dich verstehen, können wir dir besser helfen“. Es folgt eine <b>Pause</b> Jugendliche fängt an zu sprechen, Arzt stellt nur knappe Zwischenfragen, Jugendliche öffnet sich gefühlsmäßig
7.Video	2,08	Viele offene Fragen, Jugendliche sehr gesprächig, Arzt unterbricht sie bei 3:45, Arzt fragt genauer nach „Was hat deine Mutter gemeint?“
8.Video	2,75	Viele geschlossene Fragen, Ärztin wirkt etwas unsicher
9.Video	2,07	Jugendliche wird nicht zum Reden gebracht, Ärztin übernimmt das Reden für die Jugendliche
10.Video	2,51	Jugendliche spricht viel; erst offene, dann geschlossene Fragen
11.Video	2,01	Jugendliche wirkt gereizt, Arzt wiederholt sich inhaltlich oft
<b>Allgemeinärzte</b>		
1.Video	1,97	Viele geschlossene Fragen
2.Video	3,04	Sehr kurzes Gespräch, Arzt wirkt gehetzt, er fällt dem Jugendlichen ins Wort
3.Video	1,97	Ärztin spricht viel, ist sehr direkt, inhaltlich sehr gut, tritt sehr sicher und souverän auf, „Was machst du in deiner Freizeit?“

## Ergebnisse

4.Video	2,21	Lange Erklärung, gute Einführung, offene Fragen „Gibt es Probleme in der Schule?“
5.Video	2,87	Viele geschlossene Fragen seitens der Ärztin
6.Video	2,25	Viele Suggestivfragen
7.Video	2,72	Ärztin dominant im Auftreten
8.Video	1,87	Arzt sehr souverän, fasst viel zusammen
9.Video	1,91	„Small talk“ zwischen Arzt und Jugendlichen,
10.Video	2,42	Jugendliche spricht viel, Arzt stellt nur kurze offene Fragen

### 5.4.2 Redeanteile, Fragestil, nicht unterbrochene Äußerungen

Es wurde für jedes Video der Gesprächsanteil von Arzt und Patient gemessen, sowie der Fragestil bewertet, indem zwischen offenen und geschlossenen Fragen des Arztes unterschieden wurde. Die folgende Tabelle gibt zusammengefasst nach Videogruppen durchschnittliche Werte an. Eine ausführliche Auflistung für jedes Video findet sich im Anhang.

Tabelle 17 Durchschnittlicher Redeanteil und Fragestil

Videoart	Durchschn.Redanteil Arzt	Redeanteil Jugendl.	STD.	Geschlossene Fragen	Offene Fragen	STD.
Jugendliche	46%	54%	6,48	65%	35%	14,99
Notfall	48%	52%	10,7	<b>53%</b>	<b>47%</b>	13,44
Allgemeinarzt	<b>63%</b>	37%	13,3	62%	38%	16,50

In der Gruppe der Allgemeinärzte liegt der Redeanteil seitens der Ärzte etwas höher als bei den Studenten. Bei den Jugendlichen – und Notfallvideos ist der Gesprächsanteil zwischen Arzt und Jugendlichen relativ ausgewogen.

Bezüglich des Fragestils ist in der Gruppe der Notfallvideos die Anzahl der offenen Fragen mit 47% am höchsten.

Auch hinsichtlich der nicht unterbrochenen Patientenäußerungen zeigen sich Unterschiede zwischen den Gruppen, bei den Allgemeinärzten kommt es bei 4 Videos vor, dass die Jugendlichen nicht 15 Sekunden am Stück sprechen, bei 5 Videos sprechen

## Ergebnisse

---

sie einmal 15 Sekunden, nur bei einem Video spricht der Jugendliche 4-mal mehr als 15 Sekunden am Stück. Bei den Jugendlichenvideos gibt es nur bei zwei Videos die Situation, dass der Jugendliche nicht einmal 15 Sekunden am Stück spricht, bei den Notfallvideos nur einmal. Vor allem bei den Notfallvideos sprechen die Jugendlichen häufig 2-3 mal 15 Sekunden am Stück.

Es zeigt sich mittels Spearmans Korrelation ein geringfügiger Zusammenhang zwischen der Anzahl an offenen Fragen und dem Redeanteil des Jugendlichen. Einen deutlicheren Zusammenhang gibt es zwischen der Anzahl an offenen Fragen und der Anzahl zusammenhängender Jugendlichenäußerungen.

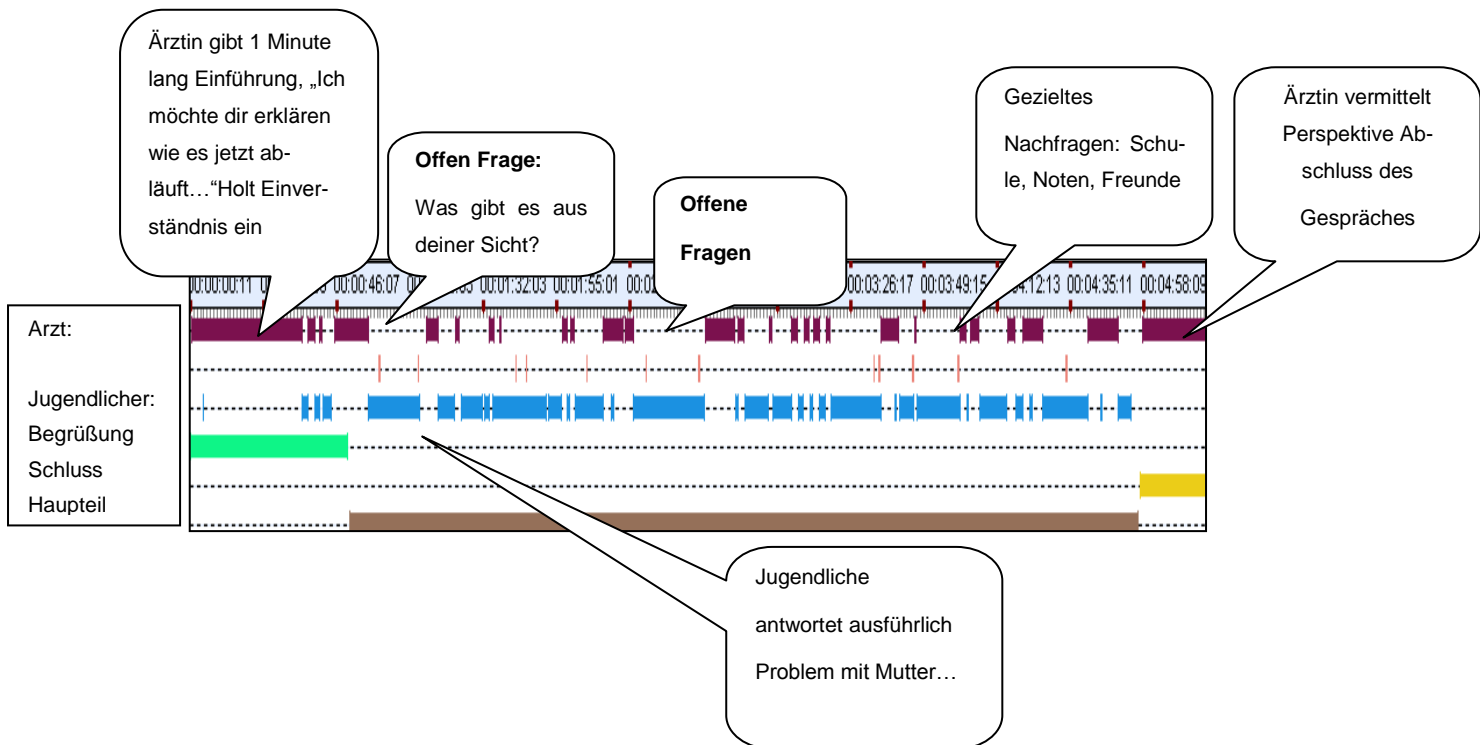
**Tabelle 18 Korrelationen zwischen Fragestil und Redeanteil bzw. nicht unterbrochener Äußerungen Jugendlicher**

<b>Korrelation offene Fragen Arzt mit Redeanteil Jugendlicher</b>	<b>Korrelation geschlossene Fragen Arzt mit Redeanteil</b>	<b>Korrelation offene Fragen Arzt mit Anzahl nicht unterbrochener Jugendlichenäußerung</b>
0,12	-0,12	<b>0,30</b>

Noch besser lässt sich dieser Zusammenhang an Einzelbeispielen hervorheben. Im folgenden Video spricht die Jugendliche über lange Passagen. Insgesamt hat sie einen Redeanteil 56% wobei die Ärztin beim Fragestil 59% offene Fragen stellt.

Auffällig sind vor allem die langen Passagen nicht unterbrochener Äußerungen der Jugendlichen. Die Jugendliche antwortet einmal 15 Sekunden, zweimal 20 Sekunden und einmal sogar 30 Sekunden am Stück – jeweils auf offene Fragen.





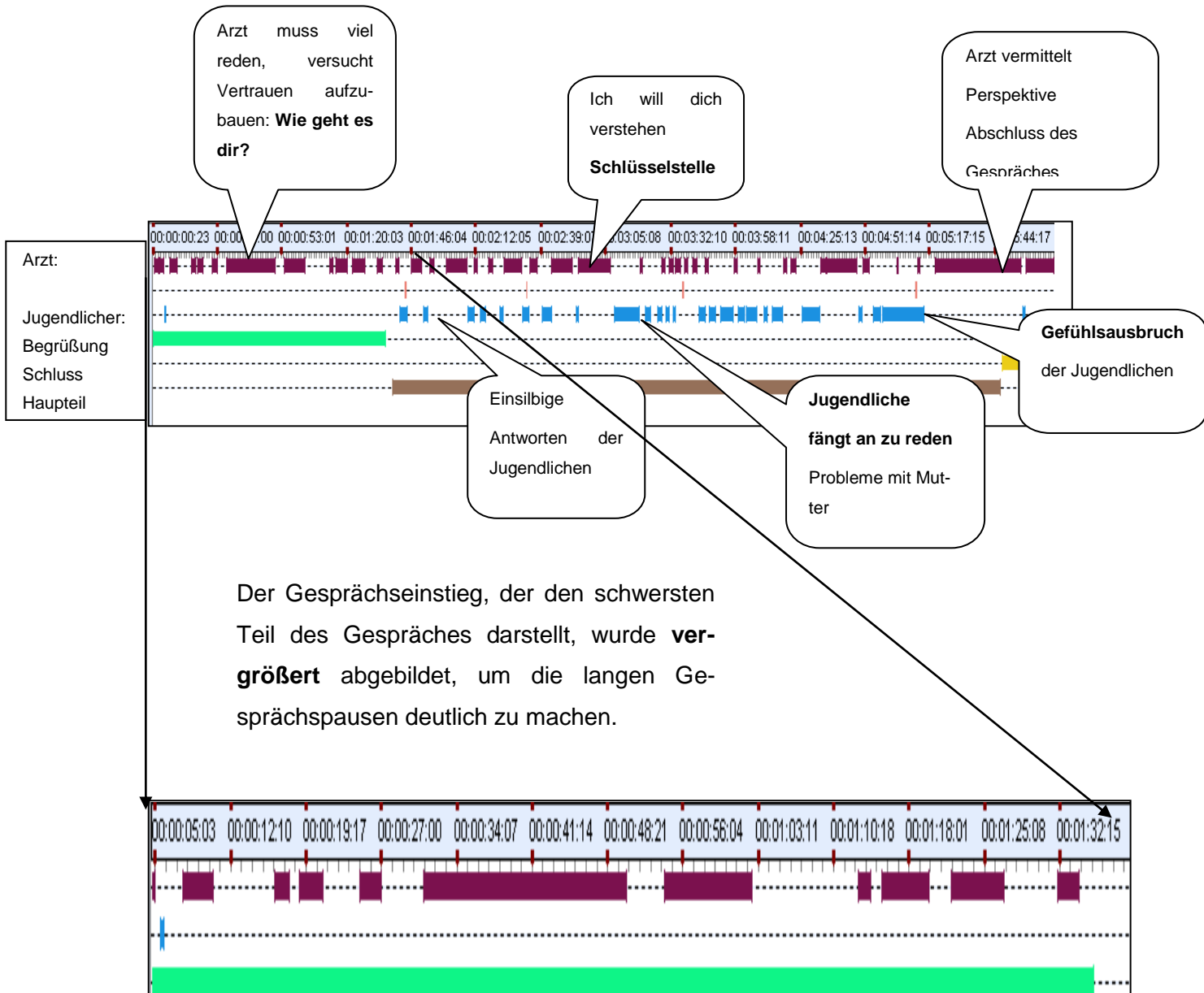
## 5.4.3 Pausen im Gespräch

Interact liefert gute Möglichkeiten, über eine grafische Darstellung der einzelnen Redeanteile auf Pausen im Gespräch zu achten und diese in ihrer Länge zu messen.

Im folgenden Beispiel gelingt es dem Gesprächsführenden in hervorragender Weise Pausen zu zulassen und so mit viel Geduld den Jugendlichen zum Reden zu bringen.

Es geht um den Suizidversuch einer Jugendlichen. Die Arztrolle ist männlich besetzt.

Zu Beginn des Gespräches versucht der Arzt, eine Verbindung zur Jugendlichen herzustellen. Die Studentin in der Rolle der Jugendlichen versetzt sich sehr überzeugend in eine Jugendliche hinein, die bedingt durch die Probleme mit einer alkoholkranken Mutter wenig Vertrauen in ihre Umgebung hat und lieber schweigt.



Wie die Vergrößerung zeigt, sind bis weit über die ersten Minuten des Gespräches hinaus 10 Pausen von mindestens 1 Sekunde bis hin zu 10 Sekunden maximal enthalten. Der Arzt versucht zwischendurch immer wieder die Jugendliche zum Reden zu bringen, er zeigt jedoch immer Verständnis und redet nicht pausenlos, sondern lässt bewusst die langen Pausen zu.

Mittels der mAGBS–Jugendliche wurde das Gespräch insgesamt von zwei Ratern mit 1,57 und 1,77 sehr hoch bewertet. Ein Rater vergab in der Durchschnittsnote über alle Items 2,23.

Die Notenvergabe bei den beiden Items, die Gesprächspause und Tempo bewerten, war jedoch unterschiedlich gut.

Tabelle 19 Notenvergabe Item 7 und Item 21 für 6.Video

	Item 7 „Pausen“	Item 21 „Gesprächstempo“
Rater 1	Note 1	Note 1
Rater 2	Note 3	Note 3
Rater 3	Note 1	Note 1

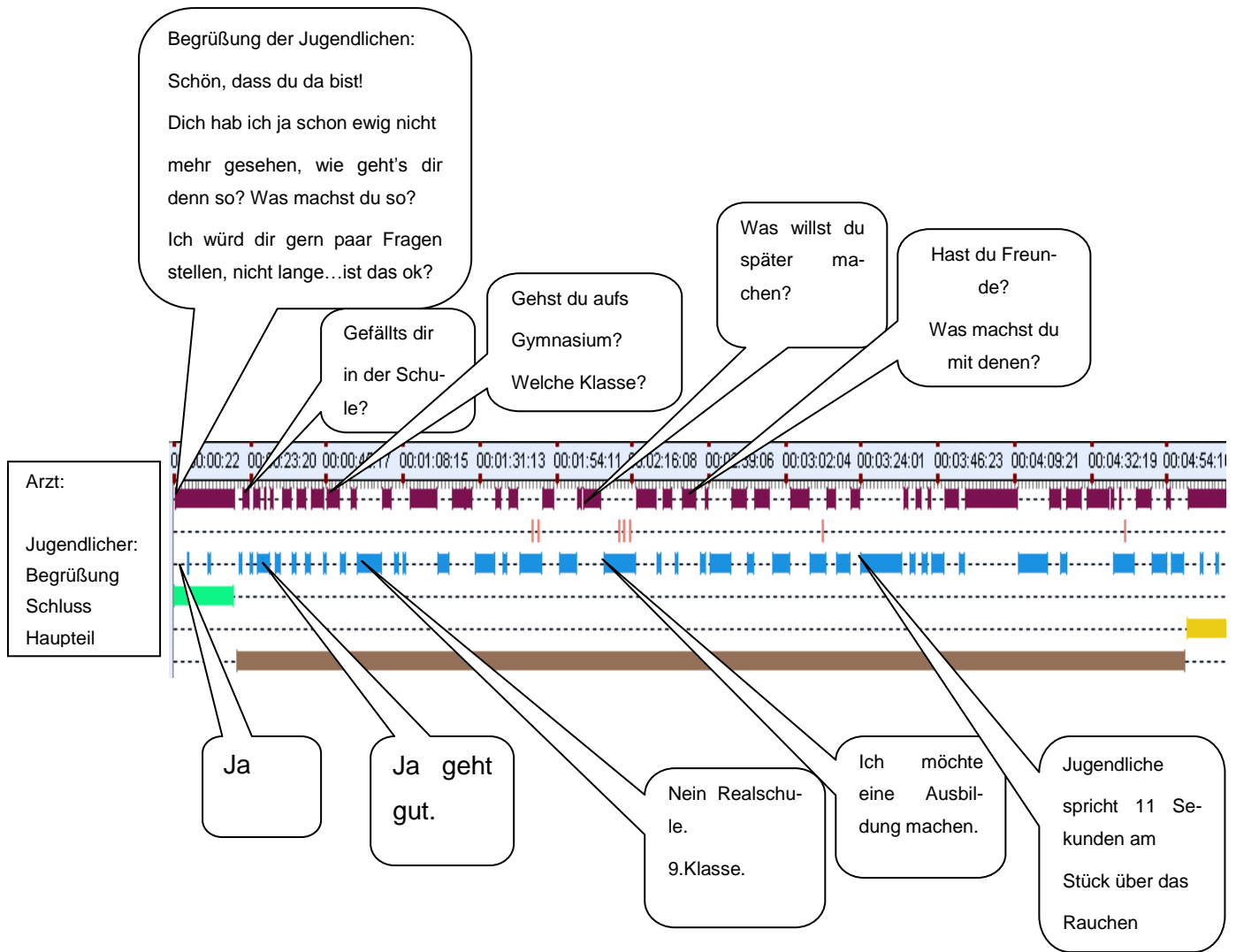
#### 5.4.4 Gesprächstempo

Am folgenden Beispiel soll gezeigt werden, dass die Ärztin durch das vor allem anfängliche Aneinanderreihen vieler Fragen ein hohes Gesprächstempo hineinbringt.

Bei diesem Gespräch fällt auf, dass es kaum Pausen im Gespräch gibt. Vor allem zu Beginn stellt die Ärztin auf jede Antwort der Jugendlichen sofort eine neue Frage.

Die längste Pause im Mittelteil des Gespräches ist 1 Sekunde lang, ansonsten finden sich Pausen mit Länge einer viertel bis halben Sekunde. Es findet sich mit 86 % auch ein sehr hoher Anteil an geschlossenen Fragen. Der Redeanteil der Jugendlichen beträgt 45 %, sie spricht nur einmal 11 Sekunden am Stück.

# Ergebnisse



Wieder beurteilten die drei Rater Pausen und Gesprächstempo uneinheitlich mittels der mAGBS-Jugendliche.

Tabelle 20 Notenvergabe für 4.Video

	Item 7 „Pausen“	Item 21 „Gesprächstempo“
<b>Rater 1</b>	Note 2	Note 2
<b>Rater 2</b>	Note 3	Note 4
<b>Rater 3</b>	Note 3	Note 3

#### **5.4.5 Hörersignale**

Die Anzahl der Hörersignale schwankt sehr stark, wobei hinzugefügt werden muss, dass ein Hörersignal nur als solches gewertet wurde, wenn vom Arzt eine akustische Äußerung wie „hm, ja, aha, okay...“, gemacht wurde. Nonverbale Äußerungen wurden nicht mit einbezogen. So gab es bei den Jugendlichenvideos eine durchschnittliche Anzahl von Hörersignalen von 15, bei den Notfallvideos von 9 und bei den Allgemeinärzten von 8.

## 6 Diskussion

Die vorliegende Arbeit wurde im Rahmen von Trainingssituationen für ärztliche Gesprächsführung durchgeführt. Ein Teil der auf Video aufgenommenen und bewerteten Gespräche stammt aus dem Wahlfach Kinder- und Jugendpsychiatrie, ein Teil aus dem „Kompaktkurs Pädiatrie“ für angehende Allgemeinärzte.

Die Arbeit bestand aus zwei Teilen. Der erste Teil umfasst die Weiterentwicklung und Reliabilitätstestung einer bereits validierten Bewertungsskala zur Evaluation der Arzt-Patient-Kommunikation. Der zweite Teil befasste sich ausgehend von den Ergebnissen der Reliabilitätstestung der modifizierten Bewertungsskala mit Fragen zu Kommunikationsstrukturen in der Arzt-Patient-Beziehung. So wurde der Frage nachgegangen, welchen Einfluss Redezeiten haben, welche Wirkung Gesprächspausen und Gesprächstempo haben und durch welchen Fragestil ein Patient zum Reden angeregt werden kann.

Im Folgenden soll die Praktikabilität und Validität des Instrumentes diskutiert werden, vor allem durch den Vergleich zu anderen Bewertungsinstrumenten. Der Einfluss verschiedener Faktoren wie Training und Berufserfahrung auf unterschiedliche Gesprächsleistung und die Möglichkeit, dies mit einem Instrument zu erfassen, wird ebenfalls dargestellt. Darüber hinaus wird der Einfluss von Rollenspielsituation und Geschlecht diskutiert.

**Die Grundfrage lautet, wie kann man Arzt-Patient-Kommunikation beurteilen?**

### 6.1 Vergleich mit den Ergebnissen anderer Bewertungsinstrumente

An dieser Stelle wird auf Ergebnisse anderer Bewertungsinstrumente eingegangen. Zum einen auf die Originalinstrumente BAS und AGBS, aber auch auf die übrigen Instrumente, die in der Einleitung vorgestellt wurden: Die Checkliste „Professionelles ärztliches Kommunikationsverhalten“(Wien), der „Structured communication adolescent guide“, sowie das Roter System.

Mit der mAGBS-Eltern, die ein weiterer Doktorand (Robert Braun) parallel zu der hier vorliegenden Version entwickelte, wurden 30 Videos ebenfalls aus dem Wahlfach „Kinder- und Jugendpsychiatrie“ zur Überbringung des psychischen Befundes an die Eltern (siehe hierzu 2.5.2 Das Wahlfach Kinder- und Jugendpsychiatrie) bewertet und anschließend ebenfalls eine Reliabilitätsanalyse durchgeführt. Es sollen Ähnlichkeiten und Unterschiede dargestellt werden.

Ein besonderer Schwerpunkt wird dabei auf die Gütekriterien Reliabilität und Validität

gelegt. Aber auch thematische und methodische Aspekte der jeweiligen Arbeiten sollen beleuchtet werden.

### **6.1.1 Einsatz von Videos zur Anwendungstestung von Instrumenten**

Bei allen Instrumenten wurden zur Anwendungstestung Videoaufnahmen von Gesprächen verwendet. Die Qualität der 32 Videos, die zur Reliabilitätstestung der mAGBS-Jugendliche benutzt wurden, ist als sehr gut anzusehen. Bei der Reliabilitätstestung des Originalinstrumentes AGBS wurde angegeben, dass ein Teil der Videos wegen eingeschränkter Tonqualität nur unvollständig bewertet werden konnte. Auch bei der BAS wurden Videos zur Anwendungstestung verwendet.

Genau wie in bei den mAGBS-Jugendlichen ist auch beim RIAS eine Anwendung direkt vom Video oder der Audioaufnahmen möglich und keine Erstellung eines Transkriptes notwendig.

Die Anzahl der Videos, die zur Reliabilitätstestung verwendet wurde, ist sehr unterschiedlich. Angefangen mit der Checkliste Wien, die aus administrativen Gründen nur an 2 Videos getestet wurde (Puchner-Matzner et al., 2006) bis hin zur Anwendungstestung der AGBS, bei der 58 Videos verwendet wurden (Wand et al., 2008). Bei der Anwendungstestung der englischen Version, der BAS, wurden 20 Videos verwendet (Miller et al., 1999), so dass die vorliegende Arbeit mit 32 Videos als ausreichend angesehen werden kann. Beim „Structured communication adolescent guide“ waren es 42 Videos, die für die Reliabilitätstestung genutzt wurden.

### **6.1.2 Raterhintergrund**

Wie bei Wand wurde Wert darauf gelegt, Rater mit verschiedenem Ausbildungshintergrund einzusetzen. So waren in der vorliegenden Arbeit zwei Medizinstudenten und eine dritte Raterin mit mehr Ausbildungshintergrund. Auch in den BAS waren es 3 Rater, die einen unterschiedlichen Ausbildungshintergrund mitbrachten. Zwei Ärzte aus unterschiedlichen Fachgebieten und eine Krankenschwester wurden als Rater eingesetzt. Die Autoren geben an, dass die Rater nur wenig Fachwissen im Bereich Kommunikation hatten. Jedoch war ein Rater Facharzt für Psychiatrie, ein anderer Facharzt für Onkologie. Beide beschäftigen sich also viel mit Kommunikation (Miller et al., 1999). Bei der Anwendungstestung der Checkliste „Wien“ wurde ein anderer Weg eingeschlagen, was die Anzahl der Beurteiler betrifft. Es wurden 14 Lehrbeauftragte der medizinischen Fakultät als Rater eingesetzt.

### **6.1.3 Raterschulung**

Die Raterschulung, die bei Wand mit zwei Tagen angegeben wurde, war in der vorliegenden Arbeit vom Aufwand vergleichbar. Über die Stundenanzahl wird bei Wand keine Auskunft gegeben (Wand et al., 2008).

Der Umfang der Raterschulung für die vorliegende Arbeit betrug 16 Stunden, wobei in diesem Zeitfenster auch das Instrument parallel weiterentwickelt wurde. Die reine Schulungszeit dürfte im Bereich von 10 Stunden liegen. In der Literatur finden sich sehr unterschiedliche Angaben für Dauer und Umfang der Raterschulungen.

Für die BAS wurden die Rater insgesamt drei Stunden mit dem Instrument vertraut gemacht, indem sie 3 Videos damit bewerteten (Miller et al., 1999).

Das RIAS ist was die Schulung betrifft, viel aufwendiger als die anderen Instrumente. Die Autoren geben an, dass die Grundregeln für die Kodierung von Ungeübten in einer Woche gelernt und hohe Reliabilität und Kodierungsgeschwindigkeit nach 6-8 Wochen Übung erreicht werden (Roter, D. & Larson, 2002). Dies erscheint gerade für den Routineeinsatz für die Prüfung kommunikativer Fertigkeiten an Universitäten sehr lange.

### **6.1.4 Ratingzeiten**

Die Ratingzeit pro Video lag mit durchschnittlich 12 Minuten 20 Sekunden in einem ähnlichen Bereich wie die bei Wand angegebene Zeit von 15-20 Minuten. Die Ratingzeit in der BAS pro Video betrug 5 Minuten. Die Videos hatten eine Länge von maximal 15 Minuten. Die Ratingzeit erscheint sehr kurz.

Beim RIAS ist genau wie die Schulungszeit der Rater die Ratingzeit pro Gespräch um einiges aufwendiger als bei anderen Instrumenten. Die Autoren geben zwar an, dass ein trainierter Rater in etwa die doppelte Gesprächszeit für die Anwendung benötigt (Roter, D. & Larson, 2002).

Ong et. al. jedoch, die mit der RIAS arbeiteten, um Gespräche im onkologischen Setting auszuwerten, gaben an, für 10-minütige Gespräche bis zu 50 Minuten – also die 5fache Zeit - mit der Kodierung beschäftigt gewesen zu sein (Ong et al., 1998).

Die Autoren der Checkliste „Wien“ erachteten es als wichtig, gerade für den Einsatz an einer „Massenuniversität“ ein Instrument zur Verfügung zu haben, das in einem kurzen Zeitrahmen von etwa 10 Minuten anwendbar ist (Puchner-Matzner et al., 2006).

### **6.1.5 Unterschiede zwischen den Ratern**

Schaut man sich verschieden Ergebnisse an - die Mittelwerte der einzelnen Items, die Durchschnittsbewertung über die einzelnen Videos, die Boxplotdarstellungen, sowie



die Trennschärfekorrelation - so wird deutlich, dass Rater 1 generell und konstant über alle Items und alle Videos eine Tendenz zur besseren Bewertung hat als Rater 2 und Rater 3.

Hier zeigt sich das Problem, dass auch bei dem Versuch, eine möglichst standardisierte Bedingung zu schaffen und die Noten klar zu umgrenzen, die Rater trotz allem persönliche Erfahrungen und Persönlichkeitsstrukturen mit einbringen. Keine Beobachtung bleibt letztendlich frei von Schlussfolgerungen und Interpretationen des Raters. Als positiv zu bewerten ist jedoch, dass alle Rater ähnliche Tendenzen in der Bewertung haben. Es war nie der Fall, dass ein Video von einem Rater als sehr schlecht von dem anderen als sehr gut wahrgenommen wird, so dass letztendlich über Abweichungen in der absoluten Notenverteilung hinweggesehen werden kann. Der erfahrenste Rater, Rater 3 schöpfte das Notenspektrum am stärksten aus, wie die Boxplotdarstellung zeigt. Hier zeigt sich, dass auch der persönliche Erfahrungshintergrund bei Ratern eine Rolle spielt. Rater 3 hat schon viel mehr Videos gesehen als die anderen beiden Rater, deswegen liegt bei ihm unter Umständen eine klarere Vorstellung über mögliche Noten vor.

Bei der Reliabilitätstestung eines anderen Instrumentes, der „Structured communication adolescent guide“ (siehe hierzu 2.23 Überblick und Gegenüberstellung gängiger Bewertungsinstrumente) zeigte sich, dass die Rater zwar grundsätzlich darin übereinstimmten, ob ein bestimmtes Verhalten gezeigt worden war oder nicht. Abstufungen, ob etwas gut oder sehr gut bzw. schlecht oder sehr schlecht war, wurden jedoch unterschiedlich vergeben (Blake et al., 2005). Dies zeigt sich auch in der vorliegenden Arbeit, bei Item 9 „Nonverbales Verhalten“ und Item 14 „Empathie“ schöpften die Rater nicht das volle Notenspektrum aus, sondern tendierten stark zu mittleren Bewertungen. Möglicherweise ist das Erkennen, ob empathisches Verhalten vorliegt oder nicht noch relativ einfach, aber die Einschätzung wie empathisch es ist, fällt dann schon schwerer. An dieser Stelle ist auch die Notwendigkeit einer Raterschulung zu betonen. Die Rater, gerade wenn sie ungeübt sind, müssen durch das Ansehen und Diskutieren von Videos geschult werden und vor allem auch darauf verwiesen werden, sich genau an das Manual zu halten. Wenn mit einer Skala ein bestimmte Interraterreliabilität erreicht wurde, heißt das nicht, dass das Instrument diesen Wert besitzt. Der Wert ist leicht vom zukünftigen Umgang und den jeweiligen Raterhintergründen abhängig (Wirtz & Caspar, 2002). Die Notwendigkeit des Aufwands der Raterschulung hängt jedoch auch von der Erfahrung der Rater ab. Es hat sich gezeigt, dass gerade ungeübtere Rater durch den Gebrauch von Hilfsinstrumenten wie Checklisten und Ratingskalen in ihrem Urteil reliabler werden – vorausgesetzt, sie halten sich genau an die definierten Kriteri-

en. Experten scheinen jedoch durch den Gebrauch von Checklisten eher eingeschränkt zu werden (Lang, McCord, Harvill & Anderson, 2004; Regehr, MacRae, Reznick & Szalay, 1998).

Dies zeigte sich auch bei der Entwicklung eines Instrumentes zur Beobachtung der Eltern–Kind-Interaktion ebenfalls an der Kinder- und Jugendpsychiatrie der Universität München. Ergebnisse der Reliabilitätstestung ergaben, dass es gerade der Experten-Gruppe, die sehr geübt im Umgang von Verhaltensbeobachtung war, schwer fiel, sich an die Vorgaben des Instrumentes zu halten (Lutz-Tittel, 1999).

In der vorliegenden Arbeit zeigt sich eine gute Korrelation zwischen „Globaler Bewertung des Arztes (Item 23) und der Durchschnittsnote über alle Items der mAGBS-Jugendliche (siehe hierzu 5.2.3 Korrelation zwischen „Globale Bewertung des Arztes“ (Item 23) mit der Durchschnittsnote (Mittelwert über alle Themenblöcke)). Dies unterstreicht die These, dass die Rater durch Schulung und häufigere Verwendung des Instrumentes ein intuitives Gefühl für unterschiedliche Gesprächsleistungen bekommen.

Auch die Tatsache, dass die vor allem die beiden weniger erfahrenen Rater, Rater 1 und Rater 2, im Gebrauch der mAGBS-Jugendliche schneller werden, zeigt, dass Training für einen effektiven Einsatz des Bewertungsinstrumentes sinnvoll ist.

## **6.2 Anwendbarkeit der modifizierten Aufklärungsgesprächsbewertungsskala Jugendliche**

Insgesamt schätzt die Autorin die mAGBS-Jugendliche als ein praktikables und in angemessener Zeit anzuwendendes Instrument ein. Es war gut möglich, die ausgewählten Videoaufnahmen mit der mAGBS-Jugendliche zu bewerten. Die einzelnen Items waren im Manual zum größten Teil klar beschrieben. Vereinzelt kamen jedoch Unklarheiten auf und es war nicht immer eindeutig festzumachen, wie die auf den unterschiedlichen Videos gezeigte Leistung der Gesprächsführenden zu bewerten war. Diese Schwierigkeit schien aber mehr von der gezeigten Leistung auf den Videos abhängig als von unklar formulierten Items. Dies wurde auch von den anderen beiden Ratern bestätigt. Da diese beiden Rater kaum in die Entwicklung der mAGBS-Jugendliche involviert waren, war vor allem ihr Urteil zur Verständlichkeit von Bedeutung. Beide Rater gaben eine positive Rückmeldung zur Anwendbarkeit der mAGBS-Jugendliche an.

Die mAGBS-Jugendliche konnte auf drei verschiedene Situationen angewandt werden und war dort problemlos übertragbar.

Im Themenblock A „Gesprächsbeginn“ waren alle Items sehr einfach zu bewerten, da

es Fakten sind, ob ein Arzt den Patienten begrüßt oder ihm einen Sitzplatz anbietet.

Lediglich bei Item 3 „Hat sich der Arzt zu Beginn des Gespräches nach dem momentanen Befinden des Jugendlichen erkundigt“ zeigte sich eine Schwierigkeit in der Definition des Wortes „zu Beginn“. Viele Gesprächsführende fragen im Laufe des Gesprächs nach dem Befinden und es ist dann schwer festzumachen, ob sie damit das Kriterium für eine sehr gute Bewertung noch erfüllen oder nicht.

Auch im Themenblock B „Information bekommen“ waren die Items im Wesentlichen gut zu bewerten. Denn auch in diesem Block werden Gesprächsfertigkeiten abgefragt, die gut zu operationalisieren sind. Teilweise waren jedoch einzelne Wortdefinitionen zu ungenau, so erschien es bei Item 7 „Hat der Arzt angemessene Pausen zwischen den Fragen gelassen?“ schwierig einzuschätzen, ab wann eine Pause als angemessen anzusehen ist.

Im Themenblock C „Bedenken des Jugendlichen herausfinden“ bereitete Item 9 „Hat der Arzt auch nonverbale Äußerungen mit einbezogen?“ Schwierigkeiten.

In den wenigsten Situationen wird die Körpersprache direkt verbalisiert im Sinne von „du siehst so traurig aus“. Es ist schwer klar zu definieren, ab wann nicht verbalisierte Körpersprache berücksichtigt wird. Dies spiegelt sich auch in den Ergebnissen wieder. Es zeigt sich eine starke Tendenz zur mittleren Bewertung (Note 3 = „der Arzt hat trotz fehlender direkter Nachfrage nonverbale Äußerungen des Jugendlichen berücksichtigt“).

Im Themenblock D „Perspektive vermitteln“ waren gerade Items wie Item 11 „Struktur“ Item 12 „Wiederholungen und Zusammenfassungen“ sowie Item 15 „Fachlich korrekt“ einfach und klar zu identifizieren. Hingegen waren Item 13 „Positive Aspekte“ und Item 14 „Empathie“ komplexer in der Beurteilung. Vor allem die Empathie ist schwer zu beurteilen. Dies liegt aber eher in der Natur des Begriffes „Empathie“ und an der Situation der Rollenspiele. Da die Situation kaum emotionsgeladen war, fielen selten gleich mehrere empathische Äußerungen, die gefordert waren, um eine sehr gute Bewertung zu bekommen. Auch bei Darstellern, die sehr gut waren gab, es in der Bewertung eine Tendenz zur Mitte. Insgesamt müssen die Rater sehr genau hinschauen, um dieses Item zu bewerten und auch versteckte Äußerungen, Haltung und freundliches Interesse dem Jugendlichen gegenüber in die Bewertung mit einfließen zu lassen.

Auch die Ergebnisse der Rangreihe zeigen, dass Items aus Themenblock D durchschnittlich schlechter bewertet wurden als andere Items (siehe hierzu 5.1.6 Zusammenfassung: Rangreihe).

Im Themenblock E- Allgemeine Überlegungen waren die Items auch größtenteils gut

zu bewerten.

Item 16 und 17 waren wegen der abgefragten „Schlüsselbereiche“ sehr eindeutig. Fragte ein Gesprächsführender die jeweiligen Lebensbereiche ab, so konnte wie bei einer Checkliste für jedes Stichwort ein Punkt vergeben werden.

Auch Item 19 „Körpersprache“, Item 21 „Gesprächstempo“ und Item 22 „Beenden des Gespräches“ bereiteten keine Schwierigkeiten. Bei zwei Items war die Bewertung etwas schwieriger. Dies waren zum einen Item 18: „Hat der Arzt es geschafft dem Jugendlichen während des Gespräches hilfreich zur Seite zu stehen?“

Dieses Item lässt nach Ansicht der Autorin durch die komplizierte Formulierung der Unterpunkte Unklarheiten aufkommen. Man lässt sich als Rater unter Umständen dazu verleiten eher nach Gefühl zu bewerten, als die Kriterien der Unterpunkte im Manual bei der Benotung mit einzubeziehen.

Und zum anderen Item 20: „Hat es der Arzt während des Gespräches vermieden, unbeholfen zu sein?“

Auch hier sind die einzelnen Unterpunkte zwar sehr ausführlich dargestellt, was aber nicht zu einer klaren Bewertung beiträgt.

**Zusammengefasst kann man sagen, dass sich Probleme in der Bewertung vor allem dann ergaben, wenn das zu bewertende Verhalten auf den Videos nicht gezeigt wurde. Andere Probleme ergaben sich bei unklar formulierten Items beziehungsweise deren Unterpunkten, sowie den Items, die sich auf Emotionen bzw. Körpersprache bezogen.**

### **6.3 Geschlechtsspezifische Unterschiede in der Bewertung der Darsteller**

In der Gesamtwertung und in den einzelnen Items, die sich auf eine Bewertung des Ansprechens psychosozialer Aspekte beziehen (Item 8 „Nachfragen bei Problemen“, Item 9 „Nonverbales Verhalten“, Item 14 „Empathie“, Item 18 „Hilfreich zur Seite stehen“, Item 19 „Angemessene Körpersprache“), lassen sich Tendenzen erkennen, dass die Frauen von allen 3 Ratern besser bewertet werden. Im formalen Gesprächsverhalten zeigen sich keine Unterschiede. So zeigte sich der Geschlechtseinfluss bezüglich Gesprächsanteile und Fragestil nicht als signifikant.

Es können in der vorliegenden Arbeit aufgrund der geringen Fallzahl und der ungleichen Verteilung – es waren nur 9 Männer in der Arztrolle verglichen mit 23 Frauen - keine endgültigen Aussagen getroffen werden. Jedoch muss der Geschlechtseinfluss in der Kommunikation – so auch in der Arzt–Patient Kommunikation - als Einflussfaktor

im Auge behalten werden.

Konsequenzen für das Wahlfach Kinder- und Jugendpsychiatrie wären beispielsweise, dass Männer verstärkt in der Arztrolle eingesetzt werden, um Übungsmöglichkeit psychosozialer Fertigkeiten zu bekommen. Wie eingangs jedoch schon erwähnt, ist es gerade durch den wachsenden Frauenanteil in der Medizin schwierig, den Männeranteil zu erhöhen.

Man darf bei Bewertungen von Gesprächen auch nicht den Geschlechtseinfluss der Rater unterschätzen. Bei der vorliegenden Arbeit war es mit einem männlichen, sowie zwei weiblichen Ratern eine gemischte Gruppe. In vergleichbaren Arbeiten finden sich teilweise rein weibliche oder männliche Ratergruppen. Dies könnte durchaus einen gewissen Verzerrungseffekt mit sich bringen (Blake et al., 2005).

## **6.4 Einfluss von Rollenspielen**

Die Notenverteilung bei Item 24 „Darstellung der Patientenrolle“ zeigt überwiegend sehr gute bis gute Bewertungen. Und auch die Rollenidentifikation der Schauspieler wird in den meisten Fällen als überzeugend wahrgenommen. Selten fallen die Rollenspieler durch Lachen und Gesprächspausen länger aus der Rolle und keines der bewerteten Videos musste wegen geringer Rollenidentifikation früher abgebrochen werden. Dies ist auch eine Erfahrung der Dozenten des Wahlfaches, dass die Rollen sehr überzeugend dargestellt werden, obwohl es die Studenten vor allem zu Beginn große Überwindung kostet, eine Rolle zu übernehmen und sie sich selbst sehr kritisch sehen (Frank et al., 2009).

Immer wieder wird auch in der Literatur berichtet, wie authentisch Rollenspiele sein können, wenn sich die Darsteller darauf einlassen (Buckman, 1992).

Gerade auch durch das Sich Hineinversetzen in die Rolle eines Jugendlichen wird den Studenten ein Zugang zu den eigenen Emotionen ermöglicht. Dadurch verbessern sie ihr Beziehungsverhalten insgesamt (Benbassat & Baupal, 2002).

So bemerken auch die Autoren der Checkliste, dass es zwar sehr schwierig ist, Beziehungsvariablen zu erfassen und zu lehren, Rollenspiele dabei aber eine große Chance bieten: „Studierende übernehmen nicht nur die Rolle des Arztes, sondern auch des Patienten – und trainieren so ihr Einfühlungsvermögen und ihr Beziehungsverhalten insgesamt. Untrainierte Studierende neigen dazu, die somatischen Aspekte der Anamnese zu fokussieren und verlieren damit wesentliche Möglichkeiten des Beobachtens und Wahrnehmens.“ (Puchner-Matzner et al., 2006)

Rollenspiele werden oft gerade aus studentischer Sicht sehr kritisch gesehen, wenn

aber einige Regeln wie klare inhaltliche Angaben und Vorgaben beachtet werden und die Studenten genügend Zeit bekommen, sich auf die Rolle vorzubereiten, können sie aber auch mit gutem Erfolg eingesetzt werden (Nestel et al. 2007).

Darin liegt neben praktischen Überlegungen – aus organisatorischen Gründen ist beispielsweise der Einsatz von Patienten oder Simulationspatienten oft sehr aufwendig, während Studenten jederzeit eine Rolle übernehmen können (Frank et al., 2009) - auch ein ganz klarer Vorteil der Rollenspielsituation. Die Studenten erhalten außer den groben thematischen Vorgaben kein Skript wie sie die Rolle spielen sollen, sondern durch Ausprobieren und Spielen nach Empfinden ergibt sich eine sehr persönliche aber auch realistische Situation (Nestel et al 2007). Abgesehen von den thematischen Vorgaben gab es im Wahlfach und Kompaktkurs keine genaue Anweisung, wie sich der Jugendliche zu verhalten hat (Frank et al., 2009).

Eine Überlegung wäre Simulationspatienten einzusetzen, wie es auch im Rahmen der Kinder- und Jugendpsychiatrie beispielsweise in Heidelberg bereits gemacht wird und die mAGBS-Jugendliche an standardisierten Interviews zu testen. Aus verschiedenen Universitäten auch aus Deutschland wird vom guten Erfolg des Einsatzes von Kindern und Jugendlichen als standardisierte Patienten berichtet (Brown et al., 2005; Hardoff & Schonmann, 2001; Koch & Resch, 2009).

Der Einsatz standardisierter Patienten ist vor allem von Vorteil, wenn Studenten definierte Krankheitsbilder lernen sollen. Rollenspiele können jedoch mehr dazu beitragen, emotionale Kompetenz zu trainieren und mit unscharf definierten Situationen zurecht zu kommen (Kopecky-Wenzel et al., 2009).

Doch dadurch ergibt sich auch ein Problem bei der Anwendungstestung der mAGBS-Jugendliche. Die Bedingungen waren ungleich, so war jedes Gespräch anders, manche schwieriger, manchen einfacher, je nachdem wie sich der Darsteller in der Rolle des Jugendlichen zeigte. Dies zeigt auch die Korrelation zwischen der Bewertung der Arztrolle und der Jugendlichenrolle. Es scheint einen zwar geringen, jedoch vorhandenen Zusammenhang zwischen der Benotung beider Rollen zu geben (siehe hierzu 5.1.4 Die Rollenidentifikation als Patient). Denkbar wäre, wenn der Darsteller des Jugendlichen seine Rolle schlecht oder unglaubwürdig spielt, dies auch Einfluss auf die Qualität des Gespräches insgesamt hat.

Unvorhersehbare Gesprächssituationen spiegeln jedoch die Realität wider, weil auch Ärzte und Patienten höchst unterschiedlich in ihren Bedürfnissen sind.

Aus den OSCE-Prüfungen z.B. bei der Prüfung von Anamnesefertigkeiten wurde der Begriff der „Intercasereliabilität“ geprägt. Darunter versteht man das Problem in struktu-

rierten Prüfungen, dass trotz möglichst objektiver Bedingungen, letztendlich doch jedes Gespräch bedingt durch das Aufeinandertreffen verschiedener Persönlichkeiten anders verläuft (Chenot & Ehrhardt, 2003).

## 6.5 Gütekriterien des Bewertungsinstrumentes

### 6.5.1 Reliabilität

#### 6.5.1.1 Varianzanalyse

Die interne Konsistenz, Cronbachs Alpha lag in den BAS bei 0,93. Die interne Konsistenz der mAGBS-Jugendliche und der mAGBS-Eltern ist mit 0,85 bzw. 0,84 gut und mit der des Originalinstrumentes 0,86 nahezu identisch.

Tabelle 21 Cronbachs Alpha

Verschiedenen Versionen der Aufklärungsgesprächsbewertungsskala	Cronbachs $\alpha$
BAS (Miller et al. 1999)	0,93
AGBS (Wand et. al 2008)	0,88
mAGBS-Jugendliche (Englmeier)	0,85
mAGBS-Eltern (Braun)	0,84

Eine weitere Varianzanalyse ist bei der Reliabilitätsanalyse nötig, da ein Teil der Übereinstimmung zwischen Ratern auch durch Zufall erklärt werden kann. Methodenkritisch ist beispielweise bei der Checkliste „Wien“ anzumerken, dass keine Varianzanalyse gemacht wurde (Puchner-Matzner et al., 2006). In der Gesamtbewertung der mAGBS-Jugendliche liegt die Varianz, die auf Kosten der Inhomogenität der Ratermittelwerte geht, mit 6 Prozent etwas höher als im Originalinstrument. Noch höher liegt sie aber bei der mAGBS-Eltern, wo sie mit 34 % ein Drittel der Varianz ausmacht. Die Interaktion zwischen Rater und Arzt, also der unsystematische Fehler in der Gesamtbewertung beträgt 22 %. Dies ist ebenfalls höher als in der AGBS. In der BAS war die Inhomogenität mit 17 % für die Gesamtbewertung und die Interaktion zwischen Rater und Proband mit 21% auch höher.

Die geringste Inhomogenität der Ratermittelwerte für Rater und Arzt wurde sowohl für die mAGBS-Jugendliche als auch die mAGBS-Eltern im Block A – Gesprächsbeginn - erreicht.

Die höchste Inhomogenität der Mittelwerte wurde bei der mAGBS-Jugendliche und dem Originalinstrument für den Themenblock E – Allgemeine Überlegungen - mit 12 bzw. 7 % erreicht. Bei der mAGBS-Eltern war der Themenblock B mit 37 % am inhomogensten.

Insgesamt zeigte sich bei den beiden modifizierten Versionen im Vergleich mit der AGBS eine größere Inhomogenität der Ratermittelwerte, während die Interaktion zwischen Rater und Proband vor allem im Themenblock D und E geringer ist.

Die folgende Tabelle gibt einen Überblick über die verschiedenen Versionen.

Tabelle 22 Varianzanalyse mAGBS-Jugendliche, mAGBS-Eltern, AGBS und BAS

	<b>Leistung des Arztes (erklärte Varianz)</b>	<b>Inhomogenität der Ratermittelwerte</b>	<b>Interaktion Rater/Arzt</b>
Miller et al. BAS			
<b>Gesamt</b>	62%	17%	21%
WAND et al. AGBS			
<b>Gesamt</b>	78%	3%	19%
<b>Block A -</b>	88%	1%	11%
<b>Block B</b>	72%	1%	26%
<b>Block C</b>	57%	0%	43%
<b>Block D</b>	55%	4%	41%
<b>Block E</b>	66%	7%	27%
mAGBS-Jugendliche (Englmeier)			
<b>Gesamt</b>	72%	6%	22%
<b>Block A</b>	85%	1%	14%
<b>Block B</b>	65%	3%	32%
<b>Block C</b>	57%	10%	33%
<b>Block D</b>	66%	8%	26%
<b>Block E</b>	71%	12%	15%
mAGBS-Eltern (Braun)			
<b>Gesamt</b>	50%	34%	16%
<b>Block A</b>	66%	5%	29%
<b>Block B</b>	37%	37%	26%
<b>Block C</b>	50%	22%	28%
<b>Block D</b>	44%	36%	20%
<b>Block E</b>	73%	12%	15%

### 6.5.1.2 Intraclasscorrelationskoeffizienten

Bei beiden modifizierten Versionen der mAGBS-Jugendliche und der mAGBS-Eltern zeigen sich im Vergleich der ICC–Werte genau wie in den AGBS Unterschiede zwischen den Themenblöcken. Die Werte in der Gesamtwertung sind nahezu identisch.

Bei beiden modifizierten Versionen zeigen sich für Themenblock E etwas bessere



Übereinstimmungen zwischen den Ratern. Es fällt auf, dass bei den mAGBS-Eltern der Themenblock A mit 0,78 einen schlechteren ICC erzielt als die mAGBS-Jugendliche bzw. im Original (AGBS).

Tabelle 23 ICC-Werte im Vergleich

<b>Vergleich der Intraclasskorrelation</b>			
<b>Themenblock</b>	<b>AGBS (Wand 2008)</b>	<b>mAGBS-Jugendliche</b>	<b>mAGBS-Eltern</b>
<b>Gesamtbewertung</b>	0,86	0,82	0,84
<b>A Gesprächsbeginn</b>	0,94	0,92	0,78
<b>B Information bekommen</b>	0,81	0,78	0,65
<b>C Bedenken herausfinden</b>	0,63	0,65	0,71
<b>D Perspektive vermitteln</b>	0,61	0,77	0,76
<b>E Allgemeine Überlegungen</b>	0,76	0,81	0,90

Die etwas abweichenden Werte bei den ICCs für die einzelnen Themenblöcke werden von Wand so erklärt: „Eine mögliche Erklärung für diese Unterschiede ist, dass sich die in den benannten Themenblöcke A und B benannten Aspekte des Gesprächsverhaltens (z.B. den Patienten mit Namen ansprechen, den Patienten in verständlicher Sprache informieren) vergleichsweise einfach operationalisieren und identifizieren lassen“ (Wand et al., 2008). So erzielten auch der Themenblock A, B und E in den mAGBS-Jugendliche gute Werte. Die mittleren Themenblöcke C und D, vor allem auch das Eingehen des Arztes auf Sorgen und Emotionen des Patienten, zeigen geringere Übereinstimmung zwischen den Ratern. Bei der Anwendungstestung der Checkliste „Wien“ wurde ebenfalls eine Raterübereinstimmung berechnet und in Prozent angegeben.

Tabelle 24 Interraterübereinstimmung der Checkliste Wien

<b>Themenbereiche</b>	<b>1.Video</b>	<b>2.Video</b>
<b>Gesamtwertung</b>	74%	75%
<b>Bereich A Gesprächsinhalt</b>	70%	70%
<b>Bereich B-Gesprächsstil</b>	85%	78%
<b>Bereich C-Beziehung</b>	79%	78%

Die Werte sind in einem ähnlich guten Bereich wie die Werte der mAGBS-Jugendliche. Es gibt zwar für keinen Gesprächsteil eine 100-prozentige Übereinstimmung, aber die Übereinstimmungen sind hoch. Die Werte sind besonders gut für den Bereich B „Gesprächsstil“, dieser Bereich entspricht inhaltlich in der mAGBS-Jugendliche dem The-

menblock E – Allgemeine Überlegungen. Dieser erzielte neben dem Themenblock A – Gesprächsbeginn - die besten Übereinstimmungen.

Auf Einzelitemebene erzielten die Items, die formale Gesprächskriterien abfragen, die höchste Übereinstimmung. Items, die psychosoziale Faktoren erfassen, wurden uneinheitlicher bewertet. Dieses Ergebnis zeigt sich ebenfalls in der mAGBS-Jugendliche. Darüber hinaus erzielte auch ein Item zum Thema „Der Arzt kann Momente des Schweigens zulassen“ schlechte Übereinstimmung zwischen den Ratern. Dies deckt sich mit Ergebnissen der mAGBS-Jugendlichen und der mAGBS-Eltern, bei denen Item 7 „Pausen“ und Item 21 „Gesprächstempo“ zu den am uneinheitlichsten bewerteten Items zählen.

**Tabelle 25 Vergleich zwischen ICC von Item „Pause“ und „Gesprächstempo“ bei der mAGBS-Jugendliche und der mAGBS-Eltern**

	mAGBS-Jugendliche	mAGBS-Eltern
Item „Pausen“	0,54	0,39
Item „Gesprächstempo“	0,50	0,36

Eine mögliche Erklärung hierfür könnte sein, dass in diesen Items relativ komplexe Fragestellungen sehr verkürzt abgefragt werden, die auf den ersten Blick trivialen Fragen: „Macht der Arzt angemessene Gesprächspausen“, „Ist das Gesprächstempo angemessen“, scheinen in der Beurteilung doch nicht so einfach, wie zuerst vermutet. Die schlechte Übereinstimmung im Bereich der Gesprächspausen wird von den Autoren der Checkliste „Wien“ damit erklärt, dass dies im studentischen Gespräch oft vergessen wird und die Studenten im Gespräch hastig und dominant vorgehen.

Bei der mAGBS-Eltern kommt hinzu, dass es sich bei den zu bewertenden Gesprächssituationen um eine Mehrpersonensituation handelt. Der Arzt trifft im Gespräch auf Mutter und Vater, also insgesamt drei Personen, die zu Wort kommen wollen. Dies könnte einen Einfluss auf das Tempo und die Pausen im Gespräch haben.

Für das RIAS bewegt sich die Interraterreliabilität durchschnittlich über alle Kategorien in vielen Studien in einem guten Bereich. Roter gibt an, dass das Instrument im Durchschnitt eine Interraterreliabilität von 0,85 erreichte, allerdings basierend auf dem Pearson Korrelationskoeffizienten, was eine direkte Vergleichbarkeit mit Werten der mAGBS-Jugendlich erschwert (Roter, D. & Larson, 2002). Jedoch kann man sagen, dass 0,85 einen ähnlich guten Zusammenhang zwischen verschiedenen Rater angibt, wie ein ICC von 0,82 für die Gesamtwertung der mAGBS-Jugendliche.

Bei der Reliabilitätstestung der modifizierten Version der Calgary Cambridge Observation Guides, der „Structured communication adolescent guide“, lag der ICC für die Gesamtwertung mit 0,93 etwas höher als bei den mAGBS–Jugendlichen. Jedoch gibt es wie auch in der mAGBS-Jugendliche in den einzelnen Themenblöcken Unterschiede in der Übereinstimmung zwischen den Ratern. So erzielte der Themenblock „Beginn des Gespräches“ einen ICC von 0,87, während die Themenblöcke „Wie wurde die Information gesammelt“ und „Sammeln von Informationen“ mit ICC–Werten von 0,24 bzw. 0,50 niedriger liegen.

Die Autoren erklären die niedrigere Übereinstimmung wie Wand damit, dass einzelne Items der Themenblöcke schwieriger zu operationalisieren und identifizieren sind. So ist z.B. „Ausreichender Einsatz von Körpersprache“ schwieriger zu identifizieren als das Item, das nach der Begrüßung fragt. Schwierig scheint es für die Rater ebenfalls bei qualitativen Abstufungen von Gesprächsleistung zu werden. So schien es kein Problem zu sein, zwischen gezeigten oder nicht gezeigtem Verhalten zu unterscheiden, doch eine Qualitätsaussage wie gut das gezeigte Verhalten der Gesprächsführenden war, führte zu abweichenden Angaben unter den Ratern (Blake et al., 2005).

Zusammengefasst lässt sich hierzu sagen, dass verschiedene Bewertungsinstrumente hier zu einem sehr ähnlichen Ergebnis kommen:

**Klar zu definierende Verhaltensweisen des Gesprächsführenden lassen sich gut einheitlich bewerten, uneinheitlicher wird es bei Körpersprache, Emotionen und Gesprächstempo.**

### 6.5.2 Validität

#### 6.5.2.1 Inhaltsvalidität

Die Validität ist das Kriterium, ob ein Instrument misst was es messen soll. Im vorliegenden Fall geht es darum, eine Antwort zu geben, inwiefern die mAGBS-Jugendliche wirkliche kommunikative Fertigkeiten misst. Zu bedenken ist an dieser Stelle, dass es keinen Goldstandard gibt. Die Reliabilität als Grundvoraussetzung der Validität ist für die mAGBS-Jugendliche hinreichend gegeben. Da ein Teil der Items vom Originalinstrument übernommen wurde und Änderungen nur nach ausführlicher Auseinandersetzung mit der Literatur durchgeführt wurden, wird die Inhaltsvalidität als relativ hoch angesehen. Die Autoren des MAAS-R-Manual von 2000 schreiben, dass die Inhaltsvalidität der MAAS-R durch intensive Auseinandersetzung mit Literatur und Expertenmeinung abgesichert wurde. So decken die Items insgesamt 85 % der wichtigen kommunikativen Fertigkeiten ab (van Thiel et al., 2000).

Eine Studie aus England verglich die Meinung von Allgemeinärzten und Patienten im Bezug auf ihre Erwartungen an ein gutes Gespräch und fand heraus das die MAAS-R inhaltlich sowohl Patienten- als auch Ärztesichtweise gut abdeckt. Gerade die Items im Instrument, die Fragetechnik von Ärzten oder die Art der Informationsvermittlung beurteilen, wurden von beiden Gruppen als relevante Kriterien beurteilt (Robinson, Walley, Pearson, Taylor & Barton, 2002).

Die mAGBS-Jugendliche hat zahlreiche Items, die auf ähnliche Fertigkeiten wie der MAAS-R abzielen, was wiederum die Inhaltsvalidität der mAGBS-Jugendliche sichert.

Die einzelnen Items decken inhaltlich die Themen ab, die in einem Arzt-Jugendlichen-Gespräch angesprochen werden sollten. Ein sehr ähnliches Konzept findet sich bei der Entwicklung der „Structured communication adolescent guides“, ein Instrument, das im Theorieteil näher beschrieben wird (Vergleiche hierzu 2.3.2.3 Die Calgary-Cambridge-Observation Guides). Die Calgary Cambridge Observation Guides wurden im Wesentlichen in ihrer Struktur beibehalten, es wurden jedoch unter Berücksichtigung des Themas inhaltliche Änderungen vorgenommen. Dieser Arbeitsschritt zeigt große Parallelen zur Entwicklung der mAGBS-Jugendliche. Eine Hauptkomponente der SCAG ist ebenfalls das Eingehen auf ein mögliches Risikoverhalten des Jugendlichen. Das Item „Lifestyle“ der SCAG, das das Fragen des Arztes nach Freunden, Beziehung, Familie, Drogen, Alkohol und Rauchen, Depressionen bewertet, hat große Überschneidungen zu Item 16 „Schlüsselbereiche“ und Item 17 „Tiefgehendes Besprechen“(Blake et al., 2005). Viele Inhalte, die das Instrument fokussiert, finden sich auch in der mAGBS-Jugendliche, was wiederum dessen Inhaltsvalidität stützt.

Es hat sich gerade im Umgang mit Checklisten bei der Bewertung von Fertigkeiten immer wieder das Problem gezeigt, dass durch ein globales Urteil die Leistung reliabler und valider eingeschätzt wurde als durch einen Gesamtwert über mehrere Kategorien einer Checkliste (Cohen, Colliver, Marcy, Fried & Swartz, 1996; Regehr et al., 1998). Dies wirft die Frage auf, ob Checklisten unter Umständen im Widerspruch zu einer Entscheidung nach „Bauchgefühl“ stehen.

In der vorliegenden Arbeit gibt es jedoch einen guten Zusammenhang zwischen dem globalen Urteil und dem Gesamtwert über alle 5 Themenblöcke der mAGBS-Jugendliche.

Jedoch ist dieser Zusammenhang zwischen globalem Urteil und Gesamtwert auch ein Hinweis auf Fertigkeiten der Rater. Durch die Schulung und die Verwendung des Instrumentes ist es ihnen möglich, ein „Bauchgefühl“ begründbar zu machen.

Genau das wäre auch das Ziel des Einsatzes eines Bewertungsinstrumentes für Prü-

fungssituationen. Gerade Lernenden genügt es nicht, wenn man ihnen sagt, ob eine Leistung gut oder schlecht war, sondern es sollten einzelne Gesprächselemente Schritt für Schritt bewertet werden. Hier kann ein Instrument eine gute Unterstützung bieten.

Bei der Validitätsprüfung der SCAG setzte die Autorin als „Goldstandard“ einen Psychologen ein, der jedes Video, das mittels der SCAG durch die Rater überprüft wurde, zusätzlich global bewertete. Eine Korrelation zwischen dem Gesamtwert der SCAG und dem globalen Urteil zeigte auch hier eine gute Übereinstimmung (Blake et al., 2005).

### **6.5.2.2 Konstruktvalidität**

Es muss möglich sein, dass verschiedene Leistungen und Ausbildungshintergründe der Gesprächsführenden auch verschieden bewertet werden und ein Instrument diesen Unterschied erfasst. Darauf wurde auch bei anderen Bewertungsinstrumenten in der Anwendungstestung geachtet. Bezüglich der Konstruktvalidität ist zu sagen, dass in mehreren Studien nachgewiesen wurde, dass die MAAS-R gut zwischen verschiedenen Ausbildungshintergründen unterscheidet (van Thiel et al., 2000).

So zeigten Medizinstudenten eine signifikante Verbesserung im Punktwert des MAAS-R über den Verlauf eines Kurses im Vorher-Nachher-Vergleich (Van Dalen et al., 1998).

In der vorliegenden Arbeit wurde durch die Testung an Gruppen mit unterschiedlicher Berufserfahrung und unterschiedlichem Training eine Konstruktvalidität belegt. Insgesamt zeigt sich über alle Videos, dass ein breit gefächertes Notenspektrum vorkommt. Im Vergleich zwischen den Gruppen zeigt sich, dass es tatsächlich einen Unterschied zwischen den Videogruppen gibt. Dieser Unterschied zeigt sich jedoch nicht darin, dass eine Gruppe konstant besser ist, sondern je nach Item jeweils eine Gruppe den anderen etwas überlegen ist in der Benotung.

Die Effektstärke als Maß für die Bedeutsamkeit eines Unterschiedes bewegt sich im Vergleich zwischen der Jugendlichen- und der Notfallgruppe mit der überwiegenden Anzahl von Werten zwischen 0,18 und 0,54 im mittleren Bereich.

Auch in der Messung mit Interact zeigten sich hinsichtlich des Gesprächsanteils des Arztes und der Verwendung des Fragestils Unterschiede zwischen den drei Gruppen.

Jedoch müsste zur Prüfung eines Trainingseffekts dieselbe Gruppe vor und nach Training von verblindeten Ratern bewertet werden. Dies war beim vorliegenden Studienansatz nicht möglich, da anhand der Gesprächssituation erkennbar war, um welche Gruppe es sich handelt.

## 6.6 Übertragbarkeit des Themas

Für das Rating wurden Videos mit verschiedenen Ausbildungshintergründen der Gesprächsführenden gewählt. Ähnlich gingen auch die Autoren bei den BAS vor. Sie wählten für die Gesprächsführung verschieden ausgebildete Probanden - vom Medizinstudenten bis zum Oberarzt - aus, um so ein möglichst großes Spektrum an Kenntnissen zu haben und sicherzustellen, dass das Instrument einen weiten Raum zur Messung von Fähigkeiten abdeckt (Miller et al., 1999). Dies war auch in der vorliegenden Arbeit der Fall. Es konnte auf drei verschiedene Situationen mit dem Ausbildungshintergrund Student, Student mit 4-tägigem Training und Arzt mit Berufserfahrung zurückgegriffen werden.

Neu ist, dass gleichzeitig bei 11 Videos ein höherer Schwierigkeitsgrad des Themas vorlag.

Bei der mAGBS–Eltern wurden Videos zur Überbringung des psychischen Befundes eines Schulkindes mit Entwicklungsrückstand an die Eltern bewertet. Dies liegt inhaltlich näher am „Breaking bad news“ -Thema des Originalinstrumentes. Deswegen wurde im Vorfeld an der mAGBS- Eltern weniger geändert, es zeigten sich jedoch bei der Reliabilitätstestung bei manchen Items Probleme der Anwendbarkeit. Ein wesentlicher Unterschied zwischen den Videos der mAGBS-Jugendliche und Eltern und der BAS und dem Originalinstrument AGBS besteht in der Rollenspielsituation auf der einen und im Einsatz von Simulationspatienten mit klar vorgegeben Krankheitsbildern auf der anderen Seite (siehe hierzu 6.4 Einfluss von Rollenspiele).

Den Videos, die beim Rating der beiden modifizierten Versionen verwendet wurden, lag ein anderes Konzept zugrunde. Während bei der Evaluation des Originalinstrumentes der AGBS standardisierte, geschulte Patientendarsteller zum Einsatz kamen und es für die Darsteller der „Ärzte“ einen umschriebenen Plan gab, was angesprochen werden sollte, waren die Darsteller der Jugendlichen in dieser Arbeit Studenten ohne klare Vorgaben. Sowohl in der mAGBS–Jugendliche, als auch in der mAGBS–Eltern wurde tendenziell schlechter bewertet als im Originalinstrument, der AGBS. Die nachfolgende Tabelle gibt dazu einen Überblick.

Tabelle 26 Vergleich der Mittelwerte von mAGBS-Jugendliche, mAGBS-Eltern und AGBS

<b>Bewertung unterschiedlicher Gesprächssituationen</b>						
	<b>mAGBS-Jugendliche(2009) allgemeines Gespräch/Suizidversuch</b>		<b>mAGBS-Eltern (Braun 2009) Schulkind mit Entwicklungsrückstand</b>		<b>AGBS (Wand et al.2008) Onkologie</b>	
	MW	Std.-Abw	MW	Std. Abw.	MW	Std. Abw.
<b>Gesamtbewertung</b>	<b>2,26</b>	0,40	<b>2,37</b>	0,37	<b>2,06</b>	0,38
<b>A - Gesprächsbeginn</b>	<b>2,13</b>	0,60	<b>2,26</b>	0,52	<b>1,84</b>	0,86
<b>B - Befund überbringen/Information bekommen</b>	<b>1,91</b>	0,51	<b>2,61</b>	0,49	<b>2,65</b>	0,60
<b>C - Bedenken des Patienten</b>	<b>2,63</b>	0,57	<b>2,89</b>	0,52	<b>2,40</b>	0,35
<b>D – Informationsübermittlung/ Perspektive vermitteln</b>	<b>2,49</b>	0,53	<b>2,44</b>	0,57	<b>1,77</b>	0,51
<b>E - Allgemeine Überlegungen</b>	<b>2,12</b>	0,57	<b>1,69</b>	0,57	<b>1,81</b>	0,50

Die Autorin ist der Meinung, dass die Differenz in der Gesamtbewertung nicht als gravierend betrachtet werden kann und alle drei in der Tabelle dargestellten Instrumente in etwa auf demselben Bewertungsniveau liegen. Auch diese Ergebnisse stützen letztendlich die Validität des Instrumentes. Unterschiedliche Gesprächssituationen werden durch das Instrument unterschiedlich abgebildet.

## **6.7 Einfluss von Training und Berufserfahrung**

### **6.7.1 Trainingseffekt**

Es zeigt sich bei den Notfallvideos eine Tendenz zur besseren Bewertung bei gut operationalisierbaren Gesprächsfertigkeiten wie Item 2 „Begrüßung“, Item 3 „Befinden“, und Item 4 „Gesprächseinstieg“. Auch bei Item 8 „Nachfragen bei Problemen“ und Item 9 „Nonverbale Äußerungen“ und Item 14 „Empathie“ werden die Notfallvideos mittels der mAGBS-Jugendliche besser bewertet.

Die Gesprächsführenden der Notfallvideos waren nicht dieselben Studenten wie auf den Jugendlichenvideos, jedoch hatten die Studenten auf den Notfallvideos ein 23-stündiges Training absolviert, und auch wenn sie davor nicht in der Rolle eines Arztes waren, die Diskussion und Rückmeldung der Gesprächsleistung anderer Kursteilnehmer mit verfolgt. Die Verbesserung der Gruppe bei einigen Items insgesamt kann deshalb, auch wenn es nicht dieselben Personen waren, durch einen Lerneffekt erklärt

werden. Verglichen mit anderen Kommunikationstrainings ist das Wahlfach „Kinder- und Jugendpsychiatrie“ im Umfang als intensiv anzusehen. Die eingangs beschriebenen Studien zum Training berichten alle von mehrstündigen Trainingsinterventionen an 1-3 Tagen mit teilweise großem Effekt gegenüber nichttrainierten Kontrollgruppen (Fallowfield et al., 2002; Hoffmann et al., 2007; Langewitz, W. A. et al., 1998).

Als nichttrainierte Gruppe kann in der vorliegenden Arbeit die Gruppe der Allgemeinärzte gesehen werden, die zwar Berufserfahrung mitbringt, aber insgesamt nur 3 Stunden spezielles Gesprächstraining bekommen hatte.

Auffällig ist auch, dass in der Gruppe der Notfallvideos der Anteil an offenen Fragen bei 47% liegt im Vergleich zu 35 % bei den Jugendlichenvideos und 38% bei den Allgemeinärzten. Zu den Lernzielen des Wahlfaches Kinder- und Jugendpsychiatrie gehören explizit neben inhaltlichen Aspekten die Vermittlung von Basisfertigkeiten wie das Stellen offener Fragen.

Auch hier ist ein Trainingseffekt sichtbar, obwohl auch zu diskutieren ist, ob es sich aufgrund des Themas „Suizidversuch“ ergibt, dass mehr offene Fragen gestellt werden. Offen bleibt deshalb, warum die trainierten Studenten bei anderen Items schlechter abschnitten als die Studenten bei den Jugendlichenvideos zu Beginn des Trainings. So erzielte diese Gruppe bei Item 5 „Annäherung“, Item 6 „Sprache“, Item 10 „Äußerungen“, Item 11 „Struktur“ sowie Item 16 „Schlüsselbereiche“ eine bessere Bewertung. Möglicherweise ist dies zulasten des schwierigeren Themas bei den Notfallvideos.

Ein schon aus anderen Studien bekanntes Problem bei der Effektmessung kommunikativer Fertigkeiten taucht jedoch auch bei der vorliegenden Arbeit auf. Es zeigt sich das Dilemma passende Evaluationsinstrumente auszuwählen, die die geschulten Fähigkeiten wirklich abbilden können. Wichtig ist dabei aber gleichzeitig, dass das Instrument nicht zu sehr auf den gewünschten Lerneffekt zugeschnitten ist. Gerade methodisch aufwendige Studiendesigns mit Vorher-Nachher-Vergleich und Kontrollgruppen, die ohne Training ein gleiches Thema bearbeiten, bringen oft als Ergebnis einen geringeren Effekt eines Kommunikationstrainings hervor als der Aufwand des Trainings erwarten ließe. Es stellt sich hierbei dann die Frage, ob das Training keinen Effekt gebracht hat oder das Instrument den Effekt nicht aufzeigen kann, weil es nicht in der Form für das Thema passt (Cegala & Lenzmeier Broz, 2002). Hier fehlt letztendlich auch wieder der Goldstandard für kommunikative Fertigkeiten. Idealerweise müssen zur umfassenden Beurteilung einer Verbesserung kommunikativer Fertigkeiten verschiedene Evaluationskriterien kombiniert werden wie auch beispielsweise die Patientsichtweise (Heaven et al., 2003).



### 6.7.2 Berufserfahrung

Die Teilnehmer des Kompaktkurses „Pädiatrie“ – alles angehende Allgemeinärzte - hatten zum Teil mehrjährige Berufserfahrung.

Ein Vergleich zwischen den Ergebnissen der mAGBS–Jugendliche zeigt jedoch, dass die Allgemeinärztergruppe bei allen Items bis auf Ausnahmen schlechter abschneidet als die Studentengruppen. Eine bessere Bewertung als die Medizinstudenten bekamen die Allgemeinärzte bei den Items 12 „Wiederholungen“, Item 13 „Positive Aspekte“ und Item 19 „Körpersprache“. Die Effektstärke war am höchsten bei Item 13 „Positive Aspekte“ mit 0,76.

Ein Unterschied zwischen den Gruppen zeigt sich auch in der Auswertung der Gesprächsverläufe mit Interact. Der Redeanteil war mit durchschnittlich 63 % bei den Allgemeinärzten, bei einzelnen Videos sogar zwischen 70 und 80% sehr hoch. Auch der Anteil an geschlossenen Fragen in der Gesprächsführung war mit durchschnittlich 62 % relativ hoch.

Im Vergleich zwischen Jugendlichen– und Notfallvideos gibt es kaum einen Unterschied bezüglich Redeanteil und Fragestil. Der Unterschied zu den Allgemeinärzten ist jedoch deutlich. Als Erklärung für die unterschiedlichen Werte bei Gesprächsanteil, Fragestil und nicht unterbrochener Redebeiträge zwischen den Gruppen können verschiedene Einflussfaktoren genannt werden. Es gibt den Geschlechtseinfluss bei beiden Gesprächspartnern, darüber hinaus besteht ein thematischer Unterschied, und der unterschiedliche Ausbildungshintergrund darf auch nicht vergessen werden. Eine univariate Varianzanalyse für Gesprächsanteile des Arztes, zeigte, dass ein relativ hoher Prozentsatz der Varianz zwischen den Gruppen nicht erklärt werden kann. Das Geschlecht des Arztes und des Jugendlichen zeigt keinen Einfluss, von Bedeutung scheint jedoch zu sein, zu welcher Gruppe das Video gehört. Es kann jedoch im vorliegenden Studienansatz nicht ausreichend geklärt werden, wie stark jeweils der Einfluss des Trainings und der Einfluss des Themas ist.

Ein direkter Gesprächsstil kann sich in einem vermehrten Redeanteil zeigen. Bedenken muss man jedoch, dass es auch stark auf den Kontext und den jeweiligen Patienten ankommt, wie viel der Patient spricht (Dorf Müller, 2001). Bei einem Aufklärungsgespräch z.B. muss der Arzt gewisse Informationen geben, es wäre hier also ein höherer Redeanteil des Arztes zu erwarten. So ist auch im Fall der Jugendgesundheitsuntersuchung die Aufgabenstellung etwas unterschiedlich, es wird vom Arzt erwartet, dass er dem Jugendlichen Anregungen zu Suchtthemen und gesunder Lebensführung mitgibt. Eine rein quantitative Messung der Gesprächsanteile, sagt noch nicht alles über die Qualität eines Gespräches aus.

Es hat sich gezeigt, dass die kommunikativen Fertigkeiten durch Berufserfahrung nicht zwangsläufig besser werden, sondern oft schleicht sich gerade durch Routine und Zeitdruck ein sehr direkter Gesprächsstil ein (Dorf Müller, 2001). Ein gezieltes Training kann unter Umständen mehr zur Verbesserung kommunikativer Fertigkeiten beitragen als mehrjährige Berufserfahrung (Fallowfield et al., 2002). Ähnliche Ergebnisse zeigt eine Arbeit aus der Psychosomatik, in der es um die Verbesserung von Allgemeinärzten bezüglich psychosomatischer Kompetenz durch ein 12-stündiges Training ging. Ein Ergebnis war unter anderem, dass je berufserfahrener ein Arzt war, desto festgelegter und direkter seine Art der Gesprächsführung (gemessen mittels des Bewertungsinstrumentes Maastricht History taking and advice checklist) schien (Günter, 2004).

Vorgabe und Gesprächsziel bei allen Gesprächen, sei es Jugendlicher, Notfallsituation, oder Allgemeinärzten waren letztendlich mit dem Jugendlichen ins Gespräch zu kommen. Dies gelang den Gesprächsführenden des Wahlfaches vor allem bei den Notfallsituationen am Ende des Trainings am besten.

Die Studenten des Wahlfaches waren alle im fortgeschrittenen Semester, man kann also davon ausgehen, dass alle vor Besuch des Wahlfaches Kinder- und Jugendpsychiatrie einen ähnlichen Trainingshintergrund bezüglich kommunikativer Fertigkeiten hatten. Im Vergleich mit den Allgemeinärzten jedoch ist durch das Wahlfach ein gewisser Verzerrungseffekt gegeben.

Es besuchen vor allem die an psychosozialen Fragestellungen interessierten Studenten das Wahlfach Kinder- und Jugendpsychiatrie.

## **6.8 Gesprächsverläufe: Interaktions-Analyse-Systeme und Interact**

### **6.8.1 Vergleich zwischen Interaktions-Analyse-Systemen und globalen Bewertungsinstrumenten**

Bei einem Interaktion–Analyse–System ist es besser möglich, den Fluss eines Gespräches widerzuspiegeln als bei einem globalen Bewertungsinstrument (Langewitz, W., Nubling & Weber, 2003). So kann bei Interaktions-Analyse-Systemen die Abfolge des Gespräches nachvollzogen werden. Bei der RIAS liegt der Blick nicht nur auf den Äußerungen des Arztes, sondern auch den Äußerungen des Patienten (Ford et al., 2000). Das RIAS wird der dialogischen Natur eines Gespräches besser gerecht (Wasserman & Inui, 1983).

Dies ist ein Kritikpunkt an den mAGBS-Jugendlichen. Das Verhalten des Patienten fließt nur indirekt in einzelne Items mit ein. Für den Fall, dass der Patient kein Verhalten zeigt, auf das der Arzt reagieren kann, geht dies jedoch auf Kosten der Bewertung des Arztes.

Vernachlässigt wird bei einem globalen System, wie der Vergleich mit dem RIAS ebenfalls zeigt, eine detailliertere Einschätzung formaler Gesprächskriterien. Beim RIAS wird jede Äußerung einer Kategorie zugeteilt, so dass beispielsweise schon bei der Erfassung gezählt wird, wie viele offene und geschlossene Fragen im Gespräch vorkommen (Roter, D. & Larson, 2002).

Der Einsatz von Interact in der vorliegenden Arbeit kann als Versuch verstanden werden, die Interaktion im Gespräch sequentiell einzufangen. In der vorliegenden Arbeit wurden mithilfe der Software der Fragestil, Pausen, die Anzahl der Hörersignale und die Abfolge von Arzt- und Patientenäußerungen kodiert.

### **6.8.2 Fragestil und Redeanteil**

Eine häufig zitierte Studie zeigt, dass Ärzte ihre Patienten nach durchschnittlich nur 18 Sekunden unterbrechen bzw. sie oft auch gar nicht so lange zu Wort kommen lassen (Beckman & Frankel, 1984). Ein Lernziel des Wahlfaches Kinder– und Jugendpsychiatrie war es, Möglichkeiten zu finden, einen Jugendlichen zum Reden zu bringen und ihn möglichst lange nach seinem eigenen Bedürfnis zu Wort kommen zu lassen.

Langewitz et al. fanden mittels dem RIAS heraus, dass auf offene Fragen durchschnittlich 2,75 Äußerungen des Patienten folgen, auf geschlossene Fragen durchschnittlich 1,78 Äußerungen (Langewitz, W. et al., 2003).

Ähnliches konnte auch für die vorliegende Arbeit herausgefunden werden. Es konnte gezeigt werden, dass zwischen dem Stellen offener Fragen seitens des Arztes und der Anzahl zusammenhängender Äußerungen des Jugendlichen ein deutlicher Zusammenhang besteht. Ein Unterschied zwischen den Gruppen zeigt sich auch in der Auswertung der Gesprächsverläufe mit Interact. Der Redeanteil war mit durchschnittlich 63 % bei den Allgemeinärzten, bei einzelnen Videos sogar zwischen 70 und 80% sehr hoch. Auch der Anteil an geschlossenen Fragen in der Gesprächsführung war mit durchschnittlich 62 % relativ hoch.

Es finden sich in der Literatur nur vereinzelt Angaben zu Redeanteilen bei Arzt-Patient-Gesprächen. Köhle et al. beschreiben in einer Studie, die eine Gruppe von Ärzten vor und nach einer Fortbildung zur psychosomatischen Grundversorgung untersuchte (Vergleiche 2.1.2 Verbesserung des ärztlichen Kommunikationsverhaltens durch Training?) Gesprächsanteile zwischen 30 und 47%, wobei gerade die Ärzte mit höherem Gesprächsanteil zu Beginn im Trainingsverlauf signifikant ihren Redeanteil verminderten. Die Autoren sehen also eine verminderten Redeanteil des Arztes als Zeichen für einen patientenzentrierten Gesprächsstil und eine Verbesserung im Kommunikationsstil (Köhle et al., 2001).

Eine Studie aus der Pädiatrie zeigt, dass die Ärzte mit durchschnittlich 61% Redeanteil am meisten zum Gespräch beitragen, während die Kinder selbst nur 4 % zum Gespräch beitragen (Wassmer et al., 2004). In der Allgemeinarztgruppe sprechen die Jugendlichen durchschnittlich 37 %.

### **6.8.3 Die Pause im Gespräch**

Pausen bereiten gerade Anfängern in der Gesprächsführung Probleme. Eine Gesprächspause wird dabei oft gleichgesetzt mit einer Lücke im Gespräch und der Arzt neigt daher intuitiv dazu, das Entstehen einer Pause zu vermeiden und über eine möglicherweise entstehende Pause hinwegzureden. Dies sind wiederkehrende Erfahrungen aus dem Studentenunterricht (Kopecky-Wenzel et al., 2009; Puchner-Matzner et al., 2006).

Eine Frage, die bei der Auseinandersetzung mit der mAGBS-Jugendliche immer wieder aufkam, war, was überhaupt eine Pause ist und ab welcher Länge sie als solche aufgefasst wird.

Bei Roter wird der Pause Bedeutung als Grenze zwischen zwei Äußerungen zugestanden. Von Sandvik et al wird bei dieser Definition kritisiert, dass die Pause ein dynamisches Element eines Gespräches ist und viel mehr Bedeutung als nur den Übergang zwischen zwei Äußerungen hat. (Sandvik et al., 2002). Darüber hinaus gibt Roter

an, dass Pausen nur ab einer Länge von einer Sekunde als relevant kodiert werden sollen (Manual Roter 2002).

Geisler schlägt für Trainingssituationen vor, bewusst Pausen von 3 Sekunden Länge einzubauen. Gleichzeitig meint er aber, dass aus Studien bekannt sei, dass schon das Aushalten von 2 Sekunden Pause teilweise für die Gesprächspartner schwer sei und einer der beiden Partner häufig in „Sprechzwang“ gerate (Geisler, 2002).

Diese Ausführungen zeigen, dass ein als auf den ersten Blick objektives und einheitliches erscheinendes Element „Pause“, doch sehr unterschiedliche Bedeutung bekommen kann und dass Pausen unabhängig von ihrer Dauer je nach Kontext des Gesprächs oft schwer auszuhalten sind.

Dies könnte auch die uneinheitliche Bewertung des Items in der mAGBS-Jugendliche zeigen, die sich in einem niedrigen ICC widerspiegelt.

Hier bietet Interact eine sehr genaue Möglichkeit, die tatsächliche Länge einer Pause auszumessen und ihre Bedeutung für die Gesprächsabfolge anschaulich darzustellen.

#### **6.8.4 Gesprächstempo**

Auch die Definition des Gesprächstempos wirft immer wieder Fragen auf. Durch die grafische Darstellung mit Interact konnte gezeigt werden, dass ein hohes Gesprächstempo nicht allein durch die Sprechgeschwindigkeit entsteht. Das Tempo ist eher eine Mischung aus Sprechgeschwindigkeit und Pausen. Auch inhaltliche Bezüge fließen mit ein. Bleibt ein Arzt nicht bei einem Thema, sondern wechselt schnell zwischen verschiedenen Themen, wird unter Umständen das Gesprächstempo höher erlebt als wenn ein Arzt mehrere Fragen zum gleichen Thema stellt. Dies wird an dem herausgegriffenen Einzelbeispiel (siehe hierzu 5.4.4 Gesprächstempo) besonders deutlich.

#### **6.8.5 Zusammenfassung**

Mittels Interact können gerade Schlüsselstellen im Gespräch herausgegriffen werden und dadurch gut dargestellt werden, ab wann der Jugendliche anfängt zu sprechen.

Es wäre durchaus denkbar „Interact“ in Lehrsituationen einzusetzen, um den Gesprächsführenden visuell zu verdeutlichen, wie ihr Gesprächsstil - gerade was formale Aspekte wie Pausensetzen und Gesprächsanteil betrifft - aussieht.

Eine Rückmeldung an den Gesprächsführenden im Beispiel des Gesprächstempos (5.4.4 Gesprächstempo) wäre an dieser Stelle: langsamer Sprechen, weniger Fragen aneinanderreihen und mehr Pausen lassen.

Jedoch darf nicht vergessen werden, dass die Bewertung mit Interact nur rein quantita-

tiv ist, Rückschlüsse auf die Qualität des Gespräches können erst in der Verbindung mit dem Inhalt gezogen werden. So würde man am Beispiel des unter Gesprächstempo gezeigten Gespräches in Verbindung mit Inhalt sagen, dass die Studentin inhaltlich alle relevanten Fragestellungen abdeckt: sie fragt nach Schule, Interessen, Sucht, Familie, ist freundlich und dem Jugendlichen zugewandt. Ähnlich wie die RIAS ist Interact von seiner ursprünglichen Idee weniger darauf ausgelegt, zwischen und guten und schlechten Arzt-Patient-Gesprächen zu trennen. Zuerst wird jede Äußerung quantitativ einer Kategorie zugeordnet bzw. Gesprächszeiten und Pausen gemessen (Ong et al., 1998). Erst aus der Auswertung der Kategorien werden dann Schlüsse über die Gesprächsqualität gezogen.

Die RIAS verfolgt wie alle Interaktionssysteme einen anderen Ansatz bei der Bewertung als global gehaltene Instrumente. Während bei der mAGBS-Jugendliche das ganze Gespräch global in verschiedenen Kategorien qualitativ bewertet wird, wird beim RIAS jede Äußerung quantitativ einer Kategorie zugeordnet (siehe hierzu 2.3.1 Anforderungen an Bewertungsinstrumenten).

Der Fokus auf rasche qualitative Bewertung eines geführten Gespräches stellt bei der mAGBS-Jugendliche einen wesentlichen Aspekt da.

## **6.9 Emotionen und nonverbales Verhalten**

Da es in den BAS betont wurde, dass die Emotionen des Patienten in der Skala zu wenig berücksichtigt wurden und dies auch in den AGBS nicht geändert wurde, erschien es der Autorin wichtig neue Items mit einzubringen, die gezielt die Empathie des Arztes und seine Reaktion auf nonverbales Verhalten ins Visier nahmen (siehe Material und Methoden 4.2.1.1 Auswahl und Entwicklung der mAGBS-Jugendliche).

Die Autoren der Checkliste „Wien“ beschreiben aus ihrer Arbeit im Studentenunterricht, dass die psychosozialen Faktoren wie nonverbales Verhalten und Emotionen am schwersten zu vermitteln und damit auch am schwersten zu bewerten sind (Puchner-Matzner et al., 2006). Ähnliches wird auch von einer Kölner Arbeitsgruppe berichtet, die sich intensiv mit ärztlicher Gesprächsführung beschäftigt und ebenfalls eine Evaluationsbogen zur Kommunikation entworfen hat. Die Studierenden erzielen gute Ergebnisse in lehr- und lernbaren Eröffnungs- und Abschlusstechniken. Der Umgang mit Emotionen bereitet aber sowohl in der Lehre als auch in der Bewertung große Probleme (Koerfer et al., 2008).

Auch im Roter System zeigen die globalen Skalen, die die Gefühle betreffen, niedrigere Übereinstimmung zwischen den Ratern als die übrigen Skalen (Ong et al., 1998). Auch hier kann die Parallele zu Reliabilitätstestung der mAGBS-Jugendliche gezogen

werden. Emotionen und nonverbales Verhalten scheinen auch hier ein Problem in der einheitlichen Bewertung und Erfassung darzustellen.

Kritiker des Roter Systems bemängeln zudem, dass im Anwendermanual der Begriff „Empathie“ zu eng gefasst wird und es kaum möglich ist, Empathie des Gesprächsführenden, die anders als verbal ausgedrückt wird zu kodieren (Sandvik et al., 2002).

Dieses Problem zeigt sich auch bei der Definition von Empathie in Item 14 der mAGBS–Jugendliche. Auch hier wird vom Gesprächsführenden, um eine gute Bewertung zu bekommen, erwartet, dass er die Empathie in Worte fasst, dies ist eine zu enge Darstellung von Empathie.

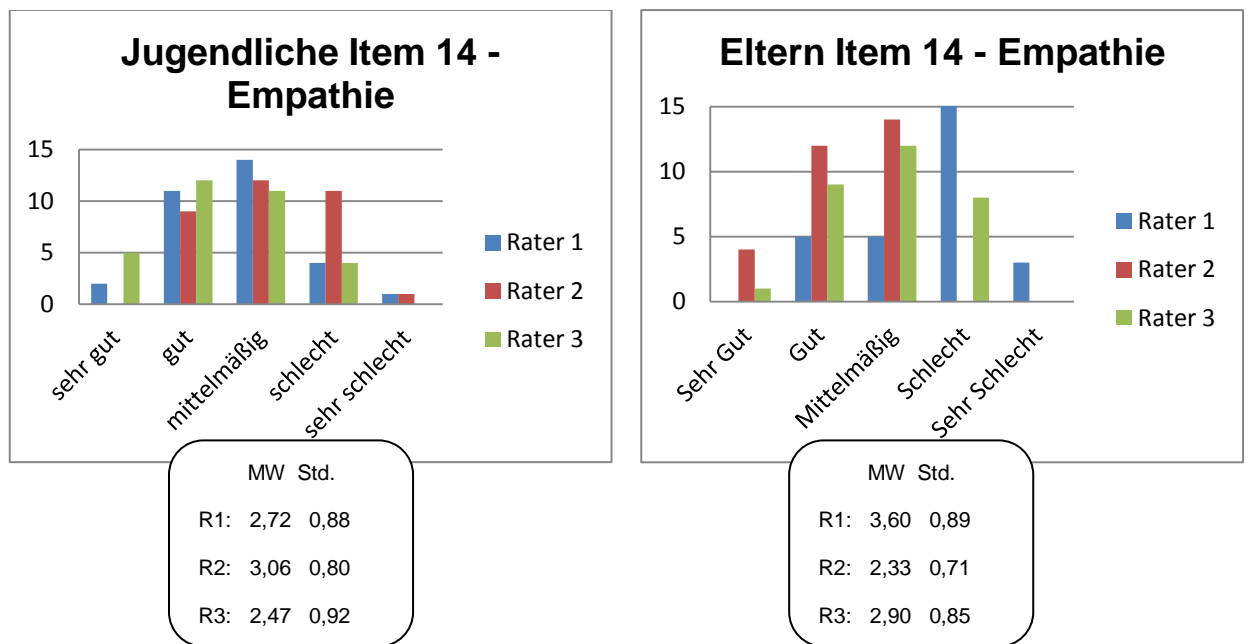
Empathie ist jedoch schwer fassbar. Ein generelles Problem bei der Reliabilitätstestung wird von Wirtz wie folgt beschrieben: „In der Psychotherapieforschung und wohl auch in anderen Gebieten sind häufig Merkmale von besonderer Relevanz, die nicht oder nur schwer reliabel zu messen sind.“ (Wirtz & Caspar, 2002). Viele Merkmale sind auch von Natur aus so schwer zu erfassen, dass sie trotz des Versuchs einer Definition von den Ratern unterschiedlich aufgefasst werden.

Ein Instrument kann eine wertvolle Unterstützung zur Erfassung kommunikativer Fertigkeiten sein, aber die Rater müssen ihre Wahrnehmung gerade im Bereich nonverbaler Kommunikation gut schulen (6.1.5 Unterschiede zwischen den Ratern).

Einige Items wurden sowohl für die mAGBS–Jugendliche als auch die mAGBS–Eltern entwickelt und neu eingeführt. Item 12 „Wiederholungen“ und Item 14 „Empathie“ findet sich identisch in der mAGBS–Eltern. Neu eingeführt wurden in beiden Versionen die Items zur Globalbeurteilung Item 23 und Item 24.

Ein Vergleich zwischen der Notenverteilung bei der m-AGBS–Jugendliche und der mAGBS–Eltern für das Item 14 „Empathie“ zeigt, dass in beiden Anwendungssituationen alle Noten vorkommen, es wird eine Tendenz zur mittleren Bewertung beobachtet. Dies zeigt sich gut in den Mittelwerten. Der ICC für dieses Item ist mit 0,66 bzw. 0,77 als gut zu beurteilen. Bei der mAGBS–Eltern kommt erschwerend hinzu, dass es sich um eine Mehrpersonensituation handelt, der Arzt muss sich sowohl dem Vater als auch der Mutter empathisch zuwenden.

Abbildung 35 Empathie in der mAGBS-Jugendliche und in der mAGBS-Eltern



Es bleibt allerdings die Frage offen, ob die Tendenz zur mittelmäßigen Bewertung an den Rater liegt, oder ob das geforderte Verhalten bei den Darstellern nicht gezeigt wird und es deswegen nicht bewertet werden kann.

## 6.10 Methodenkritik: Stärken und Schwächen der Arbeit

Das Problem des Einsatzes einer Checkliste oder eines Bewertungsinstrumentes im Unterricht ist, dass dadurch suggeriert werden könnte, dass es eine „Ideallösung“ für Kommunikation gibt. Jedoch ist Gesprächsführung in erheblichem Maße von der individuellen Persönlichkeitsstruktur eines Arztes geprägt und unterschiedliche Wege können gleichermaßen zum Ziel führen (Kopecky-Wenzel et al., 2009). Nicht nur die Vermittlung von Basisfertigkeiten wie offene Fragen stellen, Pausen machen, Zusammenfassen, sondern auch die emotionale und soziale Kompetenz sollte in einem Gesprächstraining gefördert werden (Satterfield & Hughes, 2007).

Doch wie die vorliegende Arbeit zeigt, erweisen sich gerade diese Kompetenzen schwierig in der Überprüfung mittels eines Bewertungsinstrumentes. In einer Übersichtsarbeit wurden 15 Bewertungsinstrumente, darunter auch die CCOG und die MAAS-R mit Orientierung an den Kalamazoo Consensus Statements miteinander verglichen mit dem Ergebnis, dass jedes Instrument Schwächen vor allem hinsichtlich der Erfassung psychometrischer Eigenschaften hatte (Schirmer, J. M. et al., 2005). Es zeigt sich konstant über viele Bewertungsinstrumente und in der mAGBS-Jugendliche, dass die Erfassung von psychosozialen und emotionalen Verhaltensweisen schwieri-



ger ist als beispielsweise die Identifizierung des Kriteriums, ob der Arzt den Patienten begrüßt und mit Namen anspricht.

Immer wieder wird dabei zur Diskussion gebracht, dass unterschieden werden muss, zwischen kommunikativen „Skills“, Fertigkeiten, die auch durch kurze Trainingsinterventionen vermittelbar sind, und ärztlicher Gesprächsführungskunst, die eng mit der Persönlichkeit und Charaktereigenschaften eines Arztes verbunden sind (Zoppi & Epstein, 2002). Dies kann nicht im Rahmen einer kurzen Trainingsintervention geschehen, sondern ist ein längerer Prozess.

Doch gerade hier bietet das Konzept des Wahlfaches „Kinder- und Jugendpsychiatrie“ mit Rollenspielen und Diskussion ein gutes Modell.

Neben diesen allgemeinen Kritikpunkten an globalen Ratingsystemen zeigen sich methodische Schwächen der mAGBS-Jugendliche in einer gewissen Inflexibilität. Das Instrument ist relativ starr auf bestimmte Gesprächsleistungen und Verhaltensweisen zugeschnitten. Werden diese nicht gezeigt, so kann der Gesprächsführende keine gute Benotung erreichen, obwohl er unter Umständen dennoch ein gutes Gespräch führt. Optimale Bewertungen können nur erzielt werden, wenn die Befragung relativ komplett und unkompliziert ist.

Allerdings ist es ein schon bekanntes Problem bei Bewertungsinstrumenten. Instrumente, die sehr stark auf eine Situation zugeschnitten werden, wie z.B. auf das Überbringen schlechter Nachricht, können nur in diesem Kontext angewandt werden, bedürfen dann aber jedes Mal einer Schulung der Rater und lassen kaum ein Abweichen der Situation zu, sonst wird eine Bewertung unmöglich. Allgemein gehaltene Bewertungsinstrumente sind gut für die Bewertung einsetzbar, bei speziellen Fragestellungen liefern sie jedoch nicht genügend Inhalte und die Ergebnisse verlieren an Aussagekraft. Beim Zählen und Kodieren der Äußerungen können wesentliche Punkte verloren gehen (Langewitz, W. et al., 2003).

Nach Ansicht der Autorin ist die mAGBS-Jugendliche ein Kompromiss. Es wurde versucht, wesentliche Bestandteile des Originalinstruments beizubehalten, um eine Vergleichbarkeit und Standardisierung zu erhalten. Gleichzeitig wurden jedoch Items inhaltlich abgewandelt, wo es notwendig war, um die auf den Videos dargestellte Information zu erfassen. Dieser Ansatz wie er auch schon bei der SCAG einer Modifikation der CCOG verfolgt wurde (Blake et al., 2005), könnte ein gutes Konzept sein: Ein bewährtes und validiertes Instrument zu verwenden, es in seiner Struktur zu belassen und an inhaltlichen Punkten Änderungen anzubringen.

Wie schon eingangs erwähnt, wird an dieser Stelle von den Autoren der Calgary

Cambridge Observation Guides, ein integrativer Ansatz, in dem Inhalte und Prozess des Gespraches verbunden werden, vorgeschlagen (Kurtz, S. et al., 2003).

So schreiben auch Ford et al., die eine Ratingskala zum Uberbringen schlechter Nachrichten entwickelten, dass der Kommunikationsprozess hochkomplex ist und keine Analysemethode existiert, die all seine Dimensionen abdeckt (Ford et al., 2000).

Die Autorin halt die mAGBS-Jugendliche System fur gut in Unterricht und Prufung einsetzbar. Es zeigt sich, dass das Instrument zwischen verschiedenen Gesprachisleistungen differenziert, und wesentliche Elemente eines Arzt-Patient-Gesprachs im Kontext der Kinder- und Jugendpsychiatrie abgedeckt sind. In der vorliegenden Arbeit konnte ebenfalls gezeigt werden, dass das Instrument reliabel und objektiv Gesprachsfertigkeiten misst. Es ist gut lernbar und in begrenztem Zeitrahmen anwendbar. Gerade fur Anwender, die weniger Erfahrung im Bewerten kommunikativer Fahigkeiten haben, kann das Instrument eine gute Unterstutzung sein, um zu einem differenzierten Urteil zu kommen und den Gesprachsfuhrenden Starken und Schwachen klar aufzuzeigen.

Bei den bestehenden personellen Kapazitaten der Universitaten ist es nicht immer moglich, Arzte als Prufer einzusetzen, die sehr erfahren in Kommunikation sind. Dies war beispielweise auch eine Motivation zu Entwicklung der Checkliste „Wien“ (Puchner-Matzner et al., 2006).

An den Interaktions-Analyse-Systemen wird hingegen die geringe Praktikabilitat im Alltag und der groe Aufwand bezuglich Raterschulung und Anwendung immer wieder kritisiert (Heaven et al., 2003; Ong et al., 1998).

Dieses Problem zeigt sich bei einer global angelegten Skala nicht und gerade fur Prufungssituationen kann sie gute Unterstutzung bieten.

Die Software „Interact“ liefert gute Moglichkeiten fur zur Darstellung der Verlaufsscharakteristik eines Gespraches und kann helfen formale Gesprachskriterien zu objektivieren. Es zeigt sich, dass sich die beiden Methoden, namlich der Einsatz des globalen Beurteilungsinstrumentes und der Verlaufsscharakteristik gut erganzen.

### **6.11 Verbesserungen und Ausblick**

Verbesserungen sollten vor allem die Items zur „Empathie“ und „nonverbales Verhalten“ betreffen. Daruber hinaus sollten manche Items noch genauer formuliert werden. Gerade die Items 18 „Hilfreich zur Seite stehen“ und 20 „Sicherheit im Auftreten“ erscheinen etwas unklar.

Fur eine Reliabilitatstestung war die Zahl der Videos mit 32 ausreichend, was sich vor allem in den im Vergleich zu den anderen Instrumenten mit sehr ahnlichen Ergebnis-

sen zeigt. Ein Kritikpunkt an der vorliegenden Arbeit ist, dass die Autorin das Bewertungsinstrument selbst veränderte und gleichzeitig auch als Raterin eingesetzt wurde.

Zur gezielteren Überprüfung von Trainingseffekten wäre es sinnvoll, das Instrument an denselben Gesprächsführenden vor und nach Training zu testen, um so noch besser zu untermauern, dass das Gespräch verschiedene Gesprächsleistungen abdeckt. Dazu müssten die Rater auch verblindet sein. Durch ein solches Versuchsdesign könnte auch der Trainingseffekt (beispielsweise des Wahlfaches „Kinder- und Jugendpsychiatrie“) stärker gezeigt werden.

Um den Geschlechtseffekt nachzuweisen, wäre es durchaus sinnvoll, die Fallzahl etwas zu erweitern und auch darauf zu achten, dass gleich viele Männer wie Frauen in den Vergleichsgruppen sind.

Ein weiterer Schritt, um die mAGBS-Jugendliche auf Validität zu prüfen, wäre die Videos mit einem anderen bereits seit vielen Jahren etablierten Bewertungsinstrument wie z.B. dem Calgary-Cambridge Observation Guides zu raten und zu überprüfen wie die Bewertungen zwischen den beiden Instrumenten korrelieren.

## 6.12 Zusammenfassung

**Einführung:** Kurse zur Gesprächsführung stellen deutschlandweit an vielen Fakultäten, inzwischen auch an der Ludwigs-Maximilians-Universität in München, einen wichtigen Bestandteil des Medizinstudiums dar. Eine konstruktive Rückmeldung ist wesentlich für einen Lernerfolg bei den Studenten. Ein standardisiertes Bewertungsinstrument kann dazu beitragen, Medizinstudenten und Ärzten Stärken und Schwächen in der Gesprächsführung klar aufzuzeigen. Verglichen mit englischsprachigen Bewertungsinstrumenten, gibt es nur sehr wenige validierte deutschsprachige Instrumente.

**Zielsetzung:** Ziel der vorliegenden Arbeit war es, diese Lücke an bestehenden deutschsprachigen Instrumenten zu schließen und zusätzlich einen Überblick darüber zu geben, wie man Arzt-Patient-Kommunikation umfassend beurteilen kann.

Die Arbeit besteht aus zwei Teilen. Der erste Teil umfasst die Weiterentwicklung und Reliabilitätstestung einer bereits validierten Bewertungsskala zur Evaluation der Arzt-Patient-Kommunikation.

Der zweite Teil befasst sich ausgehend von den Ergebnissen der Reliabilitätstestung der modifizierten Bewertungsskala mit Fragen zu allgemeinen Kommunikationsstrukturen in der Arzt-Patient-Beziehung. So wurde der Frage nachgegangen, welchen Einfluss Redezeiten haben, welche Wirkung Gesprächspausen und Gesprächstempo haben und durch welchen Fragestil ein Patient zum Reden angeregt werden kann. Darüber hinaus wurde eingehend diskutiert, welchen Einfluss Geschlecht, Rollenspiele, Training und Berufserfahrung auf die jeweilige Gesprächsleistung haben und in welchem Zusammenhang sie mit der der Bewertung stehen.

**Methodik:** Bei der Auswahl eines geeigneten Instruments wurde Wert gelegt, auf ein einfaches, praktikables und bereits validiertes System zurückzugreifen und dieses weiterzuentwickeln. Das Instrument sollte vor allem den Gesprächsführenden im Blickpunkt haben und seine Fähigkeiten beobachten und beurteilen.

Ein wichtiges Kriterium war außerdem eine gute Verständlichkeit und die Anwendbarkeit in einem angemessenen Zeitrahmen. So wurde ein 2008 veröffentlichtes Instrument, die Aufklärungsgesprächsbewertungsskala (AGBS), für Arzt-Jugendliche-Gespräche modifiziert.

Die vorliegende Arbeit wurde im Rahmen von Trainingssituationen für ärztliche Gesprächsführung durchgeführt. Ein Teil der auf Video aufgenommenen und bewerteten Gespräche stammt aus dem Wahlfach „Kinder- und Jugendpsychiatrie“, ein Teil aus dem „Kompaktkurs Pädiatrie“ für angehende Allgemeinärzte.

Es lagen drei Videogruppen vor: einmal Medizinstudenten ohne Training, einmal Medi-

zinstudenten mit Training und eine Videogruppe von Allgemeinärzten mit Berufserfahrung.

Neben der Bewertung mit dem Instrument wurde die Software „Interact“ eingesetzt, um Gesprächsverläufe grafisch darzustellen, sowie auf Gesprächsanteile, Fragestile und Schlüsselmomente zu untersuchen.

**Ergebnisse:** Es wurden insgesamt 32 Videos von 3 zuvor geschulten Ratern mittels des Bewertungsinstrumentes mAGBS-Jugendliche, sowie der Software „Interact“ beurteilt. Die Anwendungstestung des modifizierten Instrumentes zeigte zufriedenstellende Ergebnisse im Vergleich mit dem Original und anderen im englischsprachigen Raum häufiger verwendeten Instrumenten.

Das Instrument erwies sich als einfach in der Handhabung und der Zeitaufwand hielt sich mit etwa der zweieinhalbfachen Gesprächszeit für die Beurteilung in Grenzen.

Die Reliabilitätstestung zeigte gute Werte für interne Konsistenz (Cronbachs Alpha 0,85) sowie Interraterreliabilität (ICC 0,82) in der Gesamtwertung, jedoch variierende Testgüte auf Einzelitemebene. Es zeigte sich, dass Items, die formale Kriterien abfragen, gute Übereinstimmung zwischen den Ratern erzielten. Items, die emotionales und nonverbales Verhalten beurteilten, bekamen niedrigere Werte hinsichtlich der Interraterreliabilität.

Es ließ sich darüber hinaus ein Gruppenunterschied zwischen der Trainingsgruppe und der berufserfahrenen Gruppen nachweisen.

Die Ergebnisse zeigen außerdem, dass Frauen tendenziell besser in Gesprächsfertigkeiten beurteilt werden.

Die Analyse mit Interact zeigte, dass den Studenten das Zulassen von Pausen oft schwerfällt. Häufig wird das Gespräch durch das Stellen vieler geschlossener Fragen dominiert. Hier fanden sich ebenfalls Unterschiede zwischen den Gruppen von Medizinstudenten mit und ohne Training, sowie den Allgemeinärzten.

**Diskussion und Schlussfolgerung:** Die mAGBS-Jugendliche ist ein reliables, objektives und valides Instrument für die Beurteilung von Arztgesprächen mit Jugendlichen. Einem globalen Instrument sind jedoch auch verglichen mit einem Interaktion–Analyse–System gewisse Grenzen gesetzt. Die visuelle Darstellung mittels eines Interaktionsgraphen bei Interact ist eine sehr anschauliche Möglichkeit, Kommunikationsstrukturen darzustellen und Muster in der Kommunikation zu entdecken. Gerade der subjektive und objektive Unterschied in der Wahrnehmung von Pausen und Gesprächstempo kann so verdeutlicht werden.

Rollenspiele erwiesen sich in der vorliegenden Arbeit als realistisch.

## Diskussion

---

Ein Trainingseffekt ließ sich für leicht zu operationalisierende Fertigkeiten wie beispielsweise „Begrüßung“ und „Gesprächseinstieg“ nachweisen.

## 7 Literaturverzeichnis

- Back, A. L., Arnold, R. M., Baile, W. F., Fryer-Edwards, K. A., Alexander, S. C., Barley, G. E. et al. (2007). Efficacy of Communication Skills Training for Giving Bad News and Discussing Transitions to Palliative Care. *Archives of Internal Medicine*, 167, 453-460.
- Baile, W. F., Buckman, R., Lenzi, R., Glober, G., Beale, E. A. & Kudelka, A. P. (2000). Spikes - A Six-Step Protocol for Delivering Bad News: Application to the Patient with Cancer. *The Oncologist*, 5, 302-311.
- Beckman, H. B. & Frankel, R. M. (1984). The effect of physician behavior on the collection of data. *Ann Intern Med*, 101 (5), 692-696.
- Benbassat, J. & Baumal, R. (2002). A step-wise role playing approach for teaching patient counseling skills to medical students. *Patient Educ Couns*, 46(2), 147-152.
- Blake, K., Vincent, N., Wakefield, S., Murphy, J., Mann, K. & Kutcher, M. (2005). A structured communication adolescent guide (SCAG): assessment of reliability and validity. *Med Educ*, 39 (5), 482-491.
- Bortz, J. (2004). *Statistik für Human - und Sozialwissenschaftler* (6.). Heidelberg: Springer.
- Braun-Scharm, H. & Poustka, F. (2003). *Suizidalität im Kindes-und Jugendalter* <http://www.uni-duesseldorf.de/>
- Brown, R., Doonan, S. & Shellenberger, S. (2005). Using children as simulated patients in communication training for residents and medical students: a pilot program. *Acad Med*, 80 (12), 1114-1120.
- Buckman, R. (1992). *How to break bad news - a Guide for health care professionals* Baltimore: John Hopkins University Press.
- Buckman, R. (2002). Communications and emotions. *British Medical Journal*, 325, 672.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (Bd. 2. ). München: Pearson Studium.
- Cegala, D. J. & Lenzmeier Broz, S. (2002). Physician communication skills training: a review of theoretical backgrounds, objectives and skills. *Med Educ*, 36(11), 1004-1016.
- Chenot, J.-F. & Ehrhardt, M. (2003). Objective structured clinical examination (OSCE) in der medizinischen Ausbildung : Eine Alternative zur Klausur. *Z. Allg. Med.* , 79.
- Cohen, D. S., Colliver, J. A., Marcy, M. S., Fried, E. D. & Swartz, M. H. (1996). Psychometric properties of a standardized-patient checklist and rating-scale form used to assess interpersonal and communication skills. *Acad Med*, 71 (1 Suppl), S87-89.

- Cottrell, L. A., Nield, L. S. & Perkins, K. C. (2006). Effective interviewing and counseling of the adolescent patient. *Pediatr Ann*, 35(3), 164-166, 169-172.
- Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R., Thorpe, K. E. & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*, 296(9), 1094-1102.
- Doering, T., Steuernagel, B., Kape, A., Lieck, M. & B. (2008). Die partizipative Entscheidungsfindung mit dem Patienten - Ein Schulungskonzept für Ärzte. *Pädiatrische Praxis*, 72, 51-58.
- Dorfmueller, M. (Hrsg.). (2001). *Die ärztliche Sprechstunde Arzt, Patient und Angehörige im Gespräch* Landsberg/Lech: Ecomed.
- Ellgring, H. (1989). *Der Wert des Videos in der Psychotherapie*. Berlin Heidelberg: Springer.
- Fallowfield, L. (1996). Things to consider when teaching doctors how to deliver good, bad and sad news. *Medical Teacher*, 18(1), 27-30.
- Fallowfield, L., Jenkins, V., Farewell, V., Saul, J., Duffy, A. & Eves, R. (2002). Efficacy of a Cancer Research UK communication skills training model for oncologists: a randomised controlled trial. *Lancet*, 359(9307), 650-656.
- Ford, S., Hall, A., Ratcliffe, D. & Fallowfield, L. (2000). The Medical Interaction Process System (MIPS): an instrument for analysing interviews of oncologists and patients with cancer. *Social Science & Medicine*, 50, 553-566.
- Frank, R. (2005). Suizidalität von Kindern und Jugendlichen - Erkennen, Einschätzen, Vorgehen. *Notfall und Rettungsmedizin*, 3, 216-221.
- Frank, R., Gegenfurtner, G., Steininger, C., Kopecky-Wenzel, M. & Noterdaeme, M. (2009). Was lernen Studenten im Wahlfach Kinder- und Jugendpsychiatrie. *Z Kinder Jugendpsychiatr Psychother*, 37(2), 129-134.
- Geisler, L. (2002). *Arzt und Patient-Begegnung im Gespräch*. Frankfurt am Main: pmi Verlag AG
- Günter, J. (2004). *Die Anwendung der "Maastricht History - taking and Advice Checklist" (MAAS-R) in der Videoevaluation eines Trainingsprogramms zur Behandlung somatoformer Störungen in der Hausarztpraxis*. Freiburg
- Hardoff, D. & Schonmann, S. (2001). Training physicians in communication skills with adolescents using teenage actors as simulated patients. *Med Educ*, 35(3), 206-210.
- Heath, C., Luff, P. & Sanchez Svensson, M. (2007). Video an qualitative research: analysing medical practice and interaction. *Med Educ*, 41, 109-116.



- Heaven, C., Maguire, P. & Green, C. (2003). A patient-centred approach to defining and assessing interviewing competency. *Epidemiol Psychiatr Soc*, 12(2), 86-91.
- Hewson, M. G. & Little, M. L. (1998). Giving feedback in medical education: verification of recommended techniques. *J Gen Intern Med*, 13(2), 111-116.
- Hibbeler, B. & Korzilius, H. (2008). Arztberuf - die Medizin wird weiblich. *Deutsches Ärzteblatt*.
- Hoffmann, K., Schultz, J.-H., Conrad, C., Hancke, R., Lauber, H., Schönemann, J. et al. (2007). Kommunikationsschulung mittels "standardisierter Eltern" im Fachbereich der Pädiatrie. *GMS - Zeitschrift für Medizinische Ausbildung*, 24.
- Jungbauer, J., Alfermann, D., Kamenik, C. & Brähler, E. (2003). Vermittlung psychosozialer Kompetenzen mangelhaft. *Psychother Psychosom Med Psychol*, 53(7), 319-321.
- Kemper, K. J., Foy, J. M., Wissow, L. & Shore, S. (2008). Enhancing communication skills for pediatric visits through on-line training using video demonstrations. *BMC Med Educ*, 8, 8.
- Kiessling, C., Dieterich, A., Götz, F., Hölzer, H., Langewitz, W., Isabel, M. et al. (2008). Basler Consensus Statement "Kommunikative und soziale Kompetenzen im Medizinstudium": Ein Positionspapier des GMA - Ausschusses kommunikative und soziale Kompetenz. *GMS - Zeitschrift für Medizinische Ausbildung*, 25.
- Koch, E. & Resch, F. (2009). Das Fach Kinder- und Jugendpsychiatrie im Heidelberger Curriculum Medicinale (HeiCuMed) - Praxisorientiertes Lernen im reformierten Studiengang. *Z Kinder Jugendpsychiatr Psychother*, 37(2), 117-122.
- Koerfer, A., Köhle, K., Obliers, R., Sonntag, B., Thomas, W. & Albus, C. (2008). Training und Prüfung kommunikativer Kompetenz. Aus - und Fortbildungskonzepte zur ärztlichen Gesprächsführung. *Gesprächsforschung - Online - Zeitschrift zur verbalen Interaktion*(9), 34 - 78.
- Köhle, K., Kaerger - Sommerfeld, H., Koerfer, A., Obliers, R. & Thomas, W. (2001). Können Ärzte ihr Kommunikationsverhalten verbessern? In H. C. Deter (Hrsg.), *Psychosomatik am Beginn des 21. Jahrhunderts- Chancen einer biopsychosozialen Medizin*: Verlag Hans Huber. 301-310
- Kopecky-Wenzel, M., Maier, E. M., Muntau, A. C., Reinhardt, D. & Frank, R. (2009). Überbringen schlechter Nachrichten - videogestützte Trainingseinheit für Medizinstudenten. *Z Kinder Jugendpsychiatr Psychother*, 37(2), 139-144.
- Kruse, J. & Tress, W. (2001). Der Hausarzt im Gespräch mit seinen Patienten. In H. C. Deter (Hrsg.), *Psychosomatik am Beginn des 21. Jahrhunderts - Chancen einer bio-*

*psychosozialen Medizin*: Verlag Hans Huber. 294-300

Kurtz, S., Silverman, J., Benson, J. & Draper, J. (2003). Marrying content and process in clinical method teaching: enhancing the Calgary-Cambridge guides. *Acad Med*, 78(8), 802-809.

Kurtz, S. M. & Silverman, J. D. (1996). The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. *Med Educ*, 30(2), 83-89.

Lang, F., McCord, R., Harvill, L. & Anderson, D. S. (2004). Communication assessment using the common ground instrument: psychometric properties. *Fam Med*, 36(3), 189-198.

Langer, T. (2009). Das Gespräch zwischen Patient, Eltern und Arzt in der Pädiatrie. In T. Langer & M. Schnell (Hrsg.), *Das Arzt-Patient Patient - Arzt Gespräch*. München: Marseille Verlag. 43 - 52

Langer, T. & Schnell, M. (2009). *Das Arzt-Patient Patient - Arzt Gespräch. Ein Leitfa-*  
*den für Klinik und Praxis*. München: Marseille Verlag.

Langewitz, W., Nubling, M. & Weber, H. (2003). A theory-based approach to analysing conversation sequences. *Epidemiol Psychiatr Soc*, 12(2), 103-108.

Langewitz, W. A., Eich, P., Kiss, A. & Wossmer, B. (1998). Improving communication skills--a randomized controlled behaviorally oriented intervention study for residents in internal medicine. *Psychosom Med*, 60(3), 268-276.

Lutz-Tittel, J. (1999). *Weiterentwicklung eines Instrumentes zur standardisierten Form der Beobachtung von Eltern-Kind-Interaktion und Untersuchung der Beobachtereinflüsse*. Dissertation, Medizinische Fakultät der Ludwig Maximilians-Universität München, München.

Maguire, P. (1980). Teaching medical students to interview psychiatric patients. *Psychiatric bulletin*, 4, 188-190.

Maguire, P., Fairbairn, S. & Fletcher, C. (1986). Consultation skills of young doctors: I--Benefits of feedback training in interviewing as students persist. *Br Med J (Clin Res Ed)*, 292 (6535), 1573-1576.

Maguire, P. & Pitceathly, C. (2002). Key communication skills and how to acquire them. *BMJ*, 325(7366), 697-700.

Makoul, G. (2001). Essential elements of communication in medical encounters: the Kalamazoo consensus statement. *Acad Med*, 76(4), 390-393.

Miller, S. J., Hope, T. & Talbot, D. C. (1999). The development of a structured rating

schedule (the BAS) to assess skills in breaking bad news. *British Journal of Cancer*, 80, 792-800.

Nestel, D. S. & Tierney, T. (2007). Role-play for medical students learning about communication: Guidelines for maximising benefits. *BMC Medical Education*. <http://creativecommons.org/licenses/by/2.0/>

Ong, L. M., Visser, M. R., Kruyver, I. P., Bensing, J. M., van den Brink-Muinen, A., Stouthard, J. M. et al. (1998). The Roter Interaction Analysis System (RIAS) in oncological consultations: psychometric properties. *Psychooncology*, 7(5), 387-401.

Petersen, C., Busche, W., Bergelt, C. & Huse-Kleinstoll, G. (2005). Kommunikations-training als Teil des Medizinstudiums: Ein Modellversuch. *GMS Zeitschrift für Medizinische Ausbildung*.

Pringle, M. & Stewart-Evans, C. (1990). Does awareness of being video recorded affect doctors' consultation behaviour? *Br J Gen Pract*, 40(340), 455-458.

Ptacek, J. & Tara, L. E. (1996). Breaking bad news. A review of the literature. *Journal of the American Medical Association*, 276, 496-502.

Puchner-Matzner, I., Gleiss, A., Schmidts, M. & Frischenschlager, O. (2006). Die Checkliste PK "Professionelles ärztliches Kommunikationsverhalten" in Unterricht und Evaluation kommunikativer Fertigkeiten im Medizinstudium. *GMS - Zeitschrift für Medizinische Ausbildung*, 23.

Ram, P., Grol, R., Rethans, J. J., Schouten, B., van der Vleuten, C. & Kester, A. (1999). Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ*, 33(6), 447-454.

Ravens-Sieberer, U., Erhart, M., Wille, N. & Bullinger, M. (2008). Health-related quality of life in children and adolescents in Germany: results of the BELLA study. *Eur Child Adolesc Psychiatry*, 17 Suppl 1, 148-156.

Regehr, G., MacRae, H., Reznick, R. K. & Szalay, D. (1998). Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med*, 73(9), 993-997.

Remschmidt, H. (1992). *Psychiatrie der Adoleszenz*. Stuttgart: Thieme Verlag.

Robinson, J., Walley, T., Pearson, M., Taylor, D. & Barton, S. (2002). Measuring consultation skills in primary care in England: evaluation and development of content of the MAAS scale. *Br J Gen Pract*, 52(484), 889-893.

Ronge, J. (1997) Videografisches Arbeiten in Psychiatrie und Psychotherapie (1.) Wis-

senschaft und Praxis

Roter, D. (2006). The Roter method of interaction process analysis. Manual. [www.rias.org](http://www.rias.org)

Roter, D. & Larson, S. (2002). The Roter interaction analysis system (RIAS): utility and flexibility for analysis of medical interactions. *Patient Education and Counseling*, 46, 243-251.

Roter, D. L., Hall, J. A. & Aoki, Y. (2002). Physician gender effects in medical communication: a meta-analytic review. *JAMA*, 288(6), 756-764.

Sandvik, M., Hilde, E., Lind, M., P., G., J., T. & Finset, A. (2002). Analysing medical dialogues: strength and weakness of Roter's Interaction analysis system. *Patient Educ Couns*, 46, 235-241.

Satterfield, J. M. & Hughes, E. (2007). Emotion skills training for medical students: a systematic review. *Med Educ*, 41, 935-941.

Scheffer, S., Muehlinghaus, I., Froehmel, A. & Ortwein, H. (2008). Assessing students' communication skills: validation of a global rating. *Adv Health Sci Educ Theory Pract*, 13(5), 583-592.

Schildmann, J., Härlein, J., N., B., Schlögl, M. & Vollmann, J. (2006). Breaking bad news: evaluation study on self perceived competences and views of medical and nursing students taking part in a collaborative workshop. *Support Care in Cancer*, 14, 1157-1161.

Schildmann, J. & Schildmann, J. (2009). Das Überbringen einer schlechten Nachricht. In T. Langer & M. Schnell (Hrsg.), *Das Arzt - Patient Patient - Arzt Gespräch*. München: Marseille Verlag. 89-99

Schildmann, J. & Wand, S. (2008). Manual zur Verwendung der Aufklärungsgesprächsbewertungsskala in der Lehrveranstaltung "Wahrheit am Krankenbett" das Aufklärungsgespräch mit schwer kranken Patienten.

Schirmer, C., Stahl, K. & Günther, W. (2009). Patientenbefragung zum kommunikativen Verhalten von Ärzten in Qualitätskontrolle und -entwicklung. In T. Langer & M. Schnell (Hrsg.), *Das Arzt - Patient Patient - Arzt Gespräch*. München: Marseille Verlag. 201-213

Schirmer, J. M., Mauksch, L., Lang, F., Marvel, M. K., Zoppi, K., Epstein, R. M. et al. (2005). Assessing communication competence: a review of current tools. *Fam Med*, 37(3), 184-192.

Schmid - Mast, M. & Klöckner - Cronauer, C. (2009). Geschlechtsspezifische Aspekte

des Gesprächs zwischen Arzt bzw Ärztin und Patient bzw Patientin. In T. Langer & M. Schnell (Hrsg.), *Das Arzt - Patient Patient-Arzt - Gespräch*. München: Marseille Verlag 135-142.

Schmidt, M. (1998). *Notfall- und Krisensituationen bei Kindern- und Jugendlichen*. München Wien Baltimore: Urban & Schwarzenberg.

Schnell, M. & Langer, T. (2009). Arzt - Patient - Kommunikation im Medizinstudium. In T. Langer & M. Schnell (Hrsg.), *Das Arzt - Patient Patient - Arzt Gespräch*. München: Marseille Verlag. 189-201

Schwantes, U. (2009). Die Auflösung von Konflikten und Unzufriedenheit in der Kommunikation? In T. Langer & M. Schnell (Hrsg.), *Das Arzt - Patient Patient - Arzt Gespräch*. München: Marseille Verlag.153-163

Simmenroth-Nayda, A., Chenot, J.-F., Fischer, T., Scherer, M., Stabske, B. & Kochen, M. (2007). Mit Laienschauspielern das ärztliche Gespräch trainieren. *Deutsches Ärzteblatt*, 13, 847-852.

Simmenroth-Nayda, A., Nolte, C., Fischer, T. & Himmerl, W. (2008). Der Calgary Cambridge Observation Guide: Validierung einer ins Deutsche übersetzten Version. Greifswald: Gesellschaft für medizinische Ausbildung (GMA).

Simpson, M., Buckman, R., Stewart, M., Maguire, P., Lipkin, M., Novack, D. et al. (1991). Doctor-patient communication: the Toronto consensus statement. *BMJ*, 303(6814), 1385-1387.

Smith, R. (2002). *Patient centered interviewing an evidence -based method (2.)*. Philadelphia: Lippincott Williams and Wilkins.

Terzioglu, P., Jonitz, B., Schwantes, U. & Burger, W. (2003). Kommunikative und soziale Kompetenzen-Vermittlung muss im Medizinstudium beginnen. *Deutsches Ärzteblatt*, 10, 2277-2279.

Thiel, T. (1997). Film- und Videotechnik in der Psychologie. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung*. Bern/Göttingen: Verlag Hans Huber. 347-384

Van Dalen, J., Prince, C. J., Scherpbier, A. J. & Van Der Vleuten, C. P. (1998). Evaluating Communication Skills. *Adv Health Sci Educ Theory Pract*, 3(3), 187-195.

van Dulmen, A. M. & Holl, R. A. (2000). Effects of continuing paediatric education in interpersonal communication skills. *Eur J Pediatr*, 159(7), 489-495.

van Thiel, J., Ram, P. & van Dalen, J. (2000). MAAS-Global Manual Guidelines to the ratings of communication skills and clinical skills of doctors. Maastricht.

Wand, S., Schildmann, J., Buchardi, N. & Vollmann, J. (2008). Die

Aufklärungsgesprächbewertungsskala (AGBS): Ein Instrument zur Bewertung kommunikativer Kompetenzen bei der Aufklärung von Patienten über Tumorerkrankungen. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 101, 645-651.

Warnke, A. & Lehmkuhl, G. (2003). *Kinder-und Jugendpsychiatrie und Psychotherapie in der Bundesrepublik Deutschland* (3). Stuttgart: Schattauer.

Wasserman, R. & Inui, T. (1983). Systematic Analysis of Clinician-Patient Interactions: A Critique of recent Approaches with suggestions for Future Research. *Medical Care*, 11, 279-292.

Wassmer, E., Minnaar, G., Abdel Aal, N., Atkinson, M., Gupta, E., Yuen, S. et al. (2004). How do paediatricians communicate with children and parents? *Acta Paediatr*, 93(11), 1501-1506.

Weissenrieder N. & Stier, B. (2006). *Jugendmedizin, Gesundheit und Gesellschaft*.

Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Freiburg: Hogrefe.

Zoppi, K. & Epstein, R. M. (2002). Is communication a skill? Communication behaviors and being in relation. *Fam Med*, 34(5), 319-324.

## **8 Anhang**

---

## 8.1 Die mAGBS-Jugendliche (Bewertungsbogen)

# m(odifizierter)AGBS – Jugendliche

---

Datum: \_\_\_\_\_

Video ID: \_\_\_\_\_

Name des Raters: \_\_\_\_\_

Laufende Nummer des Raters: \_\_\_\_\_

Beginn des Ratings um \_\_\_\_\_

Uhr

Einwilligung der Teilnehmer:  Vorhanden  Keine Einwilligung

Arztrolle - Ausbildung:  Student  Arzt

Arztrolle - Geschlecht:  m  w

Darsteller - Jugendlicher – Geschlecht:  m  w

---

### BEWERTUNG

#### A. Einleitung des Gesprächs/Gesprächsbeginn/Rahmumgebung

1. Geeignete Umgebung  1  2  3  4  
 5

2. Angemessene Begrüßung  1  2  3  4  
 5

3. Befinden  1  2  3  4  
 5

#### B. Information bekommen

4. Ins Gespräch kommen  1  2  3  4  
 5

5. Langsame Annäherung  1  2  3  4  
 5

6. Verständliche Sprache  1  2  3  4  
 5



- 
7. Angemessen Pausen  1  2  3  4  
 5

**C. Bedenken des Jugendlichen herausfinden**

8. Nachfragen bei Problemen  1  2  3  4  
 5

9. Nonverbale Äusserungen  1  2  3  4  
 5

**D. Perspektive vermitteln**

10. Äußerungen. angepasst  1  2  3  4  
 5

11. Gesprächsstruktur  1  2  3  4  
 5

12. Wiederholungen und Zus.  1  2  3  4  
 5

13. Positive Aspekte  1  2  3  4  
 5

14. Empathie  1  2  3  4  
 5

15. Fachlich korrekt  1  2  3  4  
 5

---

**E. Allgemeine Überlegungen**

16. Schlüsselbereiche  1  2  3  4  
 5

17. Tiefgehendes besprechen  1  2  3  4  
 5

---

18. Hilfreich zur Seite stehen	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				
19. Körpersprache	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				
20. Unbeholfenheit	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				
21. Gesprächstempo	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				
22. Beenden	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				

---

23. Subjektiv – Arzt	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				
24. Subjektiv – Patientenrollenspiel	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
<input type="checkbox"/> 5				

Länge: \_\_\_\_\_

Minuten

Tonqualität<sup>23</sup>:  Gut  Mittelmäßig  Schlecht

Bildqualität<sup>24</sup>:  Gut  Mittelmäßig  Schlecht

Vollständigkeit:  Ja  Anfang abgeschnitten  Ende abg.  Mehr als 15 Sek. fehlen

Rollenidentifikation<sup>25</sup>:  komplett  annähernd komplett  wenig

Sprachliche Probleme:  keine oder geringe  beeinträchtigende

Störfaktoren<sup>26</sup>:  keine oder geringe  beeinträchtigende

<sup>23</sup> Tonqualität: Gut (Alles zu verstehen); Mittelmäßig (Einzelne Worte unverständlich); Schlecht (Ganze Sätze nicht zu verstehen)

<sup>24</sup> Bildqualität: Gut (Alle Beteiligten inklusive Mimik erkennbar); Mittelmäßig (Alle Beteiligten im Bild); Schlecht (Mindestens ein Beteiligter nicht im Bild)

<sup>25</sup> Rollenidentifikation: komplett (kein aus der Rolle fallen); annähernd komplett (einzelne Teilnehmer fallen für <15 Sekunden aus der Rolle)

---

Anzahl der Unterbrechungen während des Ratings

keine  1  2  3

>3

Ende des Ratings um \_\_\_\_\_ Uhr

Unterschrift (Rater) \_\_\_\_\_

---

<sup>26</sup> Störfaktoren: z.B. andere Personen die sich einmischen, störender Bau- oder Straßenlärm

---

## 8.2 Ergänzende Tabellen

### 8.2.1 Geschlecht der Arztdarsteller

Jugendlicher	Rater 1	Rater 2	Rater 3	Geschlecht Arzt
1.Video	1,97	2,35	2,53	m
2. Video	2,25	2,57	2,63	w
3.Video	2,00	2,33	2,38	w
4. Video	1,85	2,35	2,23	w
5.Video	2,32	2,72	2,37	w
6.Video	1,68	1,65	1,43	w
7. Video	2,33	2,58	2,78	m
8.Video	2,52	2,13	2,28	w
9.Video	2,18	2,75	2,92	w
10. Video	1,78	2,25	2,92	w
11.Video	1,98	2,18	2,30	w
<b>Notfall</b>				
1.Video	2,83	2,80	2,85	m
2.Video	1,67	2,20	2,67	w
3.Video	1,72	1,73	1,63	w
4.Video	2,00	1,98	1,77	w
5.Video	1,95	2,43	2,05	w
6.Video	1,57	2,23	1,77	w
7.Video	1,98	2,33	1,93	w
8.Video	2,63	2,75	2,87	w
9.Video	1,65	2,08	2,47	w
10.Video	1,95	2,80	2,78	w
11.Video	1,82	2,23	1,97	m
<b>Allgemeinarzt</b>				
Video 1	2,38	1,78	1,73	m
Video 2	3,07	2,73	3,33	m
Video 3	1,73	2,00	2,18	w
Video 4	1,87	2,32	2,43	m
Video 5	2,87	3,00	2,75	m
Video 6	2,33	2,32	2,10	m
Video 7	2,88	2,58	2,70	w

Video 8	1,78	2,05	1,77	w
Video 9	2,05	2,00	1,67	w
Video 10	2,17	2,92	2,17	w

### Item 8 „Nachfragen bei Problemen“

Weiblich: sehr gut 48%, gut 35%, 17% mittelmäßig

Männlich sehr gut 56%, gut 22%, mittelmäßig 11%, schlecht 11%

### Item 9 „Nonverbales Verhalten“

Rater 1 Weiblich: gut 9% mittelmäßig 78% schlecht 13%

Männlich: 11 % gut, 67% mittelmäßig, 22% schlecht

### Item 14 „Nonverbales Verhalten“

### Item 18 „Hilfreich zur Seite stehen“

### Item 19 „Körpersprache“

## 8.2.2 Gesprächsverläufe

Tabelle 27 - Gesprächsanteile nach Videogruppen aufgeschlüsselt

Gesprächsanteile	Arzt	Geschlecht	Jugendlicher	Geschlecht
<i>Jugendlichenvideos</i>				
1. Video	40%	w	60%	m
2. Video	51%	w	49%	w
3. Video	55%	w	45%	w
4. Video	55%	w	45%	w
5. Video	43%	w	57%	w
6. Video	44%	w	56%	w
7. Video	48%	w	52%	m
8. Video	35%	w	65%	w
9. Video	49%	m	51%	w
10. Video	47%	w	53%	w
11. Video	39%	m	61%	w
Durchschnittlicher Redeanteil	<b>46%</b>		<b>54%</b>	
<i>Notfallvideos</i>				
1. Video	37%	m	63%	m
2. Video	41%	w	59%	w
3. Video	56%	w	44%	w
4. Video	51%	w	49%	w
5. Video	52%	w	48%	w
6. Video	68%	m	32%	w

7. Video	32%	m	68%	w
8. Video	37%	w	63%	w
9. Video	47%	w	53%	w
10. Video	58%	w	42%	w
11. Video	46%	w	54%	m
Durchschnittlicher Redeanteil	<b>48%</b>		<b>52%</b>	
<i>Allgemeinärztevideos</i>				
1.Video	73%	w	27%	m
2.Video	45%	m	55%	m
3.Video	63%	w	37%	w
4.Video	57%	m	43%	m
5. Video	71%	w	29%	m
6. Video	64%	w	36%	m
7. Video	82%	w	18%	w
8. Video	77%	m	23%	w
9. Video	53%	m	47%	w
10. Video	45%	w	55%	w
Durchschnittlicher Redeanteil	<b>63%</b>		<b>37%</b>	

Tabelle 28 Fragestil nach Videogruppen

Fragestil in Prozent	geschlossen	offen	Anzahl Hörersignal Arzt
<i>Jugendlichenvideos</i>			
1. Video	61%	39%	23
2. Video	52%	48%	12
3. Video	65%	35%	10
4. Video	86%	14%	10
5. Video	80%	29%	4
6. Video	41%	59%	18
7. Video	79%	21%	6
8. Video	80%	20%	27
9. Video	59%	41%	19
10.Video	64%	36%	10
11. Video	46%	54%	23
Durchschnitt	<b>65%</b>	<b>35%</b>	
<i>Notfallvideos</i>			
1. Video	40%	60%	4
2. Video	61%	39%	5
3. Video	71%	29%	19
4. Video	70%	30%	8
5. Video	44%	56%	4

6. Video	47%	53%	5
7. Video	52%	48%	19
8. Video	74%	26%	9
9. Video	40%	60%	4
10. Video	50%	50%	13
11. Video	38%	62%	7
Durchschnitt	<b>53%</b>	<b>47%</b>	
<i>Allgemeinärzteeideos</i>			
1.Video	70%	30%	11
2.Video	63%	37%	8
3.Video	84%	16%	8
4.Video	70%	30%	10
5. Video	62%	38%	8
6. Video	72%	28%	8
7. Video	63%	37%	2
8. Video	33%	67%	8
9. Video	70%	30%	18
10. Video	33%	67%	12
Durchschnitt	<b>62%</b>	<b>38%</b>	

## 9 Verzeichnisse

### 9.1 Tabellenverzeichnis

Tabelle 1 Gegenüberstellung von Bewertungsinstrumenten.....	16
Tabelle 2 Überblick und Gegenüberstellung Bewertungsinstrumente.....	21
Tabelle 3 Lernziele Wahlfach "Kinder- und Jugendpsychiatrie" .....	29
Tabelle 4 Lernziele im Kompaktkurs Pädiatrie.....	31
Tabelle 5 Studiendesign.....	33
Tabelle 6 Allgemeine Modifikationen in den Items.....	36
Tabelle 7 Korrelation Item 23 und Item 24.....	66
Tabelle 8 Mittelwert Themenblöcke über alle Videoaufnahmen.....	69
Tabelle 9 Übersicht über Item-Skala-Korrelation .....	72
Tabelle 10 Korrelation zwischen den Ratern in den Themenblöcken.....	73
Tabelle 11 Pearsonscher Korrelationskoeffizient.....	73
Tabelle 12 Varianzanalyse für die Themenblöcke .....	75
Tabelle 13 ICC über Einzelitems .....	76
Tabelle 14 Überblick Mittelwerte und Effektstärke bei Jugendlichen- und Notfallvideos .....	78
Tabelle 15 Überblick Mittelwert und Effektstärke bei Jugendlichen- und Allgemeinärztevideos .....	79
Tabelle 16 Gesprächsstil des Jugendlichen in den einzelnen Videos .....	82
Tabelle 17 Durchschnittlicher Redeanteil und Fragestil .....	85
Tabelle 18 Korrelationen zwischen Fragestil und Redeanteil bzw. nicht unterbrochener Äußerungen Jugendlicher .....	86
Tabelle 19 Notenvergabe Item 7 und Item 21 für 6.Video.....	89
Tabelle 20 Notenvergabe für 4.Video .....	90
Tabelle 21 Cronbachs Alpha .....	101
Tabelle 22 Varianzanalyse mAGBS-Jugendliche, mAGBS-Eltern, AGBS und BAS...	102
Tabelle 23 ICC-Werte im Vergleich .....	103
Tabelle 24 Interraterübereinstimmung der Checkliste Wien.....	103
Tabelle 25 Vergleich zwischen ICC von Item „Pause“ und „Gesprächstempo“ bei der mAGBS-Jugendliche und der mAGBS-Eltern.....	104
Tabelle 26 Vergleich der Mittelwerte von mAGBS-Jugendliche, mAGBS-Eltern und AGBS.....	109
Tabelle 27 - Gesprächsanteile nach Videogruppen aufgeschlüsselt.....	139
Tabelle 28 Fragestil nach Videogruppen .....	140



## 9.2 Abbildungsverzeichnis

Abbildung 1 Übersicht zu Evaluationsmöglichkeiten des Arzt-Patient-Gesprächs .....	12
Abbildung 2 Themenblöcke von „Breaking bad news assessment scale“(BAS) und „Aufklärungsgesprächsbewertungsskala“(AGBS).....	20
Abbildung 3 Entwicklung der Bewertungsskala .....	35
Abbildung 4 Definition von Empathie.....	41
Abbildung 5 Entwicklungsschritte zur mAGBS-Jugendliche.....	42
Abbildung 6 Umfang der Raterschulung .....	46
Abbildung 7 Beurteilungszeit im Verlauf .....	52
Abbildung 8 Item 1 .....	53
Abbildung 9 Item 2 .....	53
Abbildung 10 Item 3 .....	54
Abbildung 11 Item 4 .....	54
Abbildung 12 Item 5 .....	55
Abbildung 13 Item 6 .....	55
Abbildung 14 Item 7 .....	56
Abbildung 15 Item 8 .....	57
Abbildung 16 Item 9 .....	57
Abbildung 17 Item 10 .....	58
Abbildung 18 Item 11 .....	58
Abbildung 19 Item 12 .....	59
Abbildung 20 Item 13 .....	59
Abbildung 21 Item 14 .....	60
Abbildung 22 Item 15 .....	60
Abbildung 23 Item 16 .....	61
Abbildung 24 Item 17 .....	61
Abbildung 25 Item 18 .....	62
Abbildung 26 Item 19 .....	62
Abbildung 27 Item 20 .....	63
Abbildung 28 Item 21 .....	63
Abbildung 29 Item 22 .....	64
Abbildung 30 Item 23 .....	65
Abbildung 31 Rangreihe über einzelne Items .....	68
Abbildung 32 Bewertung der einzelnen Videos .....	70
Abbildung 33 Notendurchschnitt Gesamtwertung.....	71

Abbildung 34 Korrelation Durchschnitt Gesamt und globale Bewertung .....74  
Abbildung 35 Empathie in der mAGBS-Jugendliche und in der mAGBS-Eltern ..... 118

### 9.3 Formelverzeichnis

Formel 1 Crohnbachs Alpha .....47  
Formel 2 Intraclasscorrelationskoeffizient.....48  
Formel 3 Effektstärke .....50

### 9.4 Abkürzungsverzeichnis

- Aufklärungsgesprächsbewertungsskala **AGBS**
- Breaking bad news assessment schedule **BAS**
- Calgary Cambridge Observation Guides **CCOG**
- Mittelwert **M**
- Rater **R**
- Revised Maastricht History Taking and advice checklist **MAAS-R**
- Roter Interaction Analysis System **RIAS**
- Standardabweichung **Std**
- Structured communication adolescent guide **SCAG**

## **10 Danksagung**

Bedanken möchte ich mich in erster Linie bei meinem Doktorvater und Betreuer Prof. Dr. Reiner Frank für die freundliche Übertragung des Themas und die überaus engagierte Betreuung in allen Phasen der Arbeit. Durch seine Begeisterung, Motivation und unermüdliche Unterstützung ermöglichte er eine rasche und gründliche Bearbeitung des Themas.

Außerdem möchte ich mich herzlich bei Sabine Klemz, Diplom- Sozialpädagogin für die Auswertung der Videos bedanken.

Mein Dank gilt auch meinem Doktorarbeitskollegen Robert Braun für die gute und konstruktive Zusammenarbeit.

Bedanken möchte ich mich auch bei allen Mitarbeitern der Ambulanz für Kinder- und Jugendpsychiatrie für die nette Aufnahme, angenehme Arbeitsatmosphäre und nicht zu vergessen die Teerunden.

Ebenfalls möchte ich mich bei meinem Freund Markus Thomann fürs Korrekturlesen und die Unterstützung im „Kampf mit Word“ bedanken.