



MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Beobachtungssystem INTAKT:
Vergleich zwischen Manual und
Beobachtungstraining anhand
der Kategorie Joint Attention“

verfasst von / submitted by

Theresa Pilshofer BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 289

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Psychologie

Betreut von / Supervisor:

Ass.-Prof. Dr. Ursula Kastner-Koller

Danksagung

An dieser Stelle möchte ich mich bei den Menschen bedanken, die mich während der Entstehung dieser Arbeit unterstützt haben.

Meinen wissenschaftlichen Dank möchte ich Frau Ass.-Prof. Dr. Kastner-Koller und Frau Ass.-Prof. Dr. Deimann aussprechen, für die äußerst fachkundige und verständnisvolle Betreuung, sowie für die Möglichkeit, die Masterarbeit in Zusammenhang mit dem INTAKT-Beobachtungsverfahren zu verfassen.

Einen Herzlicher Dank gilt Frau Mag. Dr. Hirschmann, für ihre fachliche und methodische Unterstützung und für die Einschulung in das Beobachtungssystem INTAKT.

Weiters möchte ich mich bei Henriette Schulz bedanken, für die konstruktiven Gespräche und die freundliche und unkomplizierte Zusammenarbeit im Zuge der Masterarbeit.

Aus tiefstem Herzen möchte ich mich bei meiner langjährigen Freundin Christine bedanken, für ihre unglaubliche Unterstützung beim Korrekturlesen und vor allem dafür, dass sie immer ein offenes Ohr für mich hatte.

Besonders möchte ich mich auch bei meinen Geschwistern Benedikt, Simon und Franziska, für aufmunternde Gespräche und für viele lustige Momente zusammen bedanken. Ein spezieller Dank gilt nochmals Benedikt, für die Hilfe beim Korrekturlesen.

Mein tiefster Dank gebührt meinen Eltern, Veronika und Friedrich, die mich während des gesamten Studiums unterstützt und bestärkt haben und mir auch im Rahmen dieser Arbeit eine außerordentliche Stütze waren.

DANKE!

2	Übereinstimmung zwischen BeobachterInnen	24
2.1	Interrater-Reliabilität	24
2.1.1	Prozentuale Übereinstimmung.....	25
2.1.2	Kappa-Koeffizient.....	26
2.1.3	Intraklassenkorrelation	28
2.1.4	Inkongruente Urteile in der Interrater-Reliabilität.....	29
2.2	Intra-Rater-Reliabilität.....	30
2.3	Beobachtungsfehler und -Verzerrungen.....	31
2.3.1	Fehler zu Lasten der Beobachtenden	33
2.3.2	Fehler zu Lasten der Beobachtungssituation	36
2.3.3	Fehler zu Lasten des Beobachtungssystems.....	37
2.3.4	Vermeidung von Beobachtungsfehlern	37
2.4	Vorgehen Zur Verbesserung der BeobachterInnenübereinstimmung.....	38
2.4.1	Beobachtungsmニュアル	39
2.4.2	BeobachterInnentraining.....	40
2.4.3	Trainingsansätze.....	41
2.5	Zusammenfassung	44

II. Empirischer Teil

3	Hintergrund & Zielsetzung der Untersuchung	46
3.1	Fragestellung.....	47
4	Untersuchungsinstrument: INTAKT	48
4.1	Beobachtungssituation	48
4.1.1	Die Bastelsituation	49
4.1.2	Die freie Spielsituation	49
4.2	Beobachtungskategorien.....	49
4.2.1	Feinfühligkeit.....	50
4.2.2	Rückmeldung	50

4.2.3	Joint Attention	51
5	Methode	54
5.1	Durchführung der Erhebung	54
5.1.1	INTAKT- Beobachtungstraining	56
5.1.2	Ablauf zur Auswertung der Daten	56
5.1.3	Berechnung der BeobachterInnenübereinstimmung	57
6	Stichprobe	59
7	Ergebnisse	60
7.1	Berechnung der Intra-Rater- und Interrater Reliabilität.....	60
7.1.1	Intra-Rater-Reliabilität	60
7.1.1.1	Übereinstimmung: Manual.....	61
7.1.1.2	Übereinstimmung: Manual und Beobachtungstraining	65
7.1.2	Interrater-Reliabilität.....	69
7.1.2.1	Übereinstimmung: Beobachtungseinschulung	69
7.1.2.2	Übereinstimmung: Manual und Beobachtungseinschulung.....	72
7.1.2.3	Die zusammengefasste Prozentuale Übereinstimmung	76
8	Diskussion	79
	Abstract Deutsch.....	88
	Abstract English	90
	Literaturverzeichnis	92
	Tabellenverzeichnis	99
	Abbildungsverzeichnis	100
	Abkürzungsverzeichnis	100
	Anhangsverzeichnis	101
	Anhang	102

EINLEITUNG

Diese Arbeit befasst sich mit der Einschulung von BeobachterInnen im Rahmen des Videobeobachtungsinstruments INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2017), welches für die Erhebung der Interaktionsqualität zwischen Müttern und ihren Kindern am Institut für Angewandte Psychologie, Entwicklung und Förderung konzipiert wurde. Mithilfe der Kategorien *Feinfühligkeit*, *Rückmeldung* und *Joint Attention* wird das Verhalten der Bezugsperson während einer Bastel- und Spielsituation beurteilt. Ziel des Verfahrens ist eine reliable, objektive und valide Abbildung des Interaktionsgeschehens zwischen Mutter und Kind und soll für entwicklungsdiagnostische Forschung und im diagnostischen Prozess eingesetzt werden (Hirschmann, Kastner-Koller, Deimann, Aigner & Svecz, 2011).

Im Zuge der stetigen Weiterentwicklung des Beobachtungsinstruments werden die Gütekriterien kontinuierlich überprüft und auf weitere Verbesserungsmöglichkeiten hin analysiert. Die Arbeit soll in diesem Sinne einen Beitrag zur Optimierung der Objektivität und Reliabilität des Verfahrens leisten. Anhand des Kategoriensystems *Joint Attention* wird die Vergleichbarkeit zwischen den Datenerhebungen mithilfe des INTAKT-Manuals und den Erhebungen mit zusätzlichem Beobachtungstraining untersucht. Ziel ist die Beantwortung der Frage, ob die Kodierung mit Manual allein ähnlich reliable und objektiv vergleichbare Ergebnisse bringt, wie eine Erhebung der Daten mit zusätzlichem Beobachtungstraining.

Zu diesem Zweck wurden Videos, welche im Rahmen von Diplomarbeiten von Huber (2013), Reischer (2013) und Vogler (2015) aufgenommen wurden, erneut kodiert. Eine zweite Erhebungsphase derselben Videos ermöglichte die Berechnung der Intra-Rater-Reliabilität. Die Mehrzahl der Videos wurde anhand der Vorgaben des Manuals erfasst, ein Teil wurde nach der Absolvierung des Beobachtungstrainings kodiert. Der Vergleich der erhobenen Daten mit den Ergebnissen einer zweiten Beurteilerin ermöglichte die Überprüfung der Interrater-Reliabilität.

Die Arbeit beginnt mit einer für die Beantwortung der Fragestellung theoretischen Auseinandersetzung inhaltlich relevanter Themengebiete. Dabei wird unter anderem auf die Beobachtung als wissenschaftliche Erhebungsmethode, insbesondere in Verbindung mit Videobeobachtung und relevanten Hintergründen zur Verhaltenskodierung

eingegangen. Des Weiteren folgt ein inhaltlicher Diskurs zur Thematik BeobachterInnenübereinstimmung und damit verknüpften Inhalten wie unterschiedliche Ansätze zur Durchführung von BeobachterInnentrainings.

I. THEORETISCHER TEIL

1 BEOBACHTUNG ALS WISSENSCHAFTLICHE ERHEBUNGSMETHODE

1.1 ALLGEMEINE BEGRIFFSDEFINITION

Unter dem Begriff *Verhaltensbeobachtung* wird der Vorgang verstanden durch „systematische und kontrollierte Herangehensweise offene oder verdeckte Verhaltensweisen einer Person, oder Interaktionsabläufe zwischen mehreren Personen, mithilfe konkreter Definitionen von Beobachtungskategorien und der Bestimmung von Beobachtungszeiträumen und Beobachtungssettings zu erfassen“ (Bodenmann, 2006, S.151).

In dieser Arbeit werden die Begriffe *Beobachtende* und *Beurteilende* sowie *KodierInnen* oder ähnliche weitgehend synonym verwendet. Zwar kann bei verschiedenen Erhebungsarten und Messungen eine Trennung in der Tätigkeit von beobachtenden und beurteilenden Personen erfolgen, im Zuge der Verhaltensbeobachtung jedoch sind jene Personen, welche eine Beobachtungsfunktion übernehmen zeitgleich auch diejenigen, durch welche die Beurteilung erfolgt. Die Begriffe *BeobachterInnentraining* und *BeobachterInneneinschulung* oder ähnliche, werden für die Beschreibung eines Trainings der beobachtenden Personen durch ein oder eine fachlich geschulte BeobachterIn eingesetzt.

Das erste Kapitel soll einen Überblick zur wissenschaftlichen Verhaltensbeobachtung geben. Zu diesem Zweck wird der Begriff der wissenschaftlichen Beobachtung erörtert und verschiedene Formen der Verhaltensbeobachtung werden dargestellt. Insbesondere wird auf die videobasierte Verhaltensbeobachtung mit Schwerpunkt auf den entwicklungspsychologischen Kontext näher eingegangen und diskutiert.

1.2 VORAUSSETZUNGEN UND FORMEN WISSENSCHAFTLICHER VERHALTENS- BEOBACHTUNG

Beobachtung wird als aktiver Prozess verstanden, wobei der Fokus der Aufmerksamkeit gezielt auf ein Geschehen gerichtet wird (Döring, Bortz & Pöschl, 2016). Die *wissenschaftliche Beobachtung* hat nachhaltig die Absicht Verhaltensmerkmale, Interaktionen und Ereignisse systematisch in der Situation unter Einbezug der Sinnesorgane oder mithilfe zusätzlicher technischer Dienste, welche einen erneuten Zugriff auf das Geschehen ermöglichen, zu erheben (Döring et al., 2016). Die Voraussetzungen für eine wissenschaftliche Beobachtung sind laut Seidel & Prenzel (2010) erfüllt, wenn ein bestimmter Forschungszweck verfolgt wird, infolge dessen die Beobachtung des Geschehens zweckgerichtet und systematisch geplant und auf einen theoretischen Bezugsrahmen aufgebaut ist. Zusätzlich ist eine schriftliche, auditive oder audiovisuelle Form der Aufzeichnung der Beobachtungssituation notwendig.

Einige psychologische Fragestellungen, z.B. in Bezug auf Interaktionsmuster und Bindungsverläufe, fordern die Beobachtung eines komplexen Prozesses, welcher sich nur unzureichend durch eine einmalige statische Erhebung erfassen lässt (Bakeman & Quera, 2011). In diesem Rahmen gilt in vielen Fällen die Verhaltensbeobachtung als die wissenschaftliche Methode der Wahl, um einen individuellen oder zwischenmenschlichen Prozess abzubilden.

Wissenschaftliche Beobachtungen ermöglichen einen differenzierten Einblick in Themenfelder, welche bei Befragungen stark subjektiv vom Individuum beeinflussbar sind, zum Beispiel bei negativ behafteten Verhaltensweisen wie Aggressivität (Rentzsch & Schütz, 2009). Trotz des Bewusstseins, dass das Verhalten einer Testperson in einer Beobachtungssituation aufgrund der Rahmenbedingungen mit hoher Wahrscheinlichkeit nicht „unbeeinflusst“ sein wird, gilt die Verhaltensbeobachtung als eine der am höchstangesehenen wissenschaftlichen Methode, Verhaltensabläufe und Interaktionsmuster auf „natürliche“ Weise zu erheben (Bakerman & Quera, 2011, S.6).

Vor Beginn einer wissenschaftlichen Verhaltensbeobachtung, welche qualitativ oder quantitativ durchgeführt werden kann, müssen viele Entscheidungen, welche die Erhebung der Daten betreffen, gegeneinander abgewogen werden. Ziel der Verhaltensbeobachtung ist der objektive Erkenntnisgewinn, in welchem beobachtbare Verhaltensweisen und deren Interpretation transparent gemacht und nachvollziehbar darge-

stellt werden können (Döring et al, 2016). Im Kontext einer wissenschaftlichen Beobachtung findet eine dokumentierte Erhebung der Daten statt, zusätzlich stützt sich der Erhebungsprozess auf eine im Vorhinein festgelegten Planung. Wichtige Begriffe angelehnt an Döring et al. (2016) im Rahmen einer wissenschaftlichen Beobachtungsstudie, deren Ziel „nachvollziehbare und objektive Erkenntnisse durch ein regelgeleitetes Vorgehen“ sind, sind *Perspektivität*, *Selektivität* und *Konstruktivität* (S. 328).

Perspektivität: Die Perspektivität der Beobachtung fragt nach dem Ziel und dem Ausgangspunkt der Forschung.

Selektivität: Die Selektivität meint die Begutachtung der Kriterien, die den Rahmen der Verhaltensbeobachtung definieren und den Fokus auf die relevanten Aspekte innerhalb der Beobachtung legt.

Konstruktivität: Wie Beobachtungen registriert, dokumentiert und interpretiert werden, wird im Rahmen des Begriffes Konstruktivität der Beobachtung offengelegt. Diese geplante Vorgangsweise zielt in der Beobachtung vorrangig darauf ab, vergleichbare Ergebnisse zu erheben und nach Möglichkeit Beobachtungsfehler zu verringern.

Den Forschenden stehen nach Schölmerich (2011) viele Möglichkeiten zur Verfügung eine Beobachtungssituation nach spezifischen Anforderungen zu gestalten. Zu Beginn gilt es Fragen zu klären, die den Beobachtungsstandort betreffen, sowie die Festlegung des Zeitpunktes, die Dauer der Beobachtung und die Anzahl der Beobachtungseinheiten, die für eine möglichst gute Erhebung des Verhaltens benötigt werden. Auch die geeignete Zielgruppe an Testpersonen ist ein Aspekt, welcher in Bezug auf das Studiendesign definiert werden muss. In der Konstruktionsphase einer Beobachtungssituation soll darauf Bezug genommen werden, welche Art von Untersuchungsgegenstand erfasst werden möchte.

Die Erhebung interpersoneller Eigenschaften, ergo zwischenmenschliche Interaktionen, benötigen andere Vorgehensweisen als die Erfassung von Personeneigenschaften (*traits*) oder Situationseigenschaften (*states*). Die Beobachtung von Interaktionen stellt eine besondere Herausforderung dar, denn eine Standardisierung des Kontextes ist oftmals mit zusätzlichen Herausforderungen verbunden, da die Personen im interaktiven Geschehen sich in ihrem Verhalten gegenseitig beeinflussen. Um eine isolierte Beobachtung der einzelnen Komponenten im Zuge des Interaktionsgeschehen zu ermöglichen, muss bereits bei der Kategorienerstellung ein genaues Wissen über das

insgesamt mögliche Verhalten, das auftreten kann, vorhanden sein. Auf dieser Basis kann ein geeignetes Vorgehen und damit einhergehendes statistisches Verfahren gewählt werden (Schölmerich, 2011).

Häufig wird zwischen folgenden Arten der Beobachtung, welche sich in der Praxis häufig überschneiden und ergänzen, unterschieden:

Systematische vs. unsystematische Beobachtung: Eine systematische Beobachtung unterscheidet sich von einer unsystematischen Beobachtung durch die hypothesengeleitete und methodenbasierte Erhebung der Daten (Faßnacht, 1997). Im nächsten Kapitel 1.3. wird die systematische Beobachtung näher erläutert.

Teilnehmende vs. nicht-teilnehmende Beobachtung: Die beobachtende Person kann bei der Explorationssituation anwesend sein und innerhalb des Geschehen selbst agieren, wie bei der teilnehmenden Beobachtung der Fall. Im Rahmen der nicht teilnehmenden ist die beurteilende Person nicht präsent. In solchen Fällen werden meist technische Hilfsmittel eingesetzt um die Situation aufzuzeichnen und später auszuwerten (Heikamp, 2014).

Offene vs. verdeckte Beobachtung: In einer offenen Beobachtung wissen alle anwesenden Personen, dass sie Teil einer Beobachtung sind, während in einer verdeckten Beobachtung, den TeilnehmerInnen nicht bewusst ist, dass sie Teil einer Beobachtungssituation sind (Schölmerich, 2011).

Labor vs. Feldbeobachtung: In einer natürlichen Beobachtungssituation (Feldbeobachtung) ist die Kontrolle über bestimmte Durchführungsaspekte und dem damit erwünschtem Verhalten niedriger als im Zuge einer Laborbeobachtung, bei der durch die künstliche Situation Umgebungsfaktoren um ein Vielfaches besser kontrolliert werden können. Die Auswirkungen der Beobachtungssituation auf den Standardisierungsgrad einer Studie sollten bei der Wahl der Beobachtungsart bedacht werden (Heikamp, 2014).

Technisch-vermittelte vs. unvermittelte Beobachtung: Die Datenerhebung wird in der technisch-vermittelnden Beobachtung mittels technischer Hilfsmittel, meist mittels Videoaufnahmen unterstützt und auf Speichermedien festgehalten. Bei unvermittelten Beobachtungen wird das Geschehen von der beobachtenden Person direkt vor Ort erfasst (Spinath & Becker, 2011).

1.3 DIE SYSTEMATISCHE VERHALTENSBEOBSACHTUNG

Das quantitative Beobachtungsdesign ermöglicht die Messung von Verhaltensmerkmalen in Form numerischer Daten anhand eines standardisierten Beobachtungsplans, insbesondere im Zusammenhang mit Untersuchungen großer Datenmengen (Döring et al., 2016).

Im Voraus definierte Aspekte des Geschehens werden in Form von Fremdbeobachtungen durch geschulte BeurteilerInnen in ihrer Häufigkeit, Dauer und Intensität ermittelt (Döring et al., 2016; Fisseni, 1997). Anhand Döring et al. (2016) lassen sich drei Hauptdimensionen zur Unterscheidung verschiedener quantitativer Beobachtungsformen bilden: *Quantitative Beobachtung mit geringem Komplexitätsgrad*, *Strukturierte Verhaltensbeobachtung* und *Nonreaktive Verhaltensbeobachtung* (S. 335).

Quantitative Beobachtung mit geringem Komplexitätsgrad: Die erste Dimension, die quantitative Beobachtung mit geringem Komplexitätsgrad, misst wenige und einfach strukturierte Merkmale.

Strukturierte (systematische) Verhaltensbeobachtung: Die zweite Möglichkeit, jene der strukturierten Verhaltensbeobachtung, erfasst ebenso Verhaltensweisen und Merkmale in ihrer Dauer, Intensität und Häufigkeit, jedoch auf komplexerer Ebene, weshalb meist auch aufwendigere Messinstrumente verwendet werden. Im nachfolgenden Absatz wird auf die Strukturierte Verhaltensbeobachtung näher eingegangen.

Nonreaktive Verhaltensbeobachtung: Die Datenerhebung der letzten Dimension, der nonreaktive Verhaltensbeobachtung, erfolgt anhand materieller Güter oder virtueller Medien.

Die *Strukturierte Verhaltensbeobachtung* wird in der Literatur oft auch als systematische oder kontrollierte Beobachtung bezeichnet. Im Rahmen der *Systematischen Beobachtung* sind die Beobachtungsmerkmale, welche erhoben werden sollen, klar festgelegt und definiert. Die systematische Beobachtung unterscheidet sich von der freien Beobachtung durch ein hypothesengeleitetes und methodenbasiertes Vorgehen, die Frage nach der Art der Erhebung, der Fixierung und der Auswertung der Daten ist ebenfalls geklärt (Faßnacht, 1997; Fisseni, 1997).

Die Systematische Verhaltensbeobachtung baut auf einzelne Beobachtungseinheiten auf, welche kleinstmögliche, nicht mehr reduzierbare Verhaltensweisen und Interaktionsprozesse erfassen, die im Anschluss zusammenfassend für eine Gesamtanalyse

des Verhaltens ausschlaggeben sind (Heikamp, 2014). Somit lassen sich Gruppenabläufe, sowie individuelle Ausdrucksweisen gut erforschen.

Der Schlüssel für eine erfolgreiche strukturierte Beobachtung liegt im Beobachtungsinstrument, welches objektiv, valide und reliable die zu beobachtenden Verhaltensweisen und Merkmale in ihrer Komplexität erfassen soll. Idealerweise wird die Erhebung unter Verwendung von Messinstrumenten durchgeführt, die die Kontrolliertheit der Beobachtung garantieren, Möglichkeit zur Selbstreflexion bieten und dabei behilflich sind die Möglichkeiten des menschlichen Wahrnehmens zu erweitern (Latz, 1993). Eine standardisierte Erhebung der Daten ist Voraussetzung um unabhängige Vergleiche zu ermöglichen, weshalb ein hohes Maß an Planung, ein umfangreiches Beobachtungstraining und eine regelmäßige Überprüfung der Gütekriterien gefordert wird (Döring et al., 2016).

Die Strukturierte Verhaltensbeobachtung ist vor allem im Zuge der klinischen, entwicklungspsychologischen und pädagogischen Forschung und Diagnostik von Bedeutung. In den meisten Fällen kommen dabei externe BeobachterInnen zum Einsatz, die für die Durchführung der Beobachtung eingeschult wurden. Unter Umständen kann auch die systematische Selbstbeobachtung, welche das Protokollieren des eigenen Verhaltens zum Ziel hat, in einer Untersuchung die Methode der Wahl sein (Heikamp, 2014).

1.4 VIDEOGESTÜTZTE VERHALTENSBEOBACHTUNG

Die Kodierung des Verhaltens kann unmittelbar während der Verhaltensbeobachtung oder unter Einsatz von Aufnahmemöglichkeiten wie Audio- oder Videoverfahren nach Abschluss der Beobachtung erfolgen. Die vollständige Erhebung eines Geschehens zum Zeitpunkt ihres Auftretens, bzw. die Erfassung mehrerer Teilaspekte davon, überfordert aufgrund der Dichte und Schnelligkeit der Situation in vielen Fällen selbst geübte BeobachterInnen. Audio- und/ oder Videoaufnahmen lassen zu einem späteren Zeitpunkt ein wiederholtes Ansehen des erhobenen Materials in unterschiedlichen Geschwindigkeiten zu und ermöglichen einen differenzierten und reflektierten Blick auf das Verhalten (Döring et al., 2016; Papousek, 2000; Thiel, 2011).

1.4.1 Techniken videogestützter Verhaltensbeobachtung

Gestützt durch die Techniken der digitalen Medien bietet die videobasierte Verhaltensbeobachtung eine Vielzahl an neuen Möglichkeiten für die Forschenden das menschliche Verhalten in verschiedenen Aspekten zu erkunden. Die Aufnahme der Beobachtungssituation auf Kamera ermöglicht eine Fixierung dynamischer, temporärer Prozesse, wie zwischenmenschliche Interaktionen, welche sich durch ihre Flüchtigkeit auszeichnen. Durch die Videoaufnahme wird ein wiederholtes Aufrufen des Geschehens im Sinne eines „ikonisch äquivalenten Abbildes“ zu einem beliebigen späteren Zeitpunkt ermöglicht (Thiel, 2003, S. 657). Somit können nicht-replizierbare Untersuchungsgegenstände vorurteilsfrei erfasst werden und im Nachhinein mithilfe objektiver Kriterien analysiert und informationsbringend genutzt zu werden. Gleichzeitig ist eine Differenzierung der eigenen subjektiv vorhandenen Sichtweise möglich, welche eine weniger durch die individuelle Wahrnehmung geprägte Auswertung der Beobachtungssituation, im Vergleich zur Verhaltensbeobachtung ohne Videoaufnahmen, ermöglicht (Thiel, 2003; Thiel, 2011).

Ein wichtiger Vorteil, den die videogestützte Verhaltensbeobachtung im Gegensatz zur unmittelbaren Beobachtung bietet, ist die Tatsache, dass aufgrund der Möglichkeit der Wiederholung des Materials die BeobachterInnen sich auf verschiedene Aspekte des Verhaltens oder der Interaktion konzentrieren und im nächsten Kodierdurchgang einen anderen Aspekt in das Zentrum der Aufmerksamkeit rücken können (Bakeman & Quera, 2011). Die zyklische Analyse, das wiederholte Bearbeiten von Teilschritten der Videoaufnahme, ist dabei eine unterstützende Form ein Maximum an Information zu erhalten (Jacob, 2016). Die Mikroanalyse ist eine weitere Methode, in welcher durch eine iterative Vorgangsweise die Datenbearbeitung in der videogestützten Verhaltensbeobachtung erleichtert wird. Die Herangehensweise von groben zu feinen Sequenzen hilft bei der Fokussierung auf spezifische Bereiche des Materials und verringert gleichzeitig die Komplexität der aufgezeichneten Situation. Zusätzlich eignet sich ein Zeitstichprobenplan, um „blinde Flecken“ offenzulegen (König, 2013, S.826).

Die Unterstützung durch technische Mittel eröffnet auch im zeitlichen Kontext neue Möglichkeiten Merkmale eines Verhaltensablaufs oder Interaktionsprozesses zu betrachten. Die Zeitlupentechnik lässt zu, dass Details im Bereich von Millisekunden erfasst werden können; die Zeitraffertechnik unterstützt bei der Beobachtung langsamer Veränderungen innerhalb eines lang dauernden Zeitrahmens und die kontrastierende

Bildmontage kann eingesetzt werden, um langfristige Prozesse, z.B. in der Entwicklung einer Person, sichtbar zu machen (Papousek, 2000).

1.4.2 Vorteile und Nachteile technisch-vermittelter Beobachtung

Die videogestützte Verhaltensanalyse kann in vielerlei Hinsicht für die forschenden Personen Erleichterung in der Durchführung und der Erhebung der Daten bringen. Die Aufzeichnung des Materials kann eine Chance für die forschenden Personen darstellen, Beobachtungssituationen im Team zu diskutieren und bei Bedarf die Forschungsfrage in Wechselwirkung mit dem beobachteten Geschehen zu präzisieren (König, 2010).

Videogestützte Verhaltensbeobachtung ermöglicht zudem eine Kontrolle der Intra-Rater-Reliabilität, die Übereinstimmung der Ergebnisse einer beobachtenden Person, welche sich in einer Verhaltensbeobachtung ohne Aufzeichnung als schwierig erweist. Durch die Möglichkeit der Wiederholung des Materials kann die Kodierung mit der vorhergehenden Kodierung derselben Person verglichen werden (Bakeman & Quera, 2011). Weiters können Beobachtungsfehler anhand videobasierter Aufnahmen für Schulungszwecke demonstriert und am „konkreten Fall“ besprochen und korrigiert werden (Fisseni, 1997, S.202). Die videogestützte Verhaltensbeobachtung ermöglicht laut König (2013) zudem eine bessere Vernetzung zwischen quantitativem und qualitativem Forschungszugang.

Nicht zu unterschätzen ist der mit dem Einsatz videobasierter Methoden einhergehende ökonomische Aufwand, welcher aufgrund der Trennung von Erhebung und Auswertung des Materials zu bedenken ist, sowie eine oftmals damit einhergehende aufwendige computergestützte Kodierung des Materials (König, 2013). Des Weiteren darf der damit einhergehende Kostenfaktor nicht unerwähnt bleiben (Döring et al., 2016). Die videogestützte Beobachtung wird jedoch aufgrund der rasanten Weiterentwicklung der Technologie in der heutigen Zeit finanziell leistbarer und in der Forschung immer mehr zur Routine, selbst wenn damit weitere Herausforderungen für die Forschenden verbunden sind, da technisches Verständnis und eine korrekte Handhabung vorausgesetzt werden (Bakeman & Quera, 2011).

Ungeachtet dessen stellt sich die Frage, inwieweit die aufgezeichnete Beobachtung einer Forschungssituation der tatsächlichen Realität nahekommen (Thiel, 2011). Obwohl viele Versuchspersonen im Laufe der Verhaltensbeobachtung die Kamera nicht mehr aktiv wahrnehmen und eine Videokamera auch ein Distanzverhältnis zwischen

den beobachtenden und den beobachteten Personen herstellen kann, ist die Problematik der sozialen Erwünschtheit und der damit verbundenen Selbstinszenierung ein kontrovers diskutierter Aspekt in der Verhaltensbeobachtung (Dinkelaker & Herrle, 2009; König, 2013). Thiel (2011) betont hierbei, dass es sich dabei nicht um eine spezifische Herausforderung der videobasierten Beobachtung handelt, sondern, dass Hindernisse wie diese generell in technisch-vermittelten, wie in unvermittelten Beobachtungssituationen vorzufinden sind.

In einigen Beobachtungssituationen sollte aufgrund ethischer Aspekte von Videoaufnahmen abgesehen werden, wenn permanente Aufnahmen als zu intensiver Eingriff die die Privatsphäre wahrgenommen werden (Bakeman & Quera, 2011).

Zusammenfassend sollte immer mitbedacht werden, dass nicht alle inneren Vorgänge eines Menschen mittels Fremdbeobachtung zugänglich sind und es klare Grenzen der Beobachtbarkeit gibt, welche den Einsatz zusätzlicher Erhebungsformen erfordern (Döring et al, 2016,). Als Beispiel dienen individuelle Fähigkeiten, Motive und Emotionen, welche Konstrukte sind, die sich nur durch die Beobachtungen von konkreten Verhaltensäußerungen erschließen lassen, jedoch nicht direkt beobachtbar sind (Fisseni, 1997).

1.4.3 Videogestützte Verhaltensbeobachtung in der Entwicklungspsychologie

Verhaltensbeobachtungen stellen eine Alternative zu anderen Erhebungsformen dar, um indirekt intuitives oder automatisiertes Verhalten bei Personen zu erfassen (Bakeman & Quera, 2011; Papousek, 2000; Thiel-Bonney, 2002; Thiel, 2011).

Die videogestützte Verhaltensbeobachtung öffnet den Weg zu einer direkten empirischen Erfassung intuitiver Verhaltensformen, welche auf automatisierten Abläufen beruhen und kognitiv unbewusst ausgeführt werden. Im Säuglings- und Kindesalter sind Erlernen motorischer Aspekte (z.B. Greifen) sowie selbstregulatorische und integrative Kompetenzen (z.B. Selbstberuhigung oder Nachahmung und Spiel) Bestandteile dieser Verhaltensabläufe (Papousek, 2000).

Es gibt mehrere Gründe, weshalb die videogestützte Verhaltensbeobachtung insbesondere im Rahmen der Entwicklungspsychologie eine essentielle Rolle in der Erhebung von Verhalten einnimmt. Zum einen sind die Kommunikationsmöglichkeiten in der Kindheit oftmals begrenzt. Die sprachlichen Fähigkeiten ermöglichen im Kleinkindesalter keine adäquate Aussage über innere Gemütszustände. Im Gegensatz dazu

zeigen Kinder ihre Emotionen und Empfindungen besonders unverfälscht in ihrem Verhalten. Vor allem Säuglinge und Kleinkinder äußern ihr Befinden vorrangig durch flüchtiges intuitives Verhalten, dessen Erhebung durch die videogestützte Verhaltensbeobachtung aufgrund der technischen Möglichkeiten neue Perspektiven bezüglich psychologischer und diagnostischer Erkenntnissen ermöglicht (Papousek, 2000; Thiel, 2011; Thiel-Bonney, 2002). In Bezug auf entwicklungspsychologische Fragestellungen hat der Einsatz videogestützter Verhaltensbeobachtungen durch die neuen Techniken der Bild zu Bild-Analysen oder Zeitlupenbetrachtungen in den letzten Jahrzehnten maßgeblich zu Erkenntnissen der frühkindlichen Entwicklung beigetragen (Papousek, 2000).

Darüber hinaus sind es im speziellen „dynamische Prozesse“ und „interaktive Phänomene“, welche in der Forschung mit Kindern von Bedeutung sind. „Dynamische Prozesse“ meint eine Veränderung über eine Zeitdauer hinweg, während „interaktive Phänomene“ das Verhalten beschreibt, welches die Grundlage für das Verhalten einer anderen Person darstellt (Schölmerich, 2011, S.769). Insbesondere in der Bindungsforschung wird daher der Einsatz videobasierter Verhaltensbeobachtung genutzt um detaillierte Informationen zum Interaktionsablauf zwischen Kind und Bezugsperson zu generieren. Der Einsatz von Videomaterial gestattet den BeobachterInnen sich abwechselnd auf die kindlichen Signale und das Verhalten der Bezugsperson zu konzentrieren, auch unter zeitlicher Abstimmung des Interaktionsverhaltens (Papousek, 2000). Die Vorteile der Videoaufnahmen lassen sich insbesondere für die Beobachtung von Eltern-Kind-Interaktionen nutzen, bei welchen ebenso meist unbewusst ablaufende Interaktionsmuster und -strukturen im Fokus der Erhebung stehen (Thiel-Bonny, 2002). Unter anderem eignen sich dafür Spielsituationen sehr gut, um das kindliche Verhalten oder Interaktionsabläufe zu beobachten (Döring et al., 2016). Darauf aufbauend können Videofeedback und Frühförderprogramme gestaltet werden, welche in der Praxis zu Präventionszwecken sowie in therapeutischen Settings und zur weiterführenden Beratung eingesetzt werden (Thiel, 2011).

Für die Beantwortung komplexer Abläufe von Entwicklungsschritten oder Eltern-Kind Interaktionen sowie andere entwicklungspsychologische Fragestellungen, ist der Einsatz technischer Geräte unabdinglich geworden und zusätzliche Hilfsmittel, wie z.B. das Messen von physiologischen Veränderungen, Stimmfrequenzanalysen oder Videoaufnahmen werden gezielt in der Forschung sowie im Kontext von Kinder- und Jugenddiagnostik eingesetzt (Papousek, 2000).

1.5 GÜTEKRITERIEN

Eine standardisierte Beobachtung setzt die Einhaltung der Gütekriterien Objektivität, Reliabilität und Validität voraus (König, 2013). Für die Gültigkeit der Ergebnisse bei wissenschaftlichen Beobachtungen ist die Nachvollziehbarkeit der Durchführung für Dritte besonders relevant, weshalb eine präzise Dokumentation über die Verwendung der Beobachtungsverfahren und Messinstrumente, sowie die Interpretation und Deutung der Beobachtung selbst von großer Wichtigkeit ist (Seidel & Prenzel, 2010).

1.5.1 Objektivität

Die Objektivität in der Verhaltensbeobachtung ist gegeben, wenn das Ergebnis nicht von der beobachtenden Person oder vom Verhalten der Testperson abhängig ist (Ulber & Imhof, 2014). Das Gütekriterium kann in drei Aspekte gegliedert werden: *Durchführungs-*, *Auswertungs-*, und *Interpretationsobjektivität*, die anhand der Erklärungen von Moosbrugger & Kelava (2012) und Ulber & Imhof (2014) folgend kurz dargestellt werden:

Durchführungsobjektivität: Die *Durchführungsobjektivität* ist gegeben, wenn das Ergebnis nicht von den eingesetzten BeobachterInnen abhängig ist. Ein gleichbleiben der Randbedingungen (z.B. Dauer, Zeitpunkt und Häufigkeit der Erhebung) trägt zu einer Verbesserung der Durchführungsobjektivität bei.

Auswertungsobjektivität: Die Beurteilung des beobachtenden Geschehens soll nicht in Abhängigkeit zu der beurteilenden Person stehen. Klare Auswertungsregeln nach einem verbindlichen Manual können den Beurteilenden helfen, ein Verhalten richtig einzuordnen.

Interpretationsobjektivität: Interpretationsobjektivität liegt vor, wenn unterschiedliche BeobachterInnen bei der Beobachtung desselben Geschehens zu gleichen und Schlussfolgerungen kommen.

Die Objektivität wird im Rahmen der Verhaltensbeobachtung geprüft, indem eine Erhebung durch voneinander unabhängigen BeobachterInnen, unter Verwendung gleicher Methoden durchgeführt wird (Faßnacht, 1995). Fisseni (1997) beschreibt diesen Vorgang als die Schaffung eines „intersubjektiven Konsens über denselben Gegenstand“ (S.186). Die Interrater-Reliabilität, die Übereinstimmung zwischen mindestens zwei BeobachterInnen, ist somit ein Maß der Objektivität eines Beobachtungsinstruments.

Die Objektivität prüft die BeobachterInnenunabhängigkeit und ist abhängig von der Standardisierung des Messinstruments. Eine standardisierte Anwendung des Beobachtungsinstruments kann durch eine Beobachtungseinschulung in Form eines Trainings oder eines Manuals erreicht werden (Döring et al., 2010). Eine zufriedenstellende Objektivität kann durch eine „Schulung der Wahrnehmungsschärfe“, durch den Einsatz videogestützter Verhaltensbeobachtung sowie der Fokussierung auf Teilaspekte des Geschehens erzielt werden (Fisseni, 1997, S.207).

In der psychologischen Praxis wird in vielen Fällen für die Erhebung der Objektivität die Berechnung der BeobachterInnenübereinstimmung durch das statistische Maß Cohens Kappa, welches sich aus den relativen Häufigkeitsdaten zweier BeobachterInnen berechnet, eingesetzt (Faßnacht, 1995). In Kapitel 2.1.2 wird Cohens Kappa näher erläutert sowie einige weitere Übereinstimmungsmaße (Prozentuale Übereinstimmung, gewichtetes Cohens Kappa, Fleiss Kappa und Intraklassenkorrelation) dargestellt. Eine hohe Anzahl unabhängiger BeobachterInnen ist ein Gewinn für die Aussagekraft der Objektivität des Beobachtungsinventars, da Rückschlüsse auf eine intersubjektive Verwendung möglich sind (Faßnacht, 1995).

1.5.2 Reliabilität

Die Reliabilität gibt den Grad der Genauigkeit, beziehungsweise der Zuverlässigkeit einer wissenschaftlichen Messung oder Erhebung an. Bei der Erhebung von Verhalten im Rahmen von Beobachtungsstudien gibt der Grad der Übereinstimmung zwischen mindestens zwei BeobachterInnen neben Informationen über die Objektivität eines Inventars auch Auskunft über die Reliabilität einer Beobachtung. Diese kann mit denselben statistischen Maßen, wie bereits in der vorhergehenden Kategorie zu Objektivität erläutert, überprüft werden (Schölmerich, 2011).

Des Weiteren ist auch die Konsistenz einer einzelnen beobachtenden Person, genannt die Intra-Rater-Reliabilität, ein Indiz für die Zuverlässigkeit der Beurteilung des Verhaltens (Seidel & Prenzel, 2010). Bei videogestützten Beobachtungen kann diese, ähnlich wie bei Retests, durch eine wiederholte Erhebung desselben Materials berechnet werden. Die *Retest-Reliabilität* gibt die Übereinstimmung der Ergebnisse derselben Person zu zwei verschiedenen Messzeitpunkten an und erfasst die zeitliche Stabilität einer Testung (Fisseni, 1997, S. 76). Theoretisch ist die Reproduzierbarkeit der Daten mittels unterschiedlicher Messinstrumente eine essentielle Voraussetzung für die Gültigkeit einer wissenschaftlichen Studie, jedoch gestaltet es sich in der Praxis oft sehr

schwierig, situative und zeitlich variable Merkmale wiederholt zu erfassen (Seidel & Prenzel, 2010).

Mees (1977) unterscheidet im Kontext der Reliabilität zwischen *Stabilität* und *Zuverlässigkeit*. Als *Zuverlässigkeit* wird die Eigenschaft des Messinstruments beschrieben. *Stabilität* bezeichnet die Eigenschaft eines Merkmales und meint den Grad mit dem ein Verhaltensmerkmal über verschiedene Beobachtungszeitpunkte hinweg gezeigt werden. Ein geringer Stabilitätswert in einer Beobachtung ohne Videounterstützung macht somit laut Mees (1997) keine Angaben über die Zuverlässigkeit der Beobachtung, da es damit zusammenhängen kann, dass das erhobene Merkmal an sich nicht stabil in seinem Auftreten ist.

Fisseni (1997) erläutert einige Faktoren, welche eine positive Auswirkung auf die Reliabilität in Verhaltensbeobachtungen haben. Zum einen kann eine Präzisierung der Beobachtungseinheiten und eine geringe Anzahl an Beobachtungskategorien zu einer hohen Reliabilität des Verfahrens beitragen. Je präziser die Formulierung der Einheiten, desto weniger subjektive Fehler werden in der Interpretation durch die BeobachterInnen auftreten und desto ähnlicher werden die Ergebnisse ausfallen. Insgesamt empfiehlt Fisseni (1997) für eine hohe Reliabilität mindestens zwei bis vier unabhängige BeobachterInnen einzusetzen, die das beobachtete Verhalten in einer geringen Anzahl von Beobachtungseinheiten protokollieren.

1.5.3 Validität

Die Validität fragt nach der Gültigkeit des Beobachtungsinstruments, indem sie in Bezug auf theoretische Modelle, Operationalisierungen und wissenschaftliche Definitionen untersucht, ob sich das untersuchte Verhalten mit dem interessierten Merkmal inhaltlich deckt (Döring et al., 2016). Die Validität einer Beobachtungsstudie ist hoch, wenn das Beobachtungsinstrument in einem hohen Grad an Genauigkeit die Merkmale misst, welche zu messen beabsichtigt wurden (Mees, 1977). Es werden drei Arten der Validität unterschieden, auf welche folgend auch näher eingegangen wird: *Inhaltsvalidität*, *Kriteriumsvalidität* und *Konstruktvalidität*. Die Definitionen orientieren sich an Mees (1977).

Inhaltsvalidität: Die *Inhaltsvalidität* gibt Auskunft darüber, inwiefern Rückschlüsse vom kodierten Verhalten der Beobachtungssituation auf das gleiche Verhalten außerhalb des Forschungsfeldes gezogen werden können. Somit bezieht sich die Inhaltsvalidität auf die „Generalisierbarkeit des Verhaltens über Setting und Zeitpunkte hinweg“

(Mees, 1977, S.89). Um eine Generalisierbarkeit und somit eine Repräsentativität des Verhaltens zu erlangen, muss das Verhaltensspektrum in seiner Gesamtheit anhand der Kodierregeln und des Aufbaus und der Zusammensetzung der Skalen und Kategorien des Beobachtungsinstruments erhebbar sein. Ziel ist Anpassung der Erhebungsstruktur an die Komplexität, Intensität und Geschwindigkeit des Verhaltens um es in seiner Flüchtigkeit bestmöglich erfassen zu können.

Weiters beschäftigt sich die *Inhaltsvalidität* mit der Frage, ob Verhaltenskategorien, welche verschiedene Verhaltensmerkmale zusammenschließen, in ihrer Zusammensetzung das erfassen, was sie laut Definition zu erfassen vorgeben. Mees (1977) zufolge sind ExpertInneneinschätzungen und eine intensive Auseinandersetzung mit dem interessierenden Verhalten essentiell zur Schätzung der Inhaltsvalidität eines Merkmals, da eine Berechnung der Inhaltsvalidität nicht möglich ist.

Kriteriumsvalidität: Die *Kriteriumsvalidität* hat, wie die Inhaltsvalidität, die Absicht, vom beobachtbaren Verhalten auf das Verhalten außerhalb der Beobachtungssituation schließen zu können. Der wesentliche Unterschied zur Inhaltsvalidität ist der Versuch dies über empirische Zusammenhänge, z. B. Korrelationen, zu berechnen. Dazu werden meist Beobachtungsmaße in Zusammenhang mit adäquaten Außenkriterien (z.B. Fragebögen, Tests etc.) gebracht und verglichen. Zusätzlich muss sich die forschende Person der Frage stellen, ob die Erhebung mittels Test oder Fragebogen dieselbe Verhaltensweise erfasst wie die Beobachtung. Ziel der Kriteriumsvalidität ist ein sinnvoller Vergleich der Ergebnisse, auch unter Berücksichtigung der Tatsache, dass das interessierende Verhalten sich zwischen den Zeitpunkten der unterschiedlichen Messerhebungen verändern kann (Mees, 1977).

Konstruktvalidität: Die *Konstruktvalidität* macht Aussagen zu der Gültigkeit des Verhaltens in der Beobachtungssituation in Verbindung mit theoretischen Konstrukten (z.B. Fähigkeiten, Charakter etc.), die diesem Verhalten zugrunde liegen (Mees 1977, S. 93). Anhand wissenschaftlicher Theorien soll mithilfe Beobachtungen überprüft werden, wie zwei Variablen zusammenhängen.

1.6 METHODEN DER QUANTIFIZIERUNG

Eine Frage, welche sich die Forschenden vorab stellen müssen, ist jene der Strukturierung des Beobachtungssystems in Anbetracht der Art und Auftrittshäufigkeit des interessierenden Verhaltensmerkmals. Basis der einzelnen Beobachtungsdimensionen des Inventars sind stets theoretische Konstrukte oder psychologische Modelle, anhand derer das Verhaltensmerkmal eines Individuums zu einer passenden Kategorie zugeordnet werden kann (Döring et al., 2016).

1.6.1 Kodierschemata und Möglichkeiten zur Messung von Beobachtungen

Kodierschemata reflektieren den wissenschaftlichen Zugang der Forschenden, die zugrundeliegenden theoretischen Modelle und das Forschungsziel einer systematischen Beobachtung. Es gibt zwei Arten von Kodierschemata, jene, die sich gegenseitig ausschließen und welche, die sich nicht gegenseitig ausschließen, sondern eine parallele Erhebung verschiedener Merkmale ermöglichen. Kodierschemata können ein Merkmal oder ein Verhalten erschöpfend oder nicht erschöpfend erfassen (Bakeman & Quera, 2011).

Ein weiterer wichtiger Punkt, welchen Bakeman & Quera (2011) hinsichtlich Datenerhebung und Datenanalyse betonen, ist die Bedeutung der Struktur und der Organisation der Kodierschemata. Darunter fällt zum Beispiel die Anzahl an vorhandenen Kodiereinheiten. Bei Videobeobachtungen empfehlen sie eine höhere Anzahl an Kodiereinheiten, da die Wiederholung der Videoaufnahme eine differenziertere Kodierung in mehreren Durchgängen ermöglicht.

Eine weitere Herausforderung stellt die Erhebung des Verhaltens mittels der entwickelten Kodierschemata und die damit verbundene Datentransformation mittels Paper-Pencil-Verfahren oder computergestützter Auswertungssoftware dar. Bakeman & Quera (2011) unterscheiden zwischen zwei Arten Verhalten zu erheben, auf welche im Folgenden näher eingegangen werden: *Event-Recording* und *Intervall-Recording/Time-Sampling*, während Faßnacht (1995) eine Einteilung in drei Verfahren zur Quantifizierung psychologischer Merkmale vornimmt: *Event-Sampling*, *Time-Sampling* und *Rating-Verfahren*, wobei die beiden ersten Methoden zur Erfassung von Häufigkeiten und Zeiten genutzt werden, während das Rating-Verfahren eingesetzt wird, wenn Intensität und Ganzheit der Inhalte zu messen beabsichtigt werden. Die folgende

Erläuterung der Erhebungsmethoden in der videogestützten Beobachtung bezieht sich vorrangig auf die Strukturierung von Bakeman & Quera (2011).

1.6.2 Event-Recording

Die BeobachterInnen protokollieren das Auftreten eines bestimmten Events. Faßnacht (1995) versteht unter dem Begriff des *Event-Sampling* Verfahren „das kontinuierliche, möglichst genaue, zeitliche Erfassen von mindestens einer aber meistens mehrerer Verhaltensweisen“ (S. 127). Bakeman und Quera (2011) unternehmen eine Einteilung in zwei Arten des Event-Recording:

Untimed-Event-Recording: Das Auftreten des gesuchten Events wird kodiert, die Zeitdauer des Events wird beim Auftreten jedoch nicht mitdokumentiert. Die Herausforderung für die beurteilenden Personen besteht vorrangig darin, stets aufmerksam zu sein und alle Events wahrzunehmen. Ein großer Vorteil dieser Kodierweise besteht darin, dass sie einfach durchzuführen ist, unkompliziert mittels der Paper-Pencil Technik angewendet werden kann und sich als kostengünstig erweist. Der Nachteil dieser Erhebungsform ist die fehlende Information über der Dauer eines Events, was zu einem wesentlichen Informationsverlust führen kann (Bakeman & Quera, 2011).

Timed-Event-Recording: Die Zeitdauer des Events wird beim Auftreten des gesuchten Verhaltens mitkodiert. Die Dauer des Events wird mit dem Notieren der Beginn- und Endzeitpunkte beim Auftreten des gesuchten Merkmals durch die Beobachtenden erfasst. Diese Kodierart wird unter anderem eingesetzt, wenn die Forschenden an der prozentualen Dauer eines Events während der gesamten Erhebung interessiert sind. Bakeman und Quera (2011) empfehlen diese Art der Datenerhebung im Rahmen von Beobachtungsstudien, da sie eine größere Informationsdichte und damit verbundene breitere Auswahl an Analysemöglichkeiten mit sich bringt. Die Methode hat jedoch auch Nachteile, da die zusätzliche Zeitkomponente bei der Erhebung von Events die Arbeit für die BeurteilerInnen erschwert. Zusätzlich werden komplexe Datenerhebungsinstrumente benötigt, meist im Rahmen von Auswertungsprogramme am PC, die zudem oft erheblich teurer in der Anwendung sind.

Aus den Erhebungsdaten mittels *Event-Recording* können verschiedene deskriptive Maße, wie reale Anfangs- und Endzeitpunkte, die absolute und relative Gesamtdauer und Auftrittshäufigkeiten sowie die Häufigkeit pro Zeitintervall, die durchschnittliche Dauer und die entsprechenden Verteilungen über ein Merkmal oder ein Verhalten berechnet werden (Faßnacht, 1995).

1.6.3 Time Sampling

Das *Time-Sampling* oder *Zeitstichproben-Verfahren* ermöglicht eine Beobachtung des Verhaltens innerhalb definierter Zeiteinheiten, deren Zeitrahmen im Vorhinein von den durchführenden Personen festgelegt wurden (Schölmerich, 2011). Faßnacht (1995) beschreibt das *Time-Sampling* als Einteilung des Zeitstroms in „kurze, kontinuierlich aufeinanderfolgende Zeiteinheiten“. Für jedes Intervall müsse die beobachtende Person im Sinne des „Alles- oder Nichts-Prinzips“ eine Entscheidung treffen, ob das interessierende Merkmal oder Verhalten auftritt oder nicht (S.138). Es folgt eine Unterteilung in zwei Arten des *Time-Samplings*: *Intervall-Recording* und *Selected-Intervall-Recording*, die folgend näher betrachtet werden (Bakeman & Quera, 2011).

Intervall-Recording: Das Verhalten einer Person wird im Zuge zeitlich zusammenhängender Intervalle einer im Vorhinein fixierten Zeitdauer (z.B. 15 Sekunden) erhoben. Das Auftreten des gesuchten Events innerhalb der Zeitspanne eines Intervalls wird kodiert. Bakeman und Quera (2011) beschreiben drei Unterformen der *Time-Sampling/Intervall Recording* Methode: *Partial-Interval Sampling (One-Zero-Sampling)*, *Momentary Sampling (Instantaneous)* und *Whole-Intervall Sampling*.

- ***Partial-Interval Sampling (One-Zero-Sampling)*:** Die BeobachterInnen kodieren innerhalb eines Intervalls, ob ein Event aufgetreten ist oder nicht. Im Falle des Auftretens eines Events während eines Intervalls wird mit der Zahl 1 kodiert, wenn kein Event auftritt, notiert die beobachtende Person die Zahl 0. Die größte Schwachstelle dieser Erhebungsform ist die Tatsache, dass sobald ein Event innerhalb eines Intervalls kodiert wurde, die beurteilende Person kein weiteres Event innerhalb desselben Intervalls kodieren kann, selbst wenn das Verhalten mehrmals in dieser Zeitspanne auftritt. Deshalb kann es zu einer Unterschätzung der Auftrittshäufigkeiten kommen und gleichzeitig zu einer Überschätzung der Proportion eines Verhaltens führen. Im Zuge des *Partial-Interval Samplings* können mehrere Verhaltensformen gleichzeitig kodiert werden, sie schließen sich somit nicht gegenseitig aus.
- ***Momentary Sampling*:** Im Kontext dieser *Time-Sampling* Technik wird nur innerhalb eines Intervalls eine Kodierung abgegeben, wenn ein Verhalten zu einem bestimmten Zeitpunkt aufgetreten ist (z. B. nach dem Auftreten eines Tones).

- **Whole-Intervall Sampling:** Eine Kodierung wird nur dann durch die BeobachterInnen vorgenommen, wenn ein Verhalten für die gesamte Dauer des Intervalls aufgetreten ist. Als Variante dazu wird oft das *Predominant-Activity-Sampling* durchgeführt, bei welchem eine Kodierung gesetzt wird, sobald das gesuchte Verhalten mehr als die Hälfte der Zeit des Intervalls auftritt. Beim *Whole-Intervall-Samplings* und *Predominant-Activity-Samplings* schließen sich die Intervalle gegenseitig aus.

Selected-Intervall-Recording: Die Selected-Intervall Recording Methode wird eingesetzt, wenn die Kontingenz oder das sequentielle Auftreten eines Merkmals keine Rolle im Zusammenhang des Ziels der Forschung spielen. Die Zeitintervalle der Kodierung sind nicht zusammenhängend.

Einige mögliche Techniken zur Erhebung von *Time-Sampling Verfahren* sind die Verwendung von Stoppuhren oder Piepstönen, beide erfordern jedoch zusätzlich zur Datenerhebung eine hohe visuelle oder auditive Aufmerksamkeit der BeobachterInnen. In den meisten Fällen werden Checklisten oder Strichlisten verwendet, welche in den Zeilen die interessierenden Verhaltensmerkmale und in den Spalten dazu die Zeitanangaben abbilden. Eine sinnvolle Ergänzung zu den Formblättern bieten visuelle Synchronisationsmarker (z.B. dicke Trennungslinien), die eine größere Anzahl von Zeitintervallen zu Blöcken zusammenfasst (Faßnacht, 1995). Faßnacht (1995) schlägt zudem vor, die Entscheidung für die passende Länge eines Intervalls an das zu erhebende Verhaltensmerkmal anzugleichen. Generell kann ein zeitlich kurzes Zeitintervall die Genauigkeit der Erfassung eines Merkmales erhöhen, gleichzeitig gilt, je kürzer ein Intervall, desto schwieriger die Erhebung für die beobachtenden Personen. In der wissenschaftlichen Beobachtung werden häufig Intervalle mit einer Zeitdauer von fünf bis dreißig Sekunden gewählt (Faßnacht, 1995).

Der größte Kritikpunkt des *Time-Samplings* ist, dass im Vergleich zu *Event-Recording* die quantitativen Maße, die durch das Verfahren erhoben wurden, weniger genau sind, insbesondere wenn die Intervalle von langer Zeitdauer sind. Die ökonomische und einfache Handhabung sowie die niederschwellige Implementierung der *Time-Sampling* Erhebungsart sind zwar klare Vorteile, die für ihren Einsatz sprechen, dem gegenüber stehen aber eine geringere Informationsdichte und eine im Vergleich zu *Timed-Event-Recording* weniger akkurate Erhebungsmöglichkeit (Faßnacht, 1995).

1.7 BEOBACHTUNGSSYSTEME ZUR VERHALTENSKODIERUNG

Um dem Anspruch gerecht zu werden einen Verhaltensablauf in seiner Komplexität während einer Beobachtungssituation zu erfassen, braucht es ein den Anforderungen entsprechendes Beobachtungssystem (Bakeman & Quera, 2011). Die Art des Beobachtungssystems hat einen wesentlichen Einfluss auf die Erhebung und Einordnung der Verhaltensmerkmale und es werden eindeutige Richtlinien benötigt, um das Verhalten einer Person akkurat zuordnen zu können (Bakeman & Quera, 2011).

Schölmerich (2011) unterscheidet zwischen *Verbal-Systemen*, *Zeichen-Systemen*, *Kategorien-Systemen* und *Rating-Verfahren*, als diverse Beschreibungssysteme, die zur Verhaltenskodierung eingesetzt werden und deren Anwendung sich nicht gegenseitig ausschließen:

1.7.1 Verbalsysteme

In Form von Tagebuchaufzeichnungen, Ereignisbeschreibungen, Verlaufsprotokollen und der Methode der kritischen Ereignisse sind Verbalsysteme eine häufig eingesetzte Möglichkeit, um Verhalten mit sprachlichen Mitteln zu erfassen (Faßnacht, 1995). Die deskriptive Form ermöglicht es dem Forschenden auf einfache Art und Weise, ohne Entwicklung komplexer Kodiersysteme, Hypothesen zu generieren und einen ersten Eindruck zu einer Forschungsfrage zu gewinnen. Ein Nachteil dieser deskriptiven Erhebungsmethode ist die damit verbundene zeitaufwändige Mitschrift, welche sich in einer komplexen Beobachtungssituation oft nur schwer mit der Flüchtigkeit des beobachtbaren Verhaltens verbinden lässt.

1.7.2 Zeichen- oder Indexsysteme

Im Zuge der Anwendung des Zeichen- oder Indexsystems werden zu Beginn Verhaltensweisen definiert, die als Anzeichen für ein bestimmtes Kriterium gelten. Beim Auftreten eines zuvor festgelegten Verhaltens wird von den BeurteilerInnen innerhalb einer Beobachtungseinheit eine Kodierung vorgenommen. Innerhalb eines Intervalls ist es möglich mehrere Zeichen zu notieren, um einen Überblick über die Auftrittshäufigkeit des untersuchten Verhaltens zu bekommen. In den meisten Fällen sind Checklisten die Methode der Wahl, um das Auftreten eines Verhaltensmerkmals zu notieren. Die Anzahl der gleichzeitig zu bearbeitenden Kategorien ist abhängig vom Beobachtungstraining, der Häufigkeit des Auftretens eines Verhaltens und der Genauigkeit der Definition des Merkmals.

1.7.3 Kategoriensysteme

Kategoriensysteme versuchen verschiedene Aspekte des beobachteten Verhalten in unterschiedliche Kategorien einzuordnen, mit dem Ziel das beobachtete Verhalten möglichst umfassend zu beschreiben. Die Kategorien sind dabei voneinander unabhängig und in sich geschlossen. Die Anwendung des Kategoriensystems hat vor allem auf der Auswertungsebene klare Vorteile, da es versucht logische Kombinationen von Verhaltensaspekten klar zu definieren, um statistisch komplexe Analysen durchführen zu können.

1.7.4 Rating-Verfahren

Im Rahmen des Rating-Verfahren werden dem Verhalten Zahlen zugeordnet, indem der Beobachter oder die Beobachterin den Ausprägungsgrad eines Merkmals oder dessen Intensität oder Häufigkeit abschätzt und den Zahlen einer Messskala zuordnet (Faßnacht, 1995). Bipolare Ratings erfordern die Einschätzung eines Verhaltens auf einer Skala mit Endpunkten, welche gegenteilig sind und sich somit wechselseitig definieren, während bei der Verwendung von unipolaren Ratings der Ausprägungsgrad einer Variable geschätzt werden muss, z.B. geringes bis häufiges Auftreten (Schölmerich, 2011). Die Skalen sind in den meisten Fällen in gleich lange Streckeneinheiten geteilt, seltener wird eine Skala ohne Zahlenwerte als Richtlinien für die Kodierung vorgegeben. Möglich ist auch die Verwendung einer inhaltlich differenzierten Skala, die sich dadurch auszeichnet, dass die Skalenpunkte spezifische Verhaltensbedeutungen abbilden, welche in Form von Beispielen erörtert werden. In der psychologischen Quantifizierung von Verhalten wird dieses Verfahren aufgrund der ökonomischen Erfassungsmöglichkeit komplexer Dimensionen und der leichten Handhabung für die beurteilenden Personen häufig eingesetzt (Faßnacht, 1995; Schölmerich, 2011).

1.8 ZUSAMMENFASSUNG

Die wissenschaftliche Verhaltensbeobachtung ist ein in der psychologischen Forschung vielfach angewandtes Mittel zur Erhebung von Daten, mit dem Ziel nachhaltiger Verhaltensmerkmale und -abläufe, sowie Interaktionsprozesse abzubilden.

Von einer *systematischen Beobachtung* kann gesprochen werden, wenn die Beobachtung einer Situation auf einen theoretischen Bezugsrahmen aufgebaut ist, das Geschehen unter einer zuvor geplanten Beobachtung stattfindet und ein gezielter

Forschungszweck erfüllt wird (Seidel & Prenzel, 2010). Vor allem im entwicklungspsychologischen Bereich werden technische Hilfsmittel, aufgrund der damit verbundenen Möglichkeit der Fixierung eines flüchtigen Geschehens, häufig eingesetzt. Im Zuge der Untersuchung der kindlichen Entwicklung oder dem Verhalten von Säuglingen und Kindern sowie der Eltern-Kind-Interaktion wird die videogestützte Verhaltensbeobachtung genutzt, um intuitives und automatisiertes Verhalten zu erfassen. In der Praxis ist die videogestützte Verhaltensbeobachtung ein fester Bestandteil der psychologischen Forschung. Sie wird im entwicklungspsychologischen Kontext, aber auch für diagnostische und therapeutische Zwecke sowie Interventionen und Beratungen eingesetzt.

Eine standardisierte Beobachtung fordert für die Gültigkeit der Ergebnisse die Einhaltung der Gütekriterien *Objektivität*, *Reliabilität* und *Validität*. Die *Objektivität*, sowie die *Reliabilität* der videogestützten Verhaltensbeobachtung lässt sich mittels BeobachterInnenübereinstimmung und standardisiertem Vorgehen verbessern. Die *Validität* eines Beobachtungsinstruments ist gegeben, wenn jene Verhaltensmerkmale erhoben wurden, welche zu messen beabsichtigt waren. Die Überprüfung der Validität gestaltet sich in der Verhaltensbeobachtung jedoch oft schwierig.

Die Wahl des Schemas der Kodierung sowie die Anzahl der Kodiereinheiten und deren Strukturierung muss dem Forschungsziel und dem zu erhebenden Verhaltensmerkmal angepasst werden. Die Quantifizierung psychologischer Merkmale kann mittels *Event-Sampling* oder *Time-Sampling* vorgenommen werden. *Verbal-Systeme*, *Zeichen/Index-Systeme*, *Kategorien-Systeme* und *Rating-Verfahren* sind in der systematischen Beobachtung häufig eingesetzte Formen von Beobachtungssysteme, um ein Verhalten zu erheben.

2 ÜBEREINSTIMMUNG ZWISCHEN BEOBACHTERINNEN

Im Folgenden wird auf die Übereinstimmungsmaße *Interrater-Reliabilität* und *Intra-Rater Reliabilität* näher eingegangen, welche Auskunft über die Objektivität und Reliabilität des eingesetzten Beobachtungsverfahrens geben (Bakeman & Quera, 2011; Döring et al., 2016; Wirtz & Caspar, 2002).

2.1 INTERRATER-RELIABILITÄT

Die BeobachterInnenübereinstimmung, auch genannt *Interrater-Reliabilität*, gibt Auskunft über das Maß der Übereinstimmung zwischen den Erhebungen von zwei oder mehreren BeurteilerInnen. Eine hohe Übereinstimmung zwischen den BeobachterInnen ist gegeben, wenn möglichst viele Verhaltensmerkmale durch die beobachtenden Personen denselben Werten oder Kategorien zugeteilt werden und gleichzeitig eine geringe Anzahl inkongruenter Urteile vorliegt (McHugh, 2012). Eine hohe Urteilspräzision ist gegeben, sobald sich die Erhebung von BeobachterInnen, mit gleichem Wissensstand über die Beobachtung, vernachlässigbar gering voneinander unterscheiden und sich innerhalb eines vorab festgelegten kritischen Wertes befinden. Wesentlich ist dabei, dass die beobachteten Personen sich nicht gegenseitig beeinflussen und unabhängige Beurteilungen des Materials vornehmen. Die Austauschbarkeit der BeobachterInnen, ohne dass dieses Vorgehen zu großen Veränderungen in den Ergebnissen führt, ist eine essentielle Anforderung, welche für die Aussagekraft der Beurteilungsgüte und der Standardisierung eines Beobachtungsinventars wichtig ist (Wirtz & Casper, 2002). Eine hohe Interrater-Reliabilität ist Maßstab für die Zuverlässigkeit des Verfahrens und essentiell, um den Anforderungen der wissenschaftlichen Forschung zu genügen (Bakeman & Quera, 2011).

Die Überprüfung der Messgenauigkeit eines standardisierten Beobachtungsinstruments in einer systematischen Verhaltensbeobachtung erfolgt mittels Berechnung der BeobachterInnenübereinstimmung in jeder einzelnen Kategorie. Eine hohe Übereinstimmung in allen Kategorien deutet auf eine störungsfreie und messgenaue Datenerfassung durch das Beobachtungsinventar hin. Bei geringer Übereinstimmung zwi-

schen BeobachterInnen kann die Ursache in unzureichenden BeobachterInnentrainings, ungenaue Abgrenzungen oder Definitionen einzelner Kategorien oder in Beobachtungsfehlern begründet sein (Döring et al., 2016).

Einerseits ist die Interrater-Reliabilität notwendig um Auskunft darüber zu erhalten, ob die beurteilenden Personen den Erwartungen entsprechend Urteile treffen, andererseits um Feedback über die Kodierschemata sowie die Beobachtungseinschulung oder das Manual zu generieren. Döring et al. (2016) empfehlen bei der Entwicklung eines Verfahrens den Einsatz von Pretests zur Prüfung der BeobachterInnenübereinstimmung der einzelnen Kategorien und des Gesamttests. Dabei wird das Beobachtungsgeschehen einer zufällig gewählten Stichprobe von zwei oder mehreren BeobachterInnen, die mit den Anforderungen der Erhebung vertraut sind, parallel kodiert. Die Ergebnisse werden im Untersuchungsbericht vermerkt, um darauf aufbauende Veränderungen oder Schlussfolgerungen nachvollziehbar zu machen.

Wirtz und Caspar (2002) merken an, dass es keine von „Randbedingungen unabhängige Reliabilität“ gibt und die Übereinstimmungsreliabilität Informationen über den Gebrauch der Skalen geben, aber nichts über die Eigenschaften der Skalen selbst aussagen (S. 24). Die Interrater-Reliabilität kann somit lediglich festzustellen, ob sich die beurteilenden Personen in ihrer Erhebung an denselben Vorgaben orientieren (Wirtz & Caspar, 2002).

Zur Berechnung der Interrater-Reliabilität können je nach Skalenniveau der Kategorie unterschiedliche statistische Maße zur Berechnung der Übereinstimmung eingesetzt werden. Auf einige häufig verwendete Übereinstimmungsmaße wird in den folgenden Abschnitten vertiefend eingegangen:

2.1.1 Prozentuale Übereinstimmung

Die *Prozentuale Übereinstimmung (Proz. Ü.)* ist ein Übereinstimmungsmaß, welches den prozentualen Anteil jener Fälle angibt, in welchen die BeobachterInnen eine kongruente Entscheidung hinsichtlich eines Merkmals getroffen haben. Die Prozentuale Übereinstimmung berechnet die Anzahl kongruenter Urteile aller BeobachterInnen im Vergleich zu der Anzahl aller abgegebenen Urteile. Dieses Übereinstimmungsmaß ist eine sehr einfache und gängige Methode zur Schätzung von BeurteilerInnenübereinstimmung bei nominalskalierten Daten (McHugh, 2012; Seidel & Prenzel, 2010).

Der Wertebereich der Prozentualen Übereinstimmung liegt zwischen 0% und 100%. Je höher der Wert, desto besser die Übereinstimmung zwischen den Beurteilenden

(Döring et al., 2016). Bei direkt beobachtbaren Variablen empfehlen Döring et al. (2016) einen Wert nahe 100% als Indiz für eine gute Übereinstimmung zwischen den beurteilenden Personen, nach McHugh (2012) sollte für eine gute BeobachterInnen-übereinstimmung eine Prozentuale Übereinstimmung von mindestens 80% erreicht werden. Spezifische Wertebereiche oder Cutoffs wurden in der Literatur jedoch keine angegeben.

Nicht berücksichtigt wird bei der Berechnung der Prozentualen Übereinstimmung, dass die BeobachterInnen auch zufällig zu gleichen Einschätzungen kommen können. Das bedeutet, sie ist nicht zufallsbereinigt (Wirtz & Caspar, 2002). Deshalb wird von einer Überschätzung der BeobachterInnenübereinstimmung durch die Berechnung der Prozentuale Übereinstimmung ausgegangen (Döring et al., 2016). Der Einsatz dieser Art der Übereinstimmungsberechnung wird bei ungleicher Verteilung und seltenem Auftreten der beobachteten Verhaltensmerkmale nicht empfohlen, sondern sollte unter diesen Bedingungen nur in Zusammenhang mit weiteren Übereinstimmungsmaßen berechnet und interpretiert werden (Seidel & Prenzel, 2010). Wenn die prozentuale Übereinstimmung bei mehr als zwei Beobachtenden berechnet werden soll, ist zuerst eine paarweise Kalkulation nötig, aus welcher der Durchschnitt gemittelt werden kann (Döring et al., 2016).

2.1.2 Kappa-Koeffizient

Es gibt je nach Anzahl der beurteilenden Personen und dem Skalenniveau des Merkmals mehrere statistische Möglichkeiten die Übereinstimmung von Beobachtenden zu überprüfen. Häufig wird die Berechnung der Kappa-Koeffizienten eingesetzt. In den nächsten Abschnitten wird *Cohens Kappa*, das *Gewichtete Cohens Kappa* und *Fleiss Kappa* vorgestellt.

Cohens Kappa: Eines der meist eingesetzten Methoden um Informationen über die Interrater-Reliabilität, wie gut zwei BeobachterInnen in ihren Urteilen übereinstimmen, zu generieren, ist die Berechnung des Cohens Kappa-Koeffizienten, die bei normalskalierten Variablen eingesetzt werden kann (Grouven, Bender, Ziegler & Lange, 2007). Der Cohens Kappa Koeffizient ist ein zufallsbereinigtes Übereinstimmungsmaß, welches seine Grundlagen in der Prozentualen Übereinstimmung hat, im Gegensatz dazu aber die durch den Zufall erwarteten Übereinstimmungen berücksichtigt. Somit

wird bei der BeobachterInnenübereinstimmung das Verhältnis der erwarteten Übereinstimmungen in Bezug zu den tatsächlich beobachteten Werten miteinberechnet (Döring et al., 2016).

Die Berechnung des Kappa-Koeffizienten wird von Bakeman & Quera (2011) besonders im Zuge von BeobachterInnentrainings und Datenzusammentragungen bei der Erstellung eines Verfahrens empfohlen.

Das Cohens Kappa berechnet mithilfe von Kontingenztabelle wie oft BeobachterInnen in ihren Erhebungen übereinstimmen, unter Berücksichtigung aller insgesamt kodierten Ereignisse und unter Berücksichtigung von Zufallsübereinstimmungen.

Der Kappa-Koeffizient kann einen Wert zwischen -1 und +1 annehmen, wobei der Wert +1 einer perfekten Konkordanz entspricht. Ein Kappa-Koeffizient von 0 bedeutet, dass die beobachtete Übereinstimmung nicht besser als eine durch Zufall erwartete Übereinstimmung ist. Der Wert -1 tritt auf, wenn keine Übereinstimmungen zwischen den Beobachtenden gegeben sind. Ein Wert unter der Ziffer 0 indiziert, dass die Übereinstimmung zwischen den Kodierergebnissen der beurteilenden Personen schlechter ist, als die, einer per Zufall durchgeführten Kodierung (Bakeman & Quera, 2011; McHugh, 2012). Die Grenzwerte für eine gute Interrater-Reliabilität liegen zwischen $\kappa = .6$ und $\kappa = .75$. Werte ab $.75$ gelten als sehr gute Übereinstimmung der BeobachterInnen (Döring et al., 2016). Werte niedriger als $.4$ deuten auf eine unzureichende Übereinstimmung zwischen den Beobachtenden hin (Döring et al., 2016). Diese Grenzwerte werden im wissenschaftlichen Diskurs häufig kritisch betrachtet, da der Kappa-Koeffizient abhängig von unterschiedlichen Faktoren ist, welche nicht alleinig die Übereinstimmung zwischen den Beobachtenden betreffen. Kappa kann wesentlich von Einflüssen, welche nicht mit dem Kodierverhalten der beobachtenden Personen in Verbindung stehen, beeinflusst werden (Wirtz & Caspar, 2002).

Die Berechnung der BeobachterInnenübereinstimmung ist an die Kontingenztabelle geknüpft. So kann eine perfekte Übereinstimmung von +1 nur dann erreicht werden, wenn die Randsummenverteilung der Kontingenztabelle für beide BeobachterInnen gleich ist (Wirtz, 2006; Wirtz und Caspar, 2002). Eine ungleiche Verteilung der Randsummen wirkt sich negativ auf die Berechnung des Kappa-Koeffizienten aus (Wirtz & Kutschmann, 2007). Eine gleichmäßige Verteilung der Daten innerhalb der Kreuztabelle kann hingegen positiven Einfluss auf die Einschätzung der Interrater-Reliabilität haben (Wirtz & Caspar, 2002).

Bei der Berechnung von Cohens Kappa muss zusätzlich bedacht werden, dass der Kappa-Koeffizient durch Faktoren beeinflusst werden kann, welche unabhängig von der Güte der Beurteilung sind und nicht die Übereinstimmung zwischen den BeurteilerInnen per se betreffen. So führen insgesamt niedrige Prävalenzen eines Merkmals zu einem niedrigeren Cohens Kappa (Bakeman & Quera, 2011; Mandrekar, 2011; Wirtz, 2006). Weiterführend kann es zu einer Unterschätzung der BeobachterInnenübereinstimmung in der Interpretation, aufgrund des niedrigen Kappa-Koeffizienten kommen. Der prozentuale Anteil, mit dem eine beobachtende Person das Merkmal als vorhanden einschätzt, ist ein zusätzlicher Faktor, welcher Einfluss auf das Cohens Kappa haben kann. Je niedriger die Einschätzung, desto niedriger Cohens Kappa (Wirtz & Kutschmann, 2007).

Manifeste Merkmale sind in der Erhebung einfacher zu erfassen als latente Konstrukte und auch die Anzahl der Ausprägungen eines Merkmals können Einfluss auf die Übereinstimmung haben (Döring et al., 2016).

Die Anzahl der Beurteilungskategorien, kann zusätzlich Einfluss auf Kappa nehmen (Bakeman, Quera, McArthur & Robinson, 1997). Je höher die Anzahl der Kodiermöglichkeiten, desto höher sollte der Kappa-Koeffizient ausfallen, um eine gute BeobachterInnenübereinstimmung zu garantieren (Bakeman & Quera, 2011).

Gewichtetes Cohens Kappa: Das gewichtete Kappa wird eingesetzt, wenn bei ordinalskalierten Kategorien die Differenzen unterschiedlichen Gewichtungen unterliegen. Vor allem bei Rating-Verfahren, kann der Einsatz des gewichteten Kappa-Koeffizienten sinnvoll sein, da Differenzen in der Erfassung eines Merkmals, je nach Grad der Unterscheidung stärker oder weniger stark gewichtet werden können. Unstimmigkeiten in der Übereinstimmung, die in der ordinalen Skala näher beisammen liegen, werden weniger stark gewichtet als Unterschiede größerer Differenz. Somit ermöglicht das gewichtete Cohens Kappa ein differenzierteres Maß der Übereinstimmung zwischen BeobachterInnen (Bakeman & Quera, 2011).

Fleiss Kappa: Bei mehr als zwei BeobachterInnen wird für die Berechnung der Übereinstimmung das Fleiss Kappa eingesetzt (Wirtz & Caspar, 2002).

2.1.3 Intraklassenkorrelation

Bei intervallskalierten Daten kann die statistische Intraklassenkorrelation (ICC) durchgeführt werden. Es wird die Stärke des Zusammenhanges der Erhebungen von zwei

oder mehreren beurteilenden Personen über mehrere Beobachtungszeitpunkte hinweg dargestellt. Die berechneten Werte der Intraklassenkorrelation können zwischen -1 und +1 liegen, dabei gilt, je näher ein Wert zu +1, umso besser der Zusammenhang der Urteile der BeobachterInnen und desto besser die Interrater-Reliabilität (Bakeman & Quera, 2011; Döring et al., 2016).

2.1.4 Inkongruente Urteile in der Interrater-Reliabilität

Eine niedrige Beobachtungsübereinstimmung kann Wirtz und Caspar (2002) zufolge vorrangig auf zwei Ursachen zurückgeführt werden: Einerseits können Uneinigkeiten der BeobachterInnen darüber, welches Merkmal erhoben werden soll, ein Grund für eine niedrige Übereinstimmung sein. Dieses Problem tritt auf, wenn Unteraspekte nicht genau definiert wurden, oder ein Unwissen besteht, wie diese gewichtet und zusammengeführt werden können. Andererseits können Inkongruenzen in der Übereinstimmung zwischen BeobachterInnen auftreten, wenn die beurteilenden Personen die Ausprägung der Merkmalsdimensionen verschieden wahrnehmen und somit unterschiedliche Schwellen zur Einschätzung eines Merkmals heranziehen.

Wirtz und Caspar (2002) benennen für diese Problematik konkret zwei mögliche Ursachen für inkongruente Urteile zwischen BeobachterInnen: Der *Effekt der mangelnden Konsistenz* und der *Effekt unterschiedlicher Grundwahrscheinlichkeiten* (S.62, S.98).

Effekt mangelnder Konsistenz: Der *Effekt der mangelnden Konsistenz* tritt auf, wenn sich das Grundverständnis der Kategorien oder Merkmale zwischen den BeobachterInnen unterscheidet oder jene nicht in der Lage sind Merkmalsausprägungen zutreffend abzuschätzen.

Effekt unterschiedlicher Grundwahrscheinlichkeiten: Der *Effekt unterschiedlicher Grundwahrscheinlichkeiten* besagt, dass BeobachterInnen die einzelnen Kategorien oder Merkmale mit unterschiedlicher Grundwahrscheinlichkeit wählen. In der Praxis bedeutet das, dass bei dichotomen Items die BeurteilerInnen verschiedene Grenzwahrscheinlichkeiten anwenden, wonach eine Person sich schon bei geringer Ausprägung eines Merkmals für eine Kodierung entscheidet, wohingegen die andere Person erst eine höhere Ausprägung der Merkmalsintensität als genügend betrachtet. Individuelle Unterschiede zwischen den Beobachtenden und die damit verbundene progressive oder konservative Urteilsbildung kann somit zu nicht unwesentlichen Verzerrungen der BeurteilerInnenübereinstimmung führen.

Um eine Aussage darüber zu machen, ob die BeobachterInnen ein Merkmal oder eine Kategorie gleichermaßen beständig notiert haben, wird im Werk von Wirtz und Caspar (2002) vorgeschlagen einen *Test auf Gleichheit der Schwellenwerte* durchzuführen, in welcher die Methode der *Zusammengefassten Prozentualen Übereinstimmung* angewandt wird (S. 98). Dieses Vorgehen wird vor allem empfohlen, wenn die Randverteilungen zwischen den BeurteilerInnen stark voneinander abweichen oder wenn das interessierende Merkmal sehr häufig oder in nur sehr geringem Ausmaß auftritt ($p < 0,2$ oder $p > 0,8$) (Wirtz und Caspar, 2002, S. 103). Wirtz und Caspar (2002) betonen, dass dieser Test ausschließlich als „Maß der Ähnlichkeit der Entdeckungsraten“ dient und somit nicht als Übereinstimmungsmaß zwischen den beurteilenden Personen interpretiert werden darf (S. 103). Je näher das Ergebnis beim Prozentrang 100 ist, desto homogener haben die beobachtenden Personen die Kategorie erhoben, und desto weniger wurden unterschiedliche Schwellen bei der Kodierung herangezogen.

2.2 INTRA-RATER-RELIABILITÄT

Die Intra-Rater-Reliabilität kontrolliert die Übereinstimmung der Auswertungen einer einzelnen beobachtenden Person über mehrere Beobachtungszeitpunkte hinweg (Grouven et al., 2007; Döring et al., 2016). Sie macht Aussagen ob und in welchem Maße eine beobachtende Person bei wiederholtem Erheben der Daten zu gleichen oder ähnlichen Ergebnissen gelangt, also inwiefern die Ergebnisse einer Person bei der Erhebung desselben Tests oder derselben Beobachtung zu zwei verschiedenen Messzeitpunkten übereinstimmen oder sich voneinander unterscheiden. Die Intra-Rater-Reliabilität misst somit, in welchem Ausmaß die erhebende Person stabil und konsistent das Testinstrument anwendet, macht jedoch keine Aussagen, inwiefern sich die Urteile dieser Person mit den Urteilen anderer BeobachterInnen deckt. Die Konsistenz der Erhebungsübereinstimmung kann durch Erinnerungseffekte oder durch das Ausmaß der dazwischenliegenden Zeitintervalle verzerrt werden (Wirtz & Caspar, 2002). Das Beurteilungsmaß ist eng verknüpft mit der Retest-Reliabilität und gibt Auskunft über die Stabilität und Konsistenz der Kodierungsergebnisse der beurteilenden Person (Mees, 1997; Wirtz & Caspar, 2002). Die Intra-Rater-Reliabilität unterscheidet sich in der Berechnungsweise nicht von der Berechnung der Interrater-Reliabilität, welche in diesem Kapitel bereits vertiefend erörtert wurde. Abhängig von der Art der Skalen der Daten wird die Intra-Rater-Reliabilität auf unterschiedliche Weise berechnet.

Bei nominalskalierten Variablen ist die Berechnung des Cohens Kappa ein gängig eingesetztes Verfahren.

2.3 BEOBACHTUNGSFEHLER UND -VERZERRUNGEN

Bei der Durchführung einer wissenschaftlichen Verhaltensbeobachtung können aufgrund der Komplexität einer Beobachtungssituation auf vielen Ebenen Beobachtungsfehler und -verzerrungen auftreten. Diese können auf die testende Person zurückzuführen sein, es können aber auch Verzerrungen durch die Situation an sich oder durch das Verhalten der zu beobachtenden Person entstehen, sowie aufgrund Mängel des eingesetzten Verfahrens verursacht werden. Urteilsfehler treten auf, wenn ein systematisches unter- oder überschätzen von Merkmalsausprägungen in der Urteilsbildung stattfindet (Büttner & Florack, 2009; Döring et al., 2016; Rentzsch & Schütz, 2009).

Schölmerich (2011) unterscheidet zwischen Beobachtungsfehler systematischer und unsystematischer Weise. Systematische Unterschiede zwischen den BeobachterInnen treten vermehrt über mehrere Durchgänge hinweg wiederholt auf, sie beruhen unter anderem auf unterschiedliche Persönlichkeitsmerkmale (z.B. eine beurteilende Person kodiert stets milder als die anderen BeurteilerInnen). Unsystematische Unterschiede bilden Momente der Unkonzentriertheit, Ermüdung oder Ablenkung der BeobachterInnen ab. Diese Unterschiede werden bei größeren Studien mit ausreichend großer Stichprobe als wenig heikel erachtet, da eine hohe Anzahl anderer Daten diese Fehler ausgleichen können. Im Gegensatz dazu führen systematische Fehler, vor allem wenn sie von einer Vielzahl der durchführenden Personen gemacht werden, zu weitergehenden Verzerrungen der Ergebnisse (Döring et al., 2016). Bereits im Vorhinein ist es daher essentiell, sich bereits bei der Erstellung eines Messinstruments und vor Beginn der Datenerhebung mit möglichen Urteilsverzerrungen auseinander zu setzen und diese auf allen Ebenen grundlegend zu bedenken.

Fisseni (1997) geht spezifisch auf Fehler ein, welche speziell bei der Verhaltensbeobachtung auftreten. Als häufige Verzerrungstendenzen werden unter anderem folgende Punkte genannt (S.201f):

Überforderte Differenzierungsfähigkeit: Da die Wahrnehmung einer jeden Person nur begrenzt aufnahmefähig ist, kommt es zu Beobachtungsfehler, wenn die subjektiven Kapazitäten der beobachtenden Person überstiegen werden (z.B. zu viele Merkmale müssen simultan erhoben werden).

Unschärfe Definition: Zu Verzerrungstendenzen führen außerdem Beobachtungsintervalle, welche nicht genau genug definiert wurden und der beurteilenden Person zu viel Interpretationsraum geben. Somit wird ein Merkmal nicht mehr objektiv erhebbar, da die beobachtende Person subjektive, eigenständig angelegte Regeln anwendet, nach welchen die Kodierung erfolgt. Um diesen Fehler zu vermeiden kann ein breit gefächertes Merkmal in kleinere Erhebungsfacetten aufgegliedert werden. Fisseni (1997) warnt zudem vor dem Auftreten von Begriffsinkonsistenz und Begriffspluralismus, da sie nur schwer kontrollierbar sind, aber die Datenerhebung erheblich beeinflussen. Ausführliche Definitionen der Kategorien können gegen ein Auftreten dieser Problematik vorbeugen.

Unvertrautheit mit Beobachtungseinheiten: Wenn eine beobachtende Person nicht genügend mit den Beobachtungseinheiten und den allgemeinen Kodierregeln vertraut ist, führt dies ebenfalls zu Verzerrungstendenzen in der Beobachtung. Deshalb ist eine Schulung mit ausreichender Einführung in das Verhaltensbeobachtungsinstrument, mit dessen Aufbau und Regeln unabdingbar.

Unvertrautheit mit der Probandengruppe: Eine beurteilende Person sollte in den meisten Beobachtungserhebungen vertraut mit dem Milieu, den Normen und Standards der Probanden sein, da sonst Fehleinschätzungen in der Beurteilung von Interaktionen oder anderen Verhaltensabläufen auftreten können, z.B. infolge von sprachlichen Problemen.

Eingriff in den Untersuchungsablauf: Wenn sich eine beobachtende Person nicht an die festgelegten Instruktionen hält, kommt es zu Verzerrungen in der Erhebung, da ein standardisiertes Vorgehen nicht mehr gewährleistet ist.

In den folgenden Unterkapiteln werden weitere Beobachtungsfehler angeführt, welche Auswirkungen auf die Ergebnisse einer Verhaltensbeobachtung zeigen. Diese gliedern sich in Beobachtungsfehler zu Lasten der BeobachterInnen, der Situation und des Beobachtungssystems (Greve und Wentura 1997; Spinath & Becker, 2011).

2.3.1 Fehler zu Lasten der Beobachtenden

Die Erhebung einer Beobachtung ist stets verbunden mit der selektiven Wahrnehmung der Testenden, die auf Vorannahmen gestützt werden, sowie dem subjektiven Interesse bestimmter Aspekte der Forschenden. Die Vertrautheit mit dem untersuchten Gegenstand sowie der Wunsch nach brauchbaren Ergebnissen, können zu einem Verlust in der erforderlichen wissenschaftlichen Distanz und somit zu einer Verringerung der Erhebungs- und Auswertungsobjektivität führen (Büttner & Florack, 2009; Döring et al., 2016).

Die Unterteilung der Fehler zu Lasten der beobachtenden Personen wird in die Kategorien Wahrnehmungsfehler, Interpretationsfehler und Erinnerungsfehler vorgenommen (Greve und Wentura, 1997; Spinath & Becker, 2011).

1. Wahrnehmungsfehler

Im Zuge der selektiven Wahrnehmung eines jeden Individuums, werden Aspekte des Geschehens beobachtet, die auf individuellen Erwartungen oder Interessen basieren, während anderen weniger Beachtung gegeben wird. Ermüdungseffekte, nachlassende Aufmerksamkeit und mangelnde Motivation wirken sich negativ auf die Gültigkeit der Ergebnisse aus. Fisseni (1997) schlägt als Lösungsansatz dazu eine Begrenzung der Beobachtungsdauer sowie BeobachterInnenrotationen vor, die effektive Maßnahmen zur Verringerung der Fehlerquellen darstellen.

Zusätzlich können auch Emotionen und die aktuelle Stimmungslage Einfluss auf die Wahrnehmung der Testenden haben. Negative Verhaltensweisen werden zum Beispiel oftmals stärker wahrgenommen und können den Gesamteindruck verzerren (Rentzsch & Schürz, 2009). In positiver Stimmung werden zudem positive Aspekte besser gefiltert und besonders gut erinnert. Bei negativer Stimmungslage werden Informationen stärker im Detail und weniger heuristisch wahrgenommen (Döring et al., 2016). Das führt zu der Schlussfolgerung, dass die Genauigkeit der Kodierung, sowie das Gesamtbild der Erhebungssituation unter anderem abhängig vom Gemütszustand der testenden Person ist (Rentzsch & Schürz, 2009).

Einige spezifische Beispiele von Wahrnehmungsfehlern, welche folgend näher erläutert werden sind *der Halo-Effekt*, *der Observer Drift*, sowie *der Logische Fehler*.

Halo-Effekt: Der Halo-Effekt oder Ausstrahlungseffekt beschreibt die Tendenz der Beurteilenden, keine Differenzierung zwischen verschiedenen unabhängigen

Merkmale in der Urteilsbildung durchzuführen. Häufig passiert dies im Zusammenhang mit auffälligen Persönlichkeitsmerkmalen, die in Verbindung mit anderen Merkmalen gebracht werden und die Wahrnehmung dieser beeinflussen. Die beurteilende Person neigt dazu, sich von einer hervorstechenden Eigenschaft blenden zu lassen und das interessierende Verhaltensmerkmal nicht differenziert, sondern unter dem Aspekt des Gesamteindrucks, zu beurteilen (Rentzsch & Schütz, 2009).

Observer Drift: Der Observer Drift beschreibt eine Veränderung der „Standards“ der beobachtenden Personen (Spinath & Becker, 2011, S.343). Die Veränderung im Kodierverhalten kann positiv wie negativ auf die Ergebnisse einwirken. Einerseits können BeobachterInnen mit der Zeit aufgrund Ermüdung etc. zu weniger genauen Erhebungen neigen, andererseits kann die Erhebung über die Beobachtungsdauer hinweg auch gewissenhafter und präziser ausfallen, da sie mit dem Verfahren vertrauter sind oder die Anforderungen des Inventars besser auf das Geschehen abstimmen können. Gewöhnungs- und Lerneffekte können langfristig gleichermaßen systematische Veränderung in der Wahrnehmung einer Beobachtung herbeiführen (Girard & Cohn, 2016; Spinath & Becker, 2011).

Logischer Fehler: Der Logische Fehler tritt auf, wenn die BeurteilerInnen versuchen eine bestimmte Persönlichkeitstheorie in der Beobachtung zu bestätigen, infolgedessen es zu einer Verzerrung in der Wahrnehmung einer Person kommt. Es wird somit unbewusst versucht ein der Theorie entsprechendes Bild zu konstruieren (Spinath & Becker, 2011).

2. Interpretationsfehler

Interpretationsfehler beziehen sich auf Fehler der Beobachtenden, das Geschehen aufgrund verschiedener Faktoren falsch zu interpretieren. *Der Ankereffekt, der Primacy-Recency Effekt, die Tendenz zur Mitte sowie der Milde-Härte-Effekt* sind dieser Fehlergruppe zuzuordnen.

Ankereffekt: Der Ankereffekt tritt auf, wenn eine Beurteilung einer Person oder eines Geschehens an das zuletzt beobachtete geknüpft wird und als Referenz oder Maßstab verwendet wird (Spinath & Becker, 2011).

Primacy-Recency Effekt: Der Primacy-Recency Effekt erklärt den speziellen Einfluss des Zeitpunktes beim Auftreten des zu beurteilenden Objekts im Zuge einer Beurteilung. Die zeitliche Abfolge von Wahrnehmung unterschiedlichem Verhaltens hat

einen nicht zu unterschätzenden Einfluss auf das Urteil von beobachtenden Personen. Der *Primacy-Effekt* bezieht sich dabei auf mögliche Effekte, welche Aufgrund des Auftretens eines Merkmals zu Beginn eines Geschehens zu einer verzerrten Einschätzung späterer Beobachtungen führen. Ebenso kann das Auftreten eines Merkmals am Ende, der sogenannte *Recency-Effekt* einen verzerrenden Effekt zu Folge haben. Dies kann dazu führen, dass Merkmale zu Beginn positiver bewertet werden als nachfolgende, oder der Eindruck eines Geschehens vorrangig an die Merkmale, welche am Ende auftraten, geknüpft ist. Die systematische Veränderung der Reihenfolge ist ein häufiger Versuch Verzerrungen so gering wie möglich zu halten (Rentzsch & Schütz, 2009).

Tendenz zur Mitte: Die zentrale Tendenz bezieht sich auf einen Fehler, der im Zusammenhang mit Ratingskalen auftritt und die Tendenz der beurteilenden Personen beschreibt zur Mitte hin zu beurteilen, das bedeutet eher Mittelwerte als extreme Randwerte zu kodieren. Auch bei Wiederholungen von Erhebungen ist häufig erkennbar, dass Beobachtende ihre Einschätzung zur Mitte korrigieren (Fisseni, 1997).

Fehler der Milde/ Fehler der Härte: Ein häufiger Fehler in der Beurteilung von Personen besteht darin, dass eine beobachtende Person das Verhalten einer Versuchsperson auf zu großzügige Weise beurteilt. Diese Verzerrung kann bei verbalen Beurteilungen sowie im Zuge von Erhebungen mit Ratingskalen auftreten (Fisseni, 1997). Diese Urteilsverzerrung kann auch in die gegenteilige Richtung erfolgen, wenn eine beobachtende Person die Versuchsperson im Vergleich übermäßig negativ einschätzt. Diese Verzerrung nennt sich dann Fehler der Härte (Rentzsch & Schütz, 2009).

3. Erinnerungsfehler

Erinnerungsfehler können auftreten, wenn die Beobachtung und die Protokollierung zeitlich versetzt stattfinden. Je mehr Zeit zwischen Beobachtung und Kodierung der Daten vergeht, umso mehr kann angenommen werden, dass das wahrgenommene Geschehen durch Erinnerungseffekten beeinflusst wird. So wird das Geschehen, welches zu Beginn als auch am Ende einer Beobachtungserhebung stattfindet, besonders gut in Erinnerung behalten und kann zu einer Verzerrung der Beurteilung führen (Spinath & Becker, 2011). Döring et al (2016) nennt auch die individuelle Gedächtniskapazität als Quelle für mögliche Erinnerungsfehler.

2.3.2 Fehler zu Lasten der Beobachtungssituation

Reaktivität erläutert den Effekt, dass die beobachtete Person in der Testsituation sich anders verhält als in einer ähnlichen Situation außerhalb der Testung (Büttner & Florack, 2009, S.189). Die Künstlichkeit, welche in einer Beobachtungssituation aufgrund situativer Einflüsse entstehen kann (z.B. Beobachtung im Labor) trägt zu einem veränderten Verhalten der Testpersonen bei (Döring et al., 2016). Dabei spielen laut Rentzsch & Schürz (2009) der *Effekt der Selbstdarstellung*, der Wunsch einen guten Eindruck beim Gegenüber zu erwecken und der *Effekt der sozialen Erwünschtheit*, die Tendenz sich den sozialen Normen entsprechend zu verhalten, eine weitgehende Rolle.

Fisseni (1997, S.224) unterscheidet konkret zwischen vier möglichen Fehlerarten, welche auf das Untersuchungsobjekt zurückzuführen sind:

Versuchspersoneneffekt: Der Effekt beschreibt eine Verhaltensänderung einer Person aufgrund des Wissens, dass sie sich in einer Beobachtungssituation befindet und ihr Verhalten erfasst und gewertet wird.

Rollenselektion: Die Rollenselektion meint die bewusste Wahl einer bestimmten Rolle seitens der Versuchsperson in der Erhebungssituation.

Messung als Agens der Veränderung: Im Kontext einer Erhebung kann es im längerfristigen Kontext zu neuen Lernprozessen kommen und würde zu damit verbundenen Veränderungen eines Merkmals aufgrund der Messerhebung selbst führen. In der Verhaltensbeobachtung könnten eine langfristige Teilnahme in einer Beobachtungsstudie Auswirkungen auf Veränderungen im Verhalten haben. Fisseni (1977) betont, dass die in der Forschung meist eher kurzfristig eingesetzten Erhebungsarten jedoch kaum einen Effekt auf Verhaltensänderungen haben.

Reaktionsneigung: Diese Verzerrungstendenz seitens der Versuchsperson spielt vor allem bei schriftlichen Erhebungen mittels Fragebögen eine Rolle. Der *JA-Sager-Effekt*, bei welchem die Versuchsperson stets „JA“ als Antwort wählt, sowie die Tendenz stets die linke oder rechte, erste oder letzte Antwort anzukreuzen, führen zu unbrauchbaren Daten für die Forschenden.

Eine Möglichkeit, einige dieser Fehler zu vermeiden, stellt die Durchführung einer verdeckten Beobachtung dar, was aufgrund ethischer oder Gründen der Durchsetzbarkeit

oft nicht möglich ist. Jedoch kann in einer offenen Beobachtungssituation davon ausgegangen werden, dass sich die teilnehmenden Personen mit der Zeit an die Tatsache gewöhnen, beobachtet zu werden (Büttner & Florack, 2009).

Das Umfeld kann in der Beobachtungssituation auch durch ungünstige Umstände, wie schlechte Lichtverhältnisse oder laute Hintergrundgeräusche, die Beobachtung negativ beeinflussen (Spinath & Becker, 2011). Ebenso können Fehler bei der Durchführung auftreten, z.B. wenn das Beobachtungsinstrument falsch eingesetzt wird oder eine schlechte Positionierung der Kamera zu unbrauchbarem Datenmaterial führt. Defekte Videoaufzeichnungen oder ein Verlust von Mitschriften des beobachteten Geschehens sind ebenfalls konkrete Beispiele, welche zu einem hohen Informationsverlust führen und die Auswertung der Erhebung beeinträchtigen (Döring et al., 2016).

2.3.3 Fehler zu Lasten des Beobachtungssystems

Das Beobachtungssystem an sich kann ebenfalls Einfluss auf Beobachtungsfehler haben. Je nach Gestaltung des Kategoriensystems können den beurteilenden Personen mehr oder weniger Möglichkeiten zur eigenen Interpretation bei der Erfassung eines Merkmals gegeben werden. Je größer der Spielraum, desto eher können die BeobachterInnen dasselbe Verhalten unterschiedlich kodieren. Je detaillierter die vorgegebene Beschreibung der Verhaltenskategorie, desto eher wird die Kodierung den Vorstellungen der forschenden Personen entsprechen. Die Komplexität des Beobachtungssystems ist ein zusätzlicher Faktor welcher zu Verzerrungen der Beobachtung führen kann. Eine hohe Anzahl an Kategorien führt im Vergleich zu einem Inventar mit geringerer Anzahl an Kategorien zu weniger reliablen Ergebnissen (Büttner & Florack, 2009).

2.3.4 Vermeidung von Beobachtungsfehlern

Sowohl Rentzsch & Schürz (2009), als auch Döring et al. (2016) sehen eine Einschulung im Umgang mit dem Manual, klare Richtlinien in der Erhebung und die Durchführung von BeobachterInnentrainings als ein effektives Mittel, um Beobachtungsfehler und -verzerrungen seitens der beobachtenden Personen entgegenzuwirken. Eine der Erhebung vorhergehende Aufklärung über Beobachtungsfehler und Verzerrungseffekte kann ebenfalls dazu beitragen diese zu verringern. Konkret helfen im Rahmen von Beobachtungserhebungen klar strukturierte Kategorien und konkrete Spezifizierungen im Regelwerk, um Urteile zu objektivieren und reliabler zu machen. Des Wei-

teren raten Rentzsch & Schürz (2009) Beobachtungen und Beurteilungen nicht simultan durchzuführen, sowie den Einsatz mehrerer BeobachterInnen um Ergebnisse abzusichern. Döring et al. (2016) propagieren zusätzlich eine Vorauswahl geeigneter Personen für den Einsatz als BeobachterInnen, sowie wiederholende Qualitätsüberprüfungen der testenden Personen bei größeren Studien.

2.4 VORGEHEN ZUR VERBESSERUNG DER BEOBACHTERINNENÜBEREINSTIMMUNG

Für eine gültige Aussagekraft der BeobachterInnenübereinstimmung dürfen sich die Beobachtenden bezüglich dem Hintergrundwissen über theoretische Grundlagen des Beobachtungssystems und Gebrauch der Skalen nicht voneinander unterscheiden (Wirtz & Caspar, 2002). Eine größere Anzahl an BeobachterInnen ist hilfreich, um Verzerrungseffekte zu vermeiden und Übereinstimmungen zu deuten und somit auch Rückschlüsse über die Reliabilität und Objektivität des Beobachtungsinstruments zu ziehen (Fieguth, 1977).

Schölmerich (2011) empfiehlt zur Bestimmung und Aufrechterhaltung der BeobachterInnenübereinstimmung einen Vorgang mit festgelegten Reihenfolgen einzuhalten: Zu Beginn soll eine Erörterung der Verhaltenskategorien mit Definitionen und Beispielen durchgeführt werden, danach erfolgt eine praktische Anwendung unter der Verwendung von Videomaterial. Die Daten werden dann mithilfe passender statistischer Mittel auf eine Übereinstimmung zwischen den BeobachterInnen überprüft. Falls die Übereinstimmung nicht gut genug sein sollte, erfolgt eine Überarbeitung und eine Wiederholung der genannten Prozedur. Wenn eine ausreichend hohe Übereinstimmung gegeben ist, können weitere Beobachtungen durchgeführt werden und die Ergebnisse in einem Datensatz angelegt werden, jedoch sollte die Übereinstimmung zwischen den BeobachterInnen immer wieder überprüft werden.

Die BeobachterInnen selbst sollten mit dem Beobachtungsfeld vertraut sein, es ist jedoch abhängig von der Situation und dem Ziel der Erhebung, ob spezielles Wissen über die Zielgruppe notwendig ist. Ein generelles Interesse an der Untersuchung ist jedoch Fieguth (1977) zufolge ratsam, um eine Verzerrung der Erhebung aufgrund Desinteresse zu vermeiden.

Fieguth (1977) empfiehlt außerdem aus mehreren Gründen eine Trennung zwischen forschenden und beobachtenden Personen: Zum einen würden die Forschenden dadurch automatisch gezwungen sein, Kategorien und Skalen ausreichend klar zu formulieren, um sie für Dritte verständlich zu machen. Die Splittung von forschenden und beurteilenden Personen fordert eine Einführung von Einschulungen für die BeobachterInnen, welche nicht im Prozess der Erstellung des Beobachtungsinventars beteiligt waren.

Ein Beobachtungsmanual oder ein spezifisches Training unter Anleitung einer geschulten Person sind essentielle Bestandteile für die Voraussetzung einer standardisierten Erhebung der Daten und der Minimierung von Beobachtungsfehler und -verzerrungen. Die Qualität der Ergebnisse ist in einem großen Rahmen abhängig von der Qualität der BeobachterInneneinschulung, die das Ziel hat die BeobachterInnen bezüglich ihrer Aufgabe zu schulen und für das interessierende Verhalten zu sensibilisieren (Fieguth, 1977).

2.4.1 Beobachtungsmanual

Ein Manual oder eine Kodieranleitung mit eindeutigen Erklärungen hinsichtlich der unterschiedlichen Beobachtungsdimensionen und deren Differenzierung ist ein essentieller Bestandteil eines strukturierten Beobachtungsverfahrens. Bakeman und Quera (2011) zufolge soll ein Manual in erster Linie einen Überblick über den generellen Aufbau der Studie geben und die Struktur und Beschaffenheit der Kodierungseinheiten erläutern. Das Manual gilt als Regelwerk, nach welchem sich die BeobachterInnen während der Datenerhebung richten müssen. Die Verhaltensmerkmale, welche erfasst werden sollen, müssen im Manual definiert und mit klaren Anweisung der Erhebung konkretisiert werden (Döring et al., 2016). Es gibt genaue Anweisungen vor, in welchen Fällen eine Kodierung vorgenommen werden darf und wann nicht. Ein wichtiger Aspekt ist dabei Begrifflichkeiten einheitlich zu definieren und die einzelnen Skalen mit anschaulichen Ankerbeispielen zu versehen (König, 2010). Bakeman & Quera (2011) befürworten die zusätzliche Verwendung von Fotografien oder grafischen Modellen. Das Beobachtungsmanual dient auch vorrangig dazu, die Beobachtenden generell mit den Kodierregeln vertraut zu machen. Diese Regeln werden eingesetzt, um eine Vereinheitlichung der Protokollierung zu erzielen. Die Dauer der einzelnen Beobachtungseinheiten werden hierbei erläutert, sowie Informationen über Zahl der zulässigen Ko-

dierung innerhalb eines Beobachtungsintervalles oder die Reihenfolge der Kodierungen. Falls die Anwendung einer Protokollschrift notwendig ist, müssen den beobachtenden Personen die dafür verwendeten Kürzel beigebracht werden. Zusätzlich sollte einheitlich geregelt sein, wie die beurteilende Person mit unerwarteten Vorfällen, wie Unterbrechung der Beobachtungssituation durch eine Versuchsperson, umzugehen hat (Fieguth, 1977).

Das Erstellen des Kodiersystems und der Durchführung nach Vorgaben des Manuals sollte nach Bakeman & Quera (2011) parallel stattfinden, damit zeitgleich Verbesserungen durchgeführt werden können. Das Manual kann auch als Möglichkeit genutzt werden die Vorgehensweise des Entstehungsprozesses zu dokumentieren und somit für weitere Forschungsprojekte hilfreich sein.

2.4.2 BeobachterInnentraining

Die Anleitung zur richtigen Handhabung des Kategoriensystems kann wie soeben dargestellt, mithilfe eines umfassenden Manuals erfolgen. Bei komplexen Erhebungsformen ist es jedoch ratsam, zum besseren Verständnis des Ablaufs und der einzelnen Systeme, ein zusätzliches Training mit einer geschulten Fachkraft für die beobachtenden Personen abzuhalten (Döring et al., 2016). Eine professionelle Vorgehensweise besteht darin, dass im Falle mehrerer geschulter Personen, welche die Trainings zur Ausbildung weiterer BeobachterInnen durchführen, zur Qualitätssicherung die BeobachterInnenübereinstimmung erhoben wird. Dazu eignet sich besonders eine Kodierung einer auf Video aufgenommenen Situation mit anschließendem Vergleich und Diskussion. Um ein standardisiertes Vorgehen der Trainings sicherzustellen wird zudem wird empfohlen mögliche Teilschritte der Einschulung schriftlich festzuhalten und den Ablauf einheitlich zu gestalten (Fieguth, 1977).

Döring et al. (2016) raten zu BeobachterInnentrainings mit eingeschultem Personal als effektive Maßnahme gegen Beobachtungsfehler und Urteilsverzerrungen. Ihnen zufolge gibt es keine strikten Rahmenbedingungen, die eingehalten werden müssten. Stattdessen könnten die Trainings ähnlich wie Informationsveranstaltungen gestaltet werden und im Idealfall praktische Übungselemente und Information miteinander verbinden. Speziell betonen sie die Notwendigkeit bei Beobachtungseinschulungen Differenzierungsfähigkeiten zwischen den einzelnen Dimensionen des Regelwerks zu schulen und den Fokus auf essentielle Beobachtungspunkte zu richten, statt zu gene-

rell auf Probleme, wie Beobachtungsfehler einzugehen. BeobachterInneneinschulungen sollen an erster Stelle Informationen über die Studie und dessen Ziel vermitteln, sowie die BeobachterInnen mit den Durchführungs- und Erhebungsmethoden vertraut machen. Nachdem die TeilnehmerInnen der Einschulung den informativ-praktischen Teil absolviert haben, wird empfohlen mehrere Probekodierungen anhand von Videoaufzeichnungen oder Rollenspiele durchzuführen und anschließend zu besprechen. Individuelles Feedback und Erfahrungsaustausch ist dabei unerlässlich. Danach erfolgt eine Erhebung unter Ernstbedingungen mit anschließender Evaluation (Döring et al., 2016).

Ebenso betont Fisseni (1997) das Potential von Videoaufzeichnungen innerhalb von Schulungen der BeobachterInnen zu nutzen, da dies die Möglichkeit bietet Fehlerquellen zu demonstrieren, zu diskutieren und an einem „konkreten Fall“ zu korrigieren (Fisseni, 1997, S. 202).

Um ausreichende Qualität zu garantieren, sollte die abschließende Prüfung der Beobachtungsübereinstimmung bei quantitativen Beobachtungen mittels passender statistischen Methoden vorgenommen werden, bei qualitativen Beobachtungen erfolgt dies interpretativ (Döring et al., 2016). Fieguth (1977) betont zudem die Wichtigkeit von Nach- oder Begleittrainings, bei welchen der korrekte Umgang mit dem Messinstrument erneut thematisiert werden kann, sowie neu auftretende Fragen besprochen werden können. Diesen Aspekt hebt auch Schölmerich (2011) hervor, indem er eine kontinuierliche Kontrolle der Übereinstimmung im Prozess der Datenerhebung empfiehlt, um gegebenenfalls auf Abweichungen reagieren zu können.

Die Dauer einer BeobachterInneneinschulung scheint nicht mit der späteren Erhebungsgenauigkeit der Daten zusammenzuhängen. In einer Metaanalyse von Roch, Woehr, Mishra & Kieszczynska (2011) konnte kein positiver Effekt zwischen Länge der BeobachterInneneinschulung und Genauigkeit der Datenerhebung ermittelt werden. Fisseni & Fennekels (1995) zufolge wirkt sich jedoch ein intensiver Diskurs der Beobachterinnen über die Beobachtungseinheiten positiv auf die Übereinstimmung der Erhebung aus.

2.4.3 Trainingsansätze

BeobachterInnentrainings können unterschiedlich aufgebaut sein und die BeobachterInnen in verschiedenen Aspekten der Wahrnehmung schulen. Woehr und Huffcutt

(1994) klassifizieren vier verschiedene Typen von Einschulungen mit jeweils unterschiedlichen Trainingsansätzen und Schwerpunkten. Diese Klassifikation wurde nach einer von ihnen erhobenen Metaanalyse zur Effektivität von Trainings zur Leistungsbeurteilung durchgeführt (Renner, 2013).

BeurteilerInnenfehler-Training (rater error training): Im Verlauf des BeurteilerInnenfehler-Trainings werden die beobachtenden Personen über verschiedene Beobachtungsfehler wie Wahrnehmungs-, Erinnerungs-, oder Interpretationsfehler aufgeklärt. Ziel dieses Trainings ist es die Selbstwahrnehmung der Beobachtenden hinsichtlich Beobachtungsfehlern zu schulen und Fehler und Verzerrungen dadurch zu verringern, beziehungsweise ganz zu vermeiden.

Beurteilungsdimensionen-Training (performance dimension training): Das Beurteilungsdimensionen-Training setzt den Fokus darauf, die BeurteilerInnen möglichst genau in Beurteilungskategorien und deren korrekte Handhabung einzuarbeiten um dadurch eine exaktere Erhebung der Daten zu generieren. Ziel ist die bessere Erhebung „dimensionsrelevanter Beobachtungen statt eines eher homogenen Gesamteindrucks“ (Spinath & Becker, 2011, S.346). Im besten Falle wirken die BeobachterInnen direkt im Entwicklungsprozess der Beobachtungskategorien mit, andernfalls erfolgt ein durch geschultes Personal angeleitetes Training.

Bezugsrahmen-Training (frame of reference training): Das Bezugsrahmen-Training baut auf das Beurteilungsdimensionen-Training auf, allerdings werden zusätzlich verbindliche Beurteilungsstandards, beziehungsweise Referenzpunkte spezifiziert, mithilfe welcher eine Relativierung der beobachteten Verhaltensweisen ermöglicht wird. Ziel ist die Vereinheitlichung individueller Maßstäbe der BeobachterInnen durch die Definition eines spezifischen Beurteilungsrahmen in Bezug auf ein theoretisches Modell. Praxisübungen in Form von Vergleichsbeispiele, Selbstevaluation und Feedback stellen wichtige Punkte dieses Trainings dar, um personenspezifische Erhebungen der Beobachtungsdimensionen zu vermeiden (Obermann, 2018; Spinath & Becker, 2011).

Verhaltensbeobachtungstraining (behavioral observation training): Das Verhaltensbeobachtungstraining wird nach Woehr und Huffcutt (1994) als ein eigener Trainingsansatz erachtet, welcher durch eine strikte zeitliche Trennung von Beobachtung und Beurteilung gekennzeichnet ist. Das Beobachten erfordert das Identifizieren,

Wahrnehmen und Abrufen von beobachteten Verhaltensmerkmalen, während die Beurteilung anderer Verarbeitungsstrategien bedarf, wie die der Kategorisierung, Integration und Evaluation (Renner, 2013).

Allgemein fanden Woehr und Huffcutt (1994) in ihrer Metastudie heraus, dass je nach Fokus des Trainings, die Leistungen der BeobachterInnen vorwiegend im Bereich der damit verbundenen Schwerpunktsetzung verbessert wurden. Das bedeutet, dass durch ein Beurteilungsfehler-Training die Beurteilungsfehler verringert wurden, während ein Training zum Bezugsrahmen und zu den Beurteilungsdimensionen die Beurteilungsgenauigkeit erhöhte. Ein Verhaltensbeobachtungstraining führt zu einer allgemeinen Verbesserung der Beobachtungsgenauigkeit. Der Erfolg der Trainingsansätze wurde durch den Vergleich von Beobachtungsfehlern von BeurteilerInnen mit und ohne Absolvierung eines Trainings gemessen. Roch et al. (2011) führten eine aktualisierte Meta-Analyse, aufbauend auf der soeben beschriebenen Studie durch. Es zeigte sich, dass in der Forschung das Bezugsrahmen-Training die meistgewählte Methode für die Einschulung neuer BeurteilerInnen ist, da es im Vergleich zu anderen Trainingsmethoden den größten Effekt auf die Genauigkeit der Erhebung zeigt. Obwohl das Bezugsrahmen-Training vorrangig auf die Verbesserung der Genauigkeit der Beurteilungen abzielt, wurde auch im Zuge der Beobachtungsgenauigkeit ein positiver Effekt gefunden (Roch et al., 2012).

In einer weiteren Studie von Thornton & Zorich (1980) wurde evaluiert, ob sich die BeobachterInnenübereinstimmung im Zuge einer vertiefenden Einschulung der BeobachterInnen verbessern. Die Studienergebnisse ergaben, dass BeobachterInnen, welche einerseits speziell zu einer aufmerksamen Beobachtung und dem Kodieren von spezifischem Verhalten ermutigt wurden und zusätzlich in die Vermeidung von Systematischen Beobachtungsfehlern eingeführt wurden, die genauesten Verhaltensbeobachtungen durchführten, im Vergleich zu BeobachterInnen, welche nur Ermutigung zu einer aufmerksamen Datenerhebung erfuhren.

Abschließend kann festgestellt werden, dass die Art des Beobachtungstrainings, je nach Beobachtungsfehler, dessen Reduktion durch die Forschenden gewünscht wird, gewählt werden sollte. Eine Kombination von verschiedenen Trainings könnte ebenfalls zu einer Verbesserung der Leistung der Beobachtenden beitragen (Spinath & Becker, 2011). Allgemein kann zusammengefasst werden, dass geschulte BeobachterInnen, die ein intensives Beobachtungstraining absolvierten, besser in der Lage waren

das Messinstrument den Anforderungen entsprechend anzuwenden, als BeobachterInnen, welche keine oder nur sehr geringe Einschulungen hatten (Girard & Cohen, 2016)

2.5 ZUSAMMENFASSUNG

Die BeobachterInnenübereinstimmung ist eine der wichtigsten Methoden die Objektivität eines Verfahrens zu überprüfen und sollte vor der Veröffentlichung eines Inventars genauestens analysiert werden. Eine statistische Überprüfung der *Interrater-* und *Intra-Rater-Reliabilität* kann bei nominalskalierten Merkmalen über die Übereinstimmungsmaße der *Prozentualen Übereinstimmung* und *Cohens Kappa* erfolgen. Die Berechnung des Kappa-Koeffizienten kann von beobachterinnenunabhängigen Faktoren, wie Prävalenz eines Merkmals, Randsummenverteilungen in der Kontingenztafel oder der Kategorienanzahl beeinflusst werden.

Die Beurteilung eines Verhaltens in einer Beobachtungssituation sollte nicht willkürlich, sondern anhand eines standardisierten Vorgehens mit klaren Vorgaben erfolgen. Die Standardisierung der Erhebung erfolgt durch eine Einschulung der BeobachterInnen, in der Praxis werden dafür einerseits Manuale und BeobachterInnentrainings mit geschultem Personal eingesetzt. Ein hoher Grad an Übereinstimmungen zwischen BeobachterInnen ist einerseits abhängig von nachvollziehbaren und verständlichen Definitionen von Beobachtungseinheiten sowie einem verständlich gestalteten Manual oder einem gut durchgeführten Beobachtungstraining (Fisseni, 1997).

Es kann zwischen mehreren Trainingsansätzen unterschieden werden: *BeurteilerInnenfehler-*, *Beurteilungsdimensionen-*, *Bezugsrahmen-* und *Verhaltensbeobachtungstraining*, wobei das *Bezugsrahmentraining* besonders hohe Effektivität zur Verbesserung der Genauigkeit der BeurteilerInnen zeigt (Roch et al., 2012).

Eine gewissenhaft durchgeführte Schulung von BeobachterInnen ist ein effizientes Vorgehen gegen systematische Beobachtungsfehler und Verzerrungstendenzen. Beobachtungsfehler zu Lasten der Beobachtenden können im Wahrnehmungs-, Interpretations- und Erinnerungsspektrum auftreten. Zusätzlich können auch auf Ebene der Beobachtungssituation sowie zu Lasten des Beobachtungssystems Beobachtungsfehler die Ergebnisse verfälschen (Spinath & Becker, 2011). Da die Erfassung vielschichtiger Persönlichkeits- und Verhaltenscharakteristika ein umfassendes Erhebungsinstrument benötigen, ist ein BeobachterInnentraining, welches den BeurteilerInnen den Aufbau des Inventars und dessen Kategoriensysteme vermittelt, selbst wenn es

zu Beginn kostenintensiv und aufwendig erscheint, in den meisten Fällen notwendig und unumgänglich (Döring et al, 2016).

II. Empirischer Teil

3 HINTERGRUND & ZIELSETZUNG DER UNTERSUCHUNG

Das Videobeobachtungsinventar INTAKT (Hirschmann et al., 2017) erhebt die Beziehungsqualität zwischen Bezugsperson und Kind anhand der Interaktion während einer Bastel- und Spielsituation. Dabei werden auf verbaler und nonverbaler Ebene drei Beobachtungskategorien erfasst: *Feinfühligkeit*, *Rückmeldung* und *Joint Attention*.

Eine Frage, die sich bei der Entwicklung von Verhaltensbeobachtungsverfahren ergibt ist jene, ob die Erhebung durch eine Erörterung der Richtlinien anhand eines Manuals ausreichend reliable und valide Ergebnisse bringt, oder ob es zusätzlich einer spezifischen Einschulung der durchführenden Personen bedarf. Verschiedene Trainingsansätze zur Einschulung von Beobachtenden, sowie Vor- und Nachteile davon, wurden bereits in Kapitel 2.4 erarbeitet.

Im Rahmen dieser Arbeit wird anhand der Beobachtungsdimension *Joint Attention* untersucht, ob, beziehungsweise inwiefern, sich die Ergebnisse bei der Kodierung anhand der Vorgaben des INTAKT-Manuals (Hirschmann et al., 2017) von den Ergebnissen der Kodierung nach erfolgtem Training durch ein geschultes Fachpersonal unterscheiden. Das Ziel dieser Arbeit ist es, Unterstützung bei der Beantwortung der Frage zu bieten, ob eine Videokodierung mit Manual alleine, ohne zusätzlichem Training, ausreichend reliable und objektive Ergebnisse für das Verhaltensbeobachtungsinstrument INTAKT erbringt, oder ob eine zusätzliche Beobachtungseinschulung mit geschultem Personal die BeobachterInnenübereinstimmung verbessert. Diese Arbeit soll somit ein breiteres Wissen zum Thema Beobachtungstraining und dessen Auswirkung auf das Kodierverhalten der BeobachterInnen bringen und einen Beitrag in der Entwicklung des Beobachtungsinventars leisten.

3.1 FRAGESTELLUNG

Unter Berücksichtigung des derzeitigen wissenschaftlichen Diskurses und themenbezogener Literatur wurden folgende Fragestellungen überprüft:

Ist eine ausreichende Übereinstimmung der Kodiererergebnisse mit dem Manual der Beobachtungsdimension *Joint Attention* des Videobeobachtungsinstruments INTAKT im Vergleich zur Kodierung mit zusätzlichem Beobachtungstraining gegeben?

Für die Beantwortung der Fragestellung, ergeben sich unter Einbezug der *Interrater-Reliabilität* sowie der *Intra-Rater-Reliabilität* folgende weitere Fragestellungen:

1. Ist eine ausreichende Übereinstimmung der Kodiererergebnisse mit Manual im Vergleich zur Kodierung mit zusätzlichem Beobachtungstraining einer beobachtenden Person gegeben?

Im Zuge dieser Fragestellung werden Informationen über die *Intra-Rater-Reliabilität*, das bedeutet die Übereinstimmung der Ergebnisse und somit die Stabilität der Auswertung einer beobachtenden Person über mehrere Beobachtungszeitpunkte hinweg, erhoben.

2. Ist eine ausreichende BeobachterInnenübereinstimmung der Dimension *Joint Attention* zwischen zwei eingeschulten BeobachterInnen gegeben?

Diese Frage nimmt Bezug auf die *Interrater-Reliabilität*, die Übereinstimmung zwischen mindestens zwei unabhängigen BeobachterInnen. Die BeobachterInnenübereinstimmung gibt Auskunft über die Gütekriterien Reliabilität und Objektivität des Verfahrens. Es wird der Unterschied zwischen zwei Beobachtenden, wenn eine Beobachterin anhand des Regelwerks die Verhaltensmerkmale erfasst, die andere hingegen bereits ein Training absolvierte, sowie die Übereinstimmung nach beidseitig erfolgtem Training erfasst.

4 UNTERSUCHUNGSSINSTRUMENT: INTAKT

Das INTAKT-Videobeobachtungsinstrument ermöglicht eine Erhebung der Bindungsqualität zwischen Bezugsperson und Kind. Das Inventar bezieht sich in seinem inhaltlichen Konstrukt auf die Grundlagen der Bindungstheorie nach John Bowlby und Mary D. Ainsworth (Hirschmann et al., 2017).

Das Beobachtungsinstrument baut einerseits auf einer unstrukturierten Form auf, dessen Absicht es ist, im Rahmen einer Bastel- und Spielsituation die Interaktion zwischen Mutter und Kind möglichst uneingeschränkt und frei durch eine Videoaufzeichnung zu erheben, um eine detaillierte Analyse des verbalen und nonverbalen Interaktionsprozesses zu erhalten (Hirschmann et al., 2017). Gleichzeitig ermöglichen die strukturierte Datenerhebung mittels der *Time-Sampling-Methode*, sowie im Vorhinein festgelegte Materialien und Instruktionen, standardisierte Aussagen und Vergleiche. Im Rahmen der Beobachtung werden verbales und nonverbales Verhalten der Bezugsperson anhand der drei Dimensionen *Feinfühligkeit*, *Rückmeldung* und *Joint Attention* erfasst.

Die Kodierung der videogestützten Verhaltensbeobachtung kann entweder computergestützt oder händisch durchgeführt werden. Ersteres erfolgt unter Zuhilfenahme der Computersoftware INTERACT®. Die computergestützte Erhebung unterscheidet sich teilweise in ihren Unterkategorien sowie in der Zeitdauer der einzelnen Intervalle von der händischen Erhebung der Daten.

4.1 BEOBACHTUNGSSITUATION

Im Idealfall sollten die untersuchenden Personen während der videogestützten Verhaltensbeobachtungen anwesend sein, um auf mögliche Komplikationen zu reagieren oder Änderungen in den Kamerapositionen vornehmen zu können. Im Rahmen der Beobachtung gibt es keine festgelegte Zeitbeschränkung, jedoch wird im Manual empfohlen bei der Terminvereinbarung eine Dauer von zirka einer Stunde als Richtwert anzugeben. Die Videos wurden in den Wohnräumen der Mütter und deren Kindern erhoben. Den Müttern wurde vor der Durchführung erklärt, dass normales Spielverhalten beobachtet werden soll. Vor Beginn der Aufnahmen wurde eine Einverständniserklärung seitens der Mütter unterschrieben (Huber, 2013; Reischer, 2013; Vogler, 2015).

4.1.1 Die Bastelsituation

In der Bastelsituation werden Mutter und Kind dazu aufgefordert, gemeinsam mit vorgegebenen Materialien ein Haus zu basteln. Die standardisierte Instruktion wird dabei direkt an das Kind gerichtet, mit dem Hinweis, dass ihm die Mutter dabei helfen kann:

Schau, ich habe hier Bastelmaterial vorbereitet. Könntest du [Kind] dieses langweilige Haus [Untersucherin zeigt dabei auf die Vorlage] in ein schönes, buntes Haus verwandeln. Deine Mama kann dir dabei helfen und ihr könnt alles verwenden, was am Tisch liegt. Das Haus ist dann fertig, wenn [Name des Kindes] sagt, dass es fertig ist. Anschließend habe ich noch etwas für dich und deine Mama zum Spielen [Untersucherin zeigt dabei auf die Kiste] (Hirschmann et al., 2017, S. 7).

Die Dauer der Bastelsituation wird nicht durch das Manual vorgegeben, im Normalfall nimmt sie ca. 15-20 Minuten in Anspruch (Hirschmann et al., 2017).

4.1.2 Die freie Spielsituation

Die Spielsituation, welche in der Regel 20-30 Minuten beansprucht, wird ebenfalls mit festgelegtem Spielmaterial wie Lego-Bausteinen, Holzpuppen und -möbel und Spielautos durchgeführt (Hirschmann et al., 2017).

4.2 BEOBACHTUNGSKATEGORIEN

Die Qualität der Interaktion ausgehend von der Bezugsperson wird anhand der Kategorien *Feinfühligkeit*, *Rückmeldung* und *Joint Attention* beurteilt. Jede Dimension hat mehrere Unterkategorien, auf welche im Folgenden näher eingegangen wird. Genauere Beschreibungen zu den einzelnen Unterkategorien können im INTAKT-Manual (Hirschmann et al., 2017) nachgelesen werden.

Es ist möglich in jedem Kategoriensystem *Unkodierbar* zu kodieren, wenn die Mutter und/oder das Kind mindestens 50% des Intervalls nicht von der Kamera gefilmt wurde, andere Störfaktoren auftreten oder die Qualität des Videos keine Beurteilung zulässt. In diesen Fällen wird in der entsprechenden Zeile die Gesamtanzahl nicht erhebbarer Sekunden eingetragen (Hirschmann et al., 2017). Nachfolgend wird ein Überblick über die INTAKT-Kategoriensysteme gegeben.

4.2.1 Feinfühligkeit

Feinfühligkeit spielt eine bedeutsame Rolle im Aufbau einer sicheren Bindungsbeziehung zwischen Kind und Bezugsperson (Hirschmann et al., 2017). Die Dimension *Feinfühligkeit* des Beobachtungsinventars beruht auf dem Konzept von M. Ainsworth, welche im Rahmen zur Bindungsforschung die Wichtigkeit dieser Variable erkannte. *Mütterliche Feinfühligkeit* nach M. Ainsworth ist die Fähigkeit kindliche Signale und Bedürfnisse wahrzunehmen und angemessen und prompt auf diese zu reagieren (Ainsworth, Bell, & Stayton, 1974; Schmücker & Buchheim, 2002). Wie im INTAKT-Manual (Hirschmann et al., 2017) näher ausgeführt, stellt dies die Grundlage zur Beurteilung der mütterlichen Feinfühligkeit dar, welche von den Untersuchenden durch eine 7-stufige Ratingskala erfasst und folgendermaßen unterteilt wird:

- 1....sehr geringe Feinfühligkeit
 - 2....geringe Feinfühligkeit
 - 3....eher geringe Feinfühligkeit
 - 4....mittlere Feinfühligkeit
 - 5....eher hohe Feinfühligkeit
 - 6....hohe Feinfühligkeit
 - 7....sehr hohe Feinfühligkeit
- Unkodierbar-Feinfühligkeit

Die Ankerpunkte 1,3,5 und 7 sind im Manual detailliert beschrieben und bieten eine Orientierungshilfe für die Beobachtenden bei der Zuordnung des Verhaltens zur angemessenen Skalenstufe.

4.2.2 Rückmeldung

Dieses Kategoriensystem beruht auf der Annahme, dass die mütterliche Rückmeldung, ob neutral, lobend oder kritisierend ein wichtiger Anhaltspunkt für das Kind darstellt, um das eigene Tun auf zwischenmenschlicher Ebene einzuordnen und sich zukünftig in seinem Verhalten daran orientieren zu können (Trudewind, Unzner, & Schneider, 1997). Nonverbale Aspekte sind hierbei bei den Aussagen genauso zu beachten, wie der Tonfall der Stimme (Hirschmann et al., 2017).

Die Dimension *Rückmeldung* gliedert sich in die drei Unterkategorien: *Positive Rückmeldung*, *Korrigierende Rückmeldung* und *Negative Rückmeldung*. Zusätzlich gibt es noch die Kategorie *Unkodierbar-Rückmeldung*.

Eine Rückmeldung wird als positiv bewertet, wenn lobende oder positive Aussagen gemacht werden, welche zum Kind selbst oder auf dessen Tun gerichtet sind (Hirschmann et al., 2017). Eine *Korrigierende Rückmeldung* bedeutet, dass in der mütterlichen Äußerung eine hilfreiche Verbesserungsmöglichkeit der kindlichen Handlung oder dessen Handlungsergebnisses aufgezeigt wird und dazu beiträgt, Informationen an das Kind zu leiten oder Entscheidungshilfen zu geben. Ablehnende und emotional negative Äußerungen seitens der Bezugsperson, die auf Handlungen des Kindes oder auf das Kind als Person gerichtet sind, werden als *Negative Rückmeldung* kodiert (Hirschmann et al., 2017).

4.2.3 Joint Attention

Joint Attention beschreibt die Fokussierung der Aufmerksamkeit der Mutter und des Kindes auf eine gemeinsame Sache oder ein Objekt (Carpenter & Liebal, 2011).

Da sich diese Arbeit allein mit der händischen Erfassung der Joint Attention Dimension beschäftigt, wird auf diese näher eingegangen.

Das Kategoriensystem Joint Attention nützt die *Time-Sampling Methode*, um das Verhalten der Mütter quantitativ abzubilden. Die Dimension *Joint Attention* wird in zwei Kodierdurchgängen erfasst. Im ersten Kodierdurchgang wird die mütterliche *Joint Attention* mittels *Predominant-Activity-Sampling* anhand der Kategorien *Joint Attention Following*, *Switching Joint Attention*, *Keine Joint Attention* oder *Unkodierbar* erhoben. Diese Skalen schließen sich gegenseitig aus. Ein Intervall beträgt fünfzehn Sekunden. Sobald ein Verhalten mehr als die Hälfte der Zeit, also mindestens acht Sekunden auftritt und eindeutig einer Skala zugeordnet werden kann, erfolgt eine Kodierung innerhalb der betroffenen Skala.

Im zweiten Kodierdurchgang, das bedeutet bei wiederholter Ansicht des Videomaterials, erfolgt mittels *One-Zero-Sampling* die Erhebung von *Switching Joint Attention* und *Keine Joint Attention*. Sobald dieses Verhalten innerhalb eines zehn Sekunden Intervalls auftritt, wird dies im entsprechenden Intervall kodiert. Innerhalb eines Intervalls kann lediglich ein einziges Auftreten der *Switching Joint Attention* und ein Auftreten von *Keiner Joint Attention* kodiert werden, auch wenn das Verhalten öfter innerhalb dieser Zeitspanne gezeigt wird. Im Gegensatz zum *Predominant-Activity Sampling*

schließen sich die beiden Unterkategorien jedoch nicht gegenseitig aus. Wenn weder *Switching Joint Attention* noch *Keine Joint Attention* gezeigt wird, bleibt eine Kodierung im entsprechenden Intervall aus.

Anknüpfend wird auf die Richtlinien der Erfassung der einzelnen Unterkategorien der Beobachtungsdimension *Joint Attention* eingegangen (Hirschmann et al., 2017).

Attention Following: *Attention Following* findet statt, wenn die Bezugsperson durch Handlungen oder verbale Kommunikation aktiv an der Beschäftigung des Kindes teilnimmt. Wichtig ist dabei, dass die Mutter den Aufmerksamkeitsfokus des Kindes berücksichtigt und darauf eingeht. Dabei kann auch eine unterstützende Funktion eingenommen werden, indem dem Kind durch verbale und nonverbale Signale etwas erklärt, rückgemeldet oder vorgezeigt wird. Die Mutter kann sich aktiv oder passiv, z.B. durch bloßes Beobachten, an der Bastel- oder Spielsituation beteiligen.

Attention Switching: *Attention Switching* wird kodiert, wenn die Bezugsperson aktiv das Verhalten oder den Aufmerksamkeitsfokus des Kindes während einer laufenden Beschäftigung zu beeinflussen versucht. Diese Lenkung des Bastel- oder Spielgeschehens kann durch Befehle, Anweisungen oder Verbote erfolgen, mit dem Ziel die Handlung des Kindes zu korrigieren oder in eine bestimmte Richtung zu steuern. Dazu zählt auch die Manipulation einer Handlung, wie das Wegnehmen eines Spielmaterials oder ein Wechsel im Spielgeschehen, sowie der Beginn einer neuen Beschäftigung ohne Rücksicht auf das Interesse des Kindes zu nehmen.

Keine Joint Attention: Die Kodierung von *Keine Joint Attention* wird vorgenommen, wenn Bezugsperson und Kind ihre Aufmerksamkeit auf unterschiedliche Tätigkeiten oder Dinge gerichtet haben, z. B. beschäftigt sich die Mutter mit dem Einräumen vom Bastelmaterial, während das Kind bereits mit dem Spielmaterial spielt.

Unkodierbar: Ein Intervall wird als *Unkodierbar* gekennzeichnet, wenn in mehr als der Hälfte der Zeit Mutter und Kind nicht sichtbar sind, eine dritte Person die Situation stört oder aufgrund anderer Vorkommnisse eine Kodierung der *Joint Attention* nicht möglich ist.

Bei der computergestützten Erhebung der Beobachtungsdimension *Joint Attention*, gliedert sich die Kategorie *Joint Attention* bei der Erhebung mittels *Predominant-Activity-Sampling* in mehrere Unterkategorien, eine Unterteilung, welche bei händischer

Erhebung nicht vorgenommen wird. Folgendermaßen wird die Erhebung der Joint Attention unterteilt: *Aktive Aufrechterhaltung der Joint Attention Episode auf Handlungsebene*, *Aktive Aufrechterhaltung der Joint Attention Episode auf verbaler Ebene* und *Passive Aufrechterhaltung der Joint Attention Ebene*. Die Unterkategorie *Attention Switching* wird in *Aufmerksamkeitslenkung im laufenden Spiel* und *Aufmerksamkeitswechsel zu einem neuen/anderen Spiel* unterteilt.

5 METHODE

Für die Beantwortung der Fragestellung wurden 30 Videos, in denen Mutter-Kind-Interaktionen im Rahmen von Spiel- und Bastelsituationen gefilmt wurden, aus Master- und Diplomarbeiten herangezogen und kodiert (Huber, 2013; Reischer, 2013; Vogler, 2015). Die dabei erhobenen Beobachtungsdaten wurden in das Statistikprogramm *IBM SPSS® Statistics 24* eingegeben, die weitere Datenanalyse erfolgte ebenfalls mithilfe dieses Programms. Bei Berechnungen, welche händisch durchgeführt wurden, wird diese Information im Ergebnisteil zusätzlich angeführt. Informationen für die Datenmaske wurden zum Teil aus den bereits abgeschlossenen Erhebungen von Huber (2013), Reischer (2013) und Vogler (2015) übernommen.

5.1 DURCHFÜHRUNG DER ERHEBUNG

Für den Vergleich zwischen Kodierung des Videomaterials mithilfe des Manuals und zwischen Kodierung mit zusätzlichem Beobachtungstraining wurden 29 aus ursprünglich 30 Videos von zwei voneinander unabhängigen BeobachterInnen kodiert. Ein Video musste aufgrund zu häufiger Abwesenheit der Mutter in der Aufnahme aus der Erhebung ausgeschlossen werden.

Die Aufzeichnung der Verhaltensbeobachtung wurde den vorgegebenen Rahmen der Zeitintervalle angepasst, nach Beendigung eines Intervalls wurde das Video gestoppt und das Auftreten eines Ereignisses im Protokollbogen vermerkt. Bei Unsicherheiten wurde das Intervall wiederholt abgespielt. Jedes Video wurde pro Erhebungsphase zweimal kodiert, im ersten Kodierdurchgang wurde mittels *Predominant-Activity-Sampling (PAS)*, im zweiten Durchgang mittels *One-Zero-Sampling (OZS)* erhoben.

In der zweiten Erhebungsphase, in welcher die Videos wiederholt protokolliert wurden, wurde gleichermaßen vorgegangen. Jedes Video wurde somit doppelt kodiert, das bedeutet insgesamt liegen von 29 Videos 58 Erhebungsdaten vor.

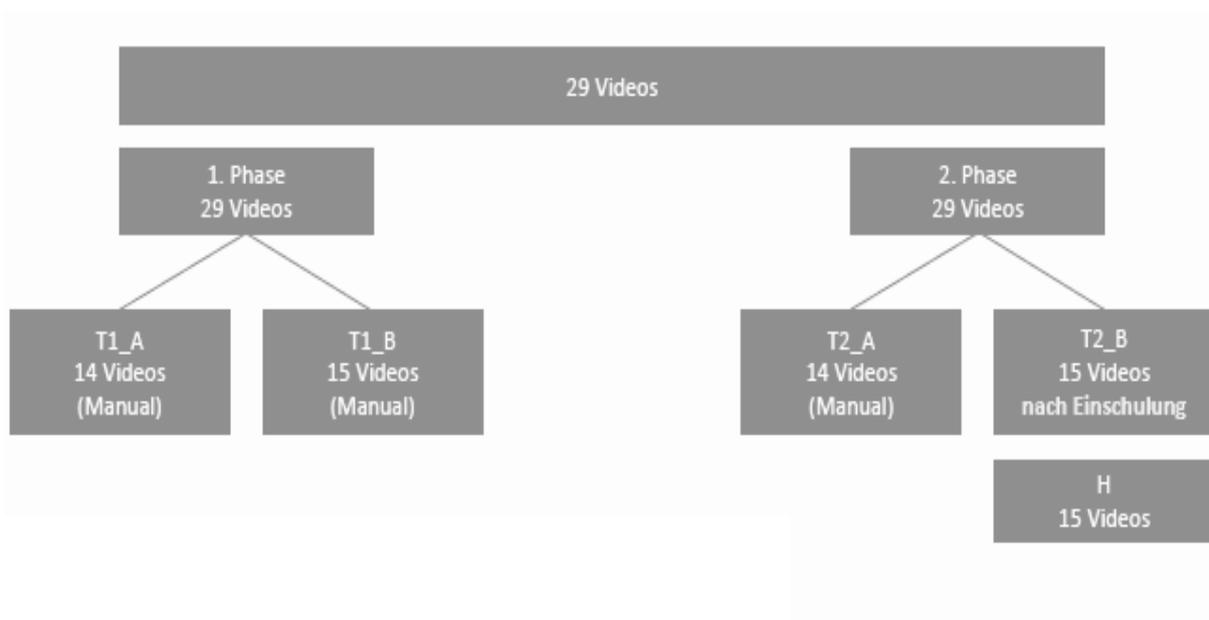
Im ersten Schritt wurden alle 29 Videos mithilfe des INTAKT-Manuals (Hirschmann et al., 2017) händisch kodiert. Diese Phase wurde in der weiteren statistischen Überprüfung in zwei Gruppen T1_A (14 Videos) und T1_B (15 Videos) eingeteilt. Um Lerneffekte zu kontrollieren und Aussagen über die Intra-Rater-Reliabilität zu ermöglichen,

Nachfolgend wird ein Überblick über die INTAKT-Kategoriensysteme gegeben. wurden in der ersten Gruppe der zweiten Erhebungsphase T2_A 14 Videos wiederholt kodiert.

Danach erfolgte ein BeobachterInnentraining zur regelkonformen Anwendung des INTAKT-Beobachtungsinventars und der Kodierung der *Joint Attention* Dimension. Nach Absolvierung des Beobachtungstrainings wurden die restlichen 15 Videos wiederholt kodiert (T2_B).

Abbildung 1 bildet die Aufteilung der einzelnen Erhebungsphasen ab und zeigt die Zuteilung der einzelnen Videos zu den für die Beantwortung der Fragestellung notwendigen Gruppen.

Abbildung 1. Erhebungsphasen der Untersuchung



T1_A, T1_B und T2_A wurden nur mit Manual kodiert, die Videos in der Gruppe T2_B wurden nach der Beobachtungseinschulung und der damit verbundenen Probekodierung mit Frau Mag. Dr. Hirschmann erhoben. Die Gruppen mit der Benennung T1 erfassen die erste Phase der Kodierung der Videos, in der Gruppe mit der Benennung T2 wurden die Videos in der zweiten Phase wiederholt kodiert.

Die Gruppen T1_A und T2_A schließen die Videos JV2, JV4, JV6, JV8, JV10, JV14, JV16, JV18, JV20, NR1, NR3, NR5, NR7, NR9 ein.

Die Gruppen T1_B und T2_B beinhalteten die Kodierung der Videos NR11, NR13, NR15, NR17, NR19, PH1, PH3, PH5, PH7, PH9, PH11, PH13, PH15, PH17, PH19.

5.1.1 INTAKT- Beobachtungstraining

Das BeobachterInnentraining wurde durch Frau Mag. Dr. Hirschmann, welche an der Entwicklung des INTAKT-Inventars maßgeblich beteiligt ist, in der ersten Oktoberwoche 2017 durchgeführt. Ziel war einerseits den richtigen Einsatz des Beobachtungsinventars zu garantieren, sowie ein verbessertes Verständnis des Konstrukts *Joint Attention* und eine ordnungsgemäße Handhabung der Beobachtungskategorien. Dafür wurde auf die Beobachtungskategorie *Joint Attention* allgemein eingegangen und die Richtlinien zur Erhebung der *Joint Attention* wurden ausführlich erläutert. Eine umfassende Differenzierung zwischen den Unterkategorien *Joint Attention Following (JA)*, *Joint Attention Switching (swJA)* und *Keine Joint Attention (kJA)* sollte eine präzisere Beobachtung und Beurteilung der Joint Attention-Merkmale garantieren sowie die Zuordnung zu den richtigen Beobachtungskategorien erleichtern.

Nach diesem Verständnisteil des Trainings wurde eine zirka 15-minütige Probekodierung durchgeführt. Im Nachhinein wurde ein Rückmeldegespräch vereinbart, um die Ergebnisse der Probekodierung zu besprechen, offene Fragen zu klären und Unklarheiten zu beseitigen. Die Datenerhebung anhand des Manuals (Erhebungsphase T1 und erste Hälfte der Videos von Erhebungsphase T2) wurde von Mai bis September kodiert, die Datenerhebung nach der Beobachtungseinschulung fand im Oktober statt.

5.1.2 Ablauf zur Auswertung der Daten

Um die Fragestellung beantworten zu können, ob das Kodieren mit Manual ohne zusätzlichem Training zu ähnlich reliablen Ergebnissen führt, wie die Kodierung nach einem zusätzlichen Beobachtungstraining durch eine geschulte Person, wurden im Zuge der Auswertung der Daten folgende Vergleiche angestellt:

1. Vergleich der Erhebungsdaten der Videos der Gruppen T1_A und T2_A

Für die Berechnung und die Interpretation der *Intra-Rater-Reliabilität* wurden die Ergebnisse der Erhebung der Videos der ersten Phase mit den Ergebnissen derselben Videos der zweiten Phase verglichen, mit der Absicht Informationen über mögliche Lerneffekte und Stabilität in der Kodierung zu generieren. Dabei wurden die Erhebungen von denselben 14 Videos der ersten (T1_A) und zweiten Phase (T2_A) verglichen, beide Male wurde nur mit Manual kodiert.

2. Vergleich der der Erhebungsdaten der Videos der Gruppen T1_B und T2_B

Im nächsten Schritt wurde die Datenerhebung der restlichen 15 Videos der ersten Phase (*T1_B*) mit jener von Phase 2 (*T2_B*) verglichen. Die Erhebungen unterschieden sich insofern, dass in der ersten Phase die Videos mit Manual kodiert wurden, während in der zweiten Phase die Beurteilerin ein zusätzliches Training durch ein geschultes Personal absolviert hatte. Das Ziel dieses Vergleichs war es, zu untersuchen, ob eine Veränderung in der Kodierung der Videos stattfinden würde, um eine Einschätzung über die Effektivität der Einschulung zu erhalten.

3. Vergleich der Erhebungsdaten der Videos der Gruppen *T2_B* und H

Für die Berechnung der Interrater-Reliabilität nach Absolvierung des Beobachtungstrainings beider Beobachterinnen, wurden die Ergebnisse der Video-Erhebung der Beobachterin T.P. der zweiten Erhebungsphase (*T2_B*) mit der Kodierung derselben Videos durch H.S. (H) verglichen (vgl. Schulz, in Vorbereitung).

4. Vergleich der Videos der Gruppen *T1_B* und H

Im zweiten Vergleich wurden die Ergebnisse der Datenerhebung der Videos mit Manual von T.P. mit den Erhebungen derselben Videos nach erfolgtem BeobachterInnen-training H.S. verglichen. Es fand somit ein Vergleich der Datenerhebung der Beobachtenden mit und ohne BeobachterInnen-training statt.

5.1.3 Berechnung der BeobachterInnenübereinstimmung

Wirtz und Caspar (2002) betonen, dass es nicht einen einzelnen Koeffizienten gibt, welcher die Übereinstimmung zwischen BeobachterInnen am besten abbilden kann, da die Koeffizienten mit den Eigenschaften der Datenstruktur in Verbindung stehen und diese abbilden. Sie raten dazu, mehrere Koeffizienten zu berechnen und diese in der Interpretation miteinander in Bezug zu stellen.

Für die Berechnung der BeobachterInnenübereinstimmung wurden die einzelnen Unterkategorien der Beobachtungsdimension *Joint Attention* als nominalskaliert angenommen. Aufgrund des Skalenniveaus und der Anzahl der beobachtenden Personen ($N=2$) wurde das statistische Maß von *Cohens Kappa* für die Berechnung der BeobachterInnenübereinstimmung herangezogen. Da das *Cohens Kappa* durch einige Faktoren limitiert ist, auf welche in Kapitel 2.1.2 näher eingegangen wurde, wurde zusätzlich die *Prozentuale Übereinstimmung* berechnet um die Interpretation der Übereinstimmung zu erleichtern.

Prozentuale Übereinstimmung: Die *Prozentuale Übereinstimmung* ist ein Übereinstimmungsmaß, welches bei nominalskalierten Kategorien ermittelt werden kann und den prozentualen Anteil jener Fälle angibt, in welchen die BeobachterInnen eine kongruente Entscheidung hinsichtlich eines Merkmals getroffen haben (Seidel & Prenzel, 2010). Die Methode zur Schätzung von Übereinstimmung zwischen Beobachtenden berücksichtigt jedoch nicht die Möglichkeit zufällig entstandener Übereinstimmungen (Wirtz & Caspar, 2002). Die Prozentuale Übereinstimmung wurde händisch unter Verwendung der Kontingenztabelle (siehe Anhang) mithilfe folgender Formel (nach Wirtz & Caspar, 2002; S. 47) berechnet:

$$\text{Proz. } \ddot{U} = \frac{\text{Anzahl der von den BeobachterInnen gleich beurteilten Merkmale}}{\text{Anzahl aller Merkmale}} * 100\%$$

Der Wertebereich der Prozentualen Übereinstimmung liegt zwischen 0% und 100%. Je höher der Wert, desto besser die Übereinstimmung zwischen den Beurteilenden. Bei direkt beobachtbaren Variablen empfehlen Döring et al. (2016) einen Wert nahe 100% als Indiz für eine gute Übereinstimmung zwischen den beurteilenden Personen, McHugh (2012) zufolge, sollte für eine gute BeobachterInnenübereinstimmung eine Prozentuale Übereinstimmung von mindestens 80% erreicht werden.

Cohens Kappa: Cohens Kappa ist ein zufallsbereinigtes Übereinstimmungsmaß, welches bei normalskalierten Kategorien die Übereinstimmung der Erhebungen von zwei BeobachterInnen erfasst (Bakeman & Quera, 2011). Die Berechnung erfolgte unter Verwendung der Statistiksoftware SPSS mithilfe von Kontingenztabelle, welche im Anhang abgebildet sind.

Der Kappa-Koeffizient kann einen Wert zwischen -1 und +1 annehmen, wobei der Wert +1 einer perfekten Konkordanz entspricht. Ein Kappa-Koeffizient von 0 bedeutet, dass die beobachtete Übereinstimmung nicht besser als eine durch Zufall erwartete Übereinstimmung ist. Der Wert -1 gibt an, dass keine Übereinstimmungen zwischen den BeobachterInnen vorhanden sind (Bakeman & Quera, 2011). Werte über dem Cutoff $\kappa = .75$ deuten auf eine sehr gute Übereinstimmung der beobachtenden Personen hin. Werte zwischen $\kappa = .6$ und $\kappa = .75$ gelten als gute Übereinstimmung der BeobachterInnen. Als mittelmäßige BeobachterInnenübereinstimmung gelten Kappa-Werte zwischen $\kappa = .4$ und $\kappa = .6$. Ein Koeffizient $\kappa < .4$ deutet auf eine geringe, nicht zufriedenstellende Übereinstimmung der Erhebungen der BeurteilerInnen hin (Döring et al., 2016).

6 STICHPROBE

Die Erhebung der Daten erfolgte aus insgesamt 29 Aufnahmen, welche im Rahmen von Diplom- und Masterarbeiten der Universität Wien von Huber (N=15), Reischer (N=15) und Vogler (N=14) erhoben wurden (Huber, 2013; Reischer, 2013; Vogler, 2015). Die Stichprobe der kodierten Videos bestand aus insgesamt 29 Müttern und deren leiblichen Kinder.

Die Videoaufzeichnungen beinhalteten, wie im INTAKT-Inventar vorgegeben allesamt eine Bastel- und eine Spielsituation. Durchschnittlich dauerten die Videos 47:48 Minuten (SD=12.2), das längste Video hat eine Aufnahmezeit von 72:52 Minuten, das kürzeste Video endet bereits nach 27:36 Minuten. Die genaue Darstellung der Dauer der einzelnen Videos kann Anhang I entnommen werden.

Die Stichprobe für die Videos setzt sich aus 29 Müttern zusammen, welche in Interaktion mit ihren 3- bis 6-jährigen Kindern (M=3.81 Jahre; SD=0.8) gefilmt wurden. Von den 29 Kindern waren 69% weiblich (N= 20) und 31% männlich (N= 9). Alle 29 Kinder hatten Deutsch als Muttersprache, drei dieser Kinder wurden zweisprachig erzogen. Insgesamt besuchten 28 der 29 Kinder einen Kindergarten, 31% ganztags (N=9), 65.5% halbtags (N=19), nur ein Kind (3.4%) war nicht in einem Kindergarten gemeldet.

Die Mütter waren zwischen 25 und 50 Jahre alt (M=34; SD=5.9). 37.9% (N=11) der Mütter hatten einen Hochschulabschluss als höchste abgeschlossene Ausbildung, 27.6% (N=8) hatten einen Maturaabschluss und 31% (N=9) hatten eine Lehre oder eine Fachschule absolviert, lediglich eine Person (3.4%) hatte einen Pflichtschulabschluss. Zum Zeitpunkt der Erhebung waren 82.8% der Mütter berufstätig (N=24), 17,2% Mütter (N=5) gaben an keine Berufstätigkeit auszuüben. Alle Mütter wohnten mit dem Kind im selben Haushalt, welcher sich entweder in Wien, Niederösterreich oder Salzburg befand.

7 ERGEBNISSE

Die Ergebnisse der Untersuchung zur Fragestellung nach Unterschieden in der BeobachterInnenübereinstimmung bei händischer Kodierung mit Manual, im Vergleich zu händischer Kodierung mit zusätzlichem Beobachtungstraining, werden in diesem Kapitel dargestellt. Für die Beantwortung der Frage wurde die Intra-Rater-Reliabilität, sowie die Interrater-Reliabilität berechnet. Die Ergebnisse werden in den meisten Fällen getrennt nach Kodierdurchgang 1 (*PAS*) und Kodierdurchgang 2 (*OZS*) dargestellt. Im Diskussionsteil dieser Arbeit (Kap.8) werden die in diesem Kapitel präsentierten Ergebnisse zueinander in Bezug gestellt, analysiert und interpretiert.

Zu Beginn der statistischen Analyse wurde die Kategorie *Unkodierbar* des ersten Kodierdurchgangs zu *Joint Attention* aufgrund fehlender inhaltlicher Relevanz in Bezug auf die Fragestellung der Arbeit ausgeschlossen und dementsprechend nicht mehr in weiteren Analysen berücksichtigt. Für die Beantwortung der Fragestellung, inwiefern die Ergebnisse der Datenerhebung anhand des Manuals mit den Ergebnissen nach erfolgtem Training übereinstimmen, wurden vor Beginn der statistischen Analysen deskriptive Statistiken durchgeführt, um einen Überblick über die Zusammensetzung der Daten zu erhalten und spätere Analysen darauf abzustimmen. Die Ergebnisse können Anhang II entnommen werden.

7.1 BERECHNUNG DER INTRA-RATER- UND INTERRATER RELIABILITÄT

7.1.1 Intra-Rater-Reliabilität

In den folgenden Unterkapiteln wird auf die Übereinstimmung der Ergebnisse von Erhebungsphase 1 und 2 der Beobachterin T.P. eingegangen. Ersteres wird die Übereinstimmung zwischen den Erhebungen mit Manual verglichen, danach folgt die Darstellung der Intra-Rater-Reliabilität Ergebnisse der Erhebung mit Manual im Vergleich zur Erhebung mit BeobachterInnentraining. Die Interpretation des Kappa-Koeffizienten bezieht sich auf die von Döring et al. (2016) empfohlenen Grenzwerte, die Interpretation der Prozentualen Übereinstimmung richtet sich ebenfalls nach der Empfehlung von Döring et al. (2016), welche beide in Kap. 5.1.3. nachzulesen sind.

7.1.1.1 Übereinstimmung: Manual

In Tabelle 1 sind die Häufigkeitsverteilungen der Beobachtungsdimension *Joint Attention* von Erhebungsphase 1 & 2 dargestellt. Im Rahmen des ersten Kodierdurchgangs wurden die Daten mittels *Predominant-Activity-Sampling (PAS)* erhoben. Im zweiten Durchgang mittels *One-Zero-Time-Sampling (OZS)*. Zusätzlich ist die *Absolute Differenz* und die *Relative Differenz* zwischen den Werten der Gruppe A der ersten Erhebungsphase mit Manual und der zweiten Erhebungsphase mit Manual (T1_A & T2_A) der Beobachterin T.P. dargestellt.

Tabelle 1: Beurteilung mit Manual: Häufigkeiten und Anteilswerte der Beobachtungskategorien

T1_A & T2_A	T1_A		T2_A		T1_A & T2_A		
	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Differenz	Differenz der relativen Häufigkeit (%)	Relative Differenz (%)
Kodierdurchgang 1: Joint Attention Following							
Joint Attention Following	2453	91.4	2480	92.3	27	0.9	1.1
Switching Joint Attention	184	6.9	166	6.2	18	0.7	9.8
Keine Joint Attention	48	1.8	40	1.5	8	0.3	16.7
Kodierdurchgang 2: Switching Joint Attention							
swJA	868	21.7	823	20.4	45	1.3	5.7
Kodierdurchgang 2: Keine Joint Attention							
kJA	114	2.8	84	2.1	30	0.8	26.3

Kodierdurchgang 1: Joint Attention Following

Die Tabelle zeigt, dass es zwischen erster Erhebungsphase (T1_A) und wiederholtem Erheben desselben Videomaterials (T2_A) bei der Kategorie *Joint Attention Following* nur einen relativen Anstieg um 1.1% gab. Gleichzeitig nahmen die damit in Verbindung stehenden Werte *Switching Joint Attention* und *Keine Joint Attention* in Relation um 9.8% bzw. 16.7% ab.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Unter Beobachtung mit Manual erfolgte zwischen den Gruppen T1_A und T2_A eine Reduktion der Kodierung von *Switching Joint Attention-Events* von 21.7% auf 19.9%, eine relative Differenz von 5.7%. Die Ergebnisse des Auftretens von *Keine Joint Attention* Events zeigen einen relativen Rückgang von 26.3% zwischen erster und wiederholter Erhebung der Events.

Tabelle 2 zeigt den Wert des berechneten Cohens Kappa, sowie die Prozentuale Übereinstimmung der Übereinstimmung zwischen den einzelnen Videos der Gruppen T1_A & T2_A, welche beide von T.P. mit Manual kodiert wurden. Die Ergebnisse beziehen sich auf die Übereinstimmung der Kodierungen zwischen Phase 1 und Phase 2 in Bezug auf beide Kodierdurchgänge. K1_JA bildet den Kodierdurchgang 1 (PAS) ab, K2_swJA (*Switching Joint Attention*) und K2_kJA (*Keine Joint Attention*) sind Teil des Kodierdurchgangs 2 (OZS).

Tabelle 2: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T1_A & T2_A (beide mit Manual kodiert)

Video	K1_JA		K2_swJA		K2_kJA	
	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %
JV2	.303	92.9	.457	85.5	.282	98.5
JV4	.556	90.8	.7	89.7	.495	99.1
JV6	.602	94.2	.847	94.8	-	100
JV8	.876	95.8	.839	96.8	.793	92.1
JV10	.399	94.2	.85	94.7	.798	99.5
JV14	.484	92.2	.831	93.6	-	99.4
JV16	.660	98.5	.6	90.5	.491	98.0
JV18	.431	92.4	.831	92.7	-	99.4
JV20	.472	94.6	.697	92.0	-	100
NR1	.506	92.8	.734	92.0	.951	99.7
NR3	-.005	99.0	.697	94.9	-	99.8
NR5	.632	89.6	.791	89.9	1	100
NR7	.563	89.8	.778	89.5	.397	99.1
NR9	.488	97.4	.792	94.1	-	100
Insgesamt	.607	94.0	.766	92.2	.791	99.0

Kodierdurchgang 1: Joint Attention Following

Der berechnete Cohens Kappa Wert aller Videos des ersten Kodierdurchgangs beträgt $\kappa=.607$ und ist somit ein Indiz für eine gute Übereinstimmung der Beobachterin zwischen erster und zweiter Erhebungsphase. Die Prozentuale Übereinstimmung deutet mit 94% auf eine sehr gute Übereinstimmung der Beobachterin hin.

In K1_JA wurde die niedrigste übereinstimmende Kodierung bei Video JV2, mit einem Cohens-Kappa Wert von $\kappa=.303$ berechnet, den Grenzwerten zufolge eine mangelnde Übereinstimmung. Die höchste Übereinstimmung betrug $\kappa=.876$ (Video JV8), ein Wert, welcher einer sehr guten Beobachtungsübereinstimmung entspricht. Das Video NR3 ergab mit einem Kappa-Koeffizienten von $\kappa= -.005$ keine Übereinstimmung in den Datenerhebungen, welche besser ist, als durch Zufall erwartet worden wäre. Auffällig ist, dass die Prozentuale Übereinstimmung bei demselben Video die höchste der gesamten Gruppe ist und bei 99% liegt. Generell ist erkennbar, dass das Video JV2 mit dem niedrigsten Kappa-Koeffizienten nicht jenes mit der niedrigsten Prozentualen Übereinstimmung ist. Insgesamt liegt die Prozentuale Übereinstimmung in allen Videos über 89%.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Bei der Kodierung der *Switching Joint Attention* wurde ein Kappa-Wert $\kappa=.766$ erzielt, auch die Beurteilung der Kategorie *Keine Joint Attention* war mit $\kappa=.791$ ähnlich hoch. Die Ergebnisse beider Kategorien des zweiten Kodierdurchgangs zeigen somit eine sehr gute Beobachtungsübereinstimmung. Die Prozentuale Übereinstimmung mit Werten über 92% deutet ebenso auf eine hohe Übereinstimmung der Beobachterin zwischen erster und zweiter Erhebungsphase hin.

In K2_swJA ist die niedrigste Übereinstimmung bei Erhebung des Videos JV2 zu finden ($\kappa=.457$). Dieser Wert ist als mittelmäßige BeobachterInnenübereinstimmung interpretierbar. Gleichzeitig ist im selben Video auch die Prozentuale Übereinstimmung mit 85.5% die niedrigste. Der höchste Cohens Kappa Wert $\kappa=.847$ wurde im Video JV6 erreicht. Dieser Wert weist auf eine sehr gute Übereinstimmung hin. Die höchste Prozentuale Übereinstimmung wurde mit 96.8% im Video JV8 erzielt.

In K2_kJA konnten in sechs der insgesamt vierzehn Videos keine Cohens Kappa Werte berechnet werden. Die Ursache dafür, dass im Rahmen der Videoaufnahmen keine Berechnung des Cohens Kappa möglich war, liegt einerseits darin, dass generell weder in der ersten, noch in der zweiten Erhebungsphase ein Auftreten der Kategorie *Keine Joint Attention* kodiert wurde. Das Auftreten dieses Falles ist in der Tabelle

dadurch erkennbar, dass der Kappa-Koeffizient nicht berechnet werden konnte, die Prozentuale Übereinstimmung jedoch 100% beträgt.

Ein weiterer Grund für die fehlende Berechnung des Cohens Kappa liegt der Tatsache zugrunde, dass in nur einer Erhebungsphase ein Auftreten des Events *Keine Joint Attention* vorgenommen wurde, für dasselbe Video bei wiederholtem Beobachten in Erhebungsphase 2 jedoch kein einziges Mal ein Auftreten der Kategorie *Keine Joint Attention* vermerkt wurde. Dieser Umstand wird in der Tabelle deutlich, wenn kein Cohens Kappa Wert durch SPSS berechnet wurde und in derselben Spalte die Prozentuale Übereinstimmung einen Wert unter 100% anzeigt. Details zu den Kontingenztabellen der einzelnen Videos dieser Kategorie können Anhang IX entnommen werden. Die niedrigste berechenbare Übereinstimmung beträgt $\kappa=.282$ (Video JV2), was auf eine sehr geringe Übereinstimmung schließen lässt. Der höchste Kappa-Koeffizient wurde bei Erhebung des Videos JV10 mit $\kappa=.798$ berechnet und deutet auf eine sehr gute Übereinstimmung der Beobachterin zwischen erster und zweiter Erhebungsphase hin.

7.1.1.2 Übereinstimmung: Manual und Beobachtungstraining

Tabelle 3 zeigt die Häufigkeitsverteilungen der Beobachtungsdimension *Joint Attention* bei der Erhebung der Daten durch Beobachterin T.P., in der ersten Phase mit Manual (T1_B) und in der zweiten Phase mit BeobachterInnentraining (T2_B).

Tabelle 3: Häufigkeiten und Anteilswerte der Beobachtungskategorien: Beurteilung mit Manual (T1_B) im Vergleich zu Beurteilung mit Training (T2_B)

T1_B & T2_B	T1_B		T2_B		T1_B & T2_B		
	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Differenz	Differenz der relativen Häufigkeit (%)	Relative Differenz (%)
Kodierdurchgang 1: Joint Attention Following							
Joint Attention Following	2471	90.6	2555	93.7	84	2.1	3.3
Switching Joint Attention	237	8.7	149	5.5	88	3.2	37.1
Keine Joint Attention	19	0.7	24	0.9	5	0.2	20.8
Kodierdurchgang 2: Switching Joint Attention							
SwJA	1199	29.2	1160	28.3	39	0.9	3.3
Kodierdurchgang 2: Keine Joint Attention							
kJA	101	2.5	57	1.4	56	1.1	43.6

Kodierdurchgang 1: Joint Attention Following

Es kann ein leichter Anstieg in der Kodierung von *Joint Attention Following* von der ersten (T1_B) zur zweiten Erhebungsphase (T2_B) der Videos beobachtet werden. Die Kodierung der Aufrechterhaltung der *Joint Attention* nach der Beobachtungseinschulung (T2_B) stieg um 2.1%, von 90.6% auf 93.7%, die relative Differenz beträgt 3.3%. Nach der Beobachtungseinschulung (T2_B) gibt es im Vergleich zu der Kodierung mit Manual in der Unterkategorie *Switching Joint Attention* eine relative Abnahme in der Erhebung dieser um 37.1%. Die Kategorie *Keine Joint Attention* zeigte nach der Beobachtungseinschulung einen Abfall von 20.8% in der Protokollierung.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Die Häufigkeitsverteilung der Skala *Keine Joint Attention* des *One-Zero-Samplings*, zeigt lediglich einen minimalen Unterschied von 3.3%, bei der Erhebung der Kategorie *Keine Joint Attention* gab es jedoch einen Rückgang um 43.6%.

Die Übereinstimmungen zwischen den Videos, welche mithilfe des Manuals erhoben wurden und jenen welche nach Beobachtungstraining kodiert wurden, werden in Tabelle 4 dargestellt.

Tabelle 4: Intra-Rater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T1_B & T2_B (T1_B mit Manual kodiert; T2_B mit zusätzlichem Beobachtungstraining)

Video	K1_JA		K2_swJA		K2_kJA	
	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %
NR11	.484	96.6	.84	93.7	.659	98,3
NR13	.616	96.5	.862	95.1	-	100
NR15	.645	90.6	.868	94.6	.832	96.6
NR17	.761	984	.837	94.2	-	100
NR19	.703	97.5	.784	92.2	-	100
PH1	.718	98.0	840	94.2	-	99.6
PH3	389	96.3	.846	94.2	.457	94.6
PH5	1	100	.829	96.0	-	100
PH7	.532	97.2	822	94.1	-	99.6
PH9	.778	96.0	.858	93.7	.799	99.8
PH11	.621	92.6	.836	92.1	.831	99.5
PH13	.649	89.9	784	89.6	.394	98.2
PH15	.644	89.6	.787	89.7	.432	96.5
PH17	.362	89.0	.818	91.9	.887	99.6
PH19	.615	91.8	.797	91.6	.579	97.7
Insgesamt	.642	94.8	.816	92.5	.63	98.6

Kodierdurchgang 1: Joint Attention Following

Die Übereinstimmung $\kappa=.642$ aller Videos der Gruppe B zeigt eine gute Übereinstimmung des ersten Kodierdurchgangs zwischen der Erhebung mit Manual und der Erhebung mit zusätzlicher Beobachtungseinschulung.

Im ersten Kodierdurchgang zeigte sich die niedrigste Übereinstimmung im Video PH17 mit einem Kappa-Wert $\kappa=.362$ und einer prozentualen Übereinstimmung von 89%. Der Kappa-Koeffizient deutet auf eine unzureichende Übereinstimmung der Beurteilung

hin, während die prozentuale Übereinstimmung auf eine hohe Übereinstimmung schließen lässt. Bei der Kodierung des Videos PH5 konnte zwischen erstem und zweitem Erhebungsdurchgang eine absolut kongruente Übereinstimmung mit einem Cohens Kappa Wert $\kappa=1$ und einer prozentualen Übereinstimmung von 100% erzielt werden.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Eine sehr gute Übereinstimmung zeigt die Kategorie *Switching Joint Attention* ($\kappa=.816$; Proz.Ü=92.5%) hinsichtlich der Kodierung mit und ohne Beobachtungseinschulung. Auch das Ergebnis der Kategorie *Keine Joint Attention* weist mit einem Wert $\kappa=.63$ auf eine gute Übereinstimmung hin. Die Prozentuale Übereinstimmung ist wie in Tabelle 4 zu sehen, über alle Kategorien sehr hoch, auffällig ist jedoch, dass die Kategorie *Keine Joint Attention* mit dem niedrigsten Cohens Kappa-Koeffizienten, die höchste prozentuale Übereinstimmung hat ($\kappa=.63$; Proz.Ü.=98.6%).

Die Ergebnisse der Kategorie *Switching Joint Attention* des zweiten Kodierdurchgangs zeigen durchgehend sehr gute Übereinstimmungen mit Kappa Werten zwischen $\kappa=.784$ (Video PH13) und $\kappa=.868$ (Video NR15). Die prozentuale Übereinstimmung deutet ebenfalls mit Werten $>89\%$ auf eine hohe Übereinstimmung zwischen der Erhebung mit Manual und der Erhebung mit zusätzlicher Beobachtungseinschulung hin. In der Kategorie *Keine Joint Attention* des zweiten Kodierdurchgangs ist wieder auffällig, dass bei einigen Videos keine Kappa-Werte berechnet werden konnten. Die Gründe dafür sind dieselben, wie bereits zuvor bei Berechnung der Intra-Rater-Reliabilität mit Manual (siehe Erklärung Tabelle 3) erklärt. Die Übereinstimmungen der Erhebungen sind breit gefächert, PH13 ist mit einem Kappa-Koeffizienten von $\kappa=.394$ sehr gering, während die Übereinstimmung in der Kodierung des Videos PH17 mit einem Cohens Kappa $\kappa=.887$ sehr hoch ist. Ebenfalls deutlich sichtbar ist der Unterschied zwischen niedrigen Kappa-Werten und sehr hohen Ausprägungen der Prozentualen Übereinstimmung mit Werten ab 94.6% (Video PH3). Zusätzlich ist in der Tabelle auch erkennbar, dass ein niedriger Kappa-Wert der Kategorie *Keine Joint Attention* in vielen Fällen nicht mit niedrigen Werten der Prozentualen Übereinstimmung einhergeht (z.B. Video PH13: $\kappa=.394$; Proz.Ü=98.2%).

In Tabelle 5 wird die Übereinstimmung der Beobachterin T.P. über beide Kodierdurchgänge (PAS & OZS) der Beobachtungskategorie *Joint Attention* zusammengefasst.

Tabelle 5: Intra-Rater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung der Beobachtungsdimension *Joint Attention* insgesamt & der Kategorien JA, swJA, kJA

	Kappa	Prozentuale Übereinstimmung (%)
Insgesamt (K1 & K2)		
T1_A & T2_A	.896	95.2
T1_B & T2_B	.902	95.3
K1_JA		
T1_A & T2_A	.607	94.0
T1_B & T2_B	.642	94.8
K2_swJA		
T1_A & T2_A	.766	92.2
T1_B & T2_B	.816	92.5
K2_kJA		
T1_A & T2_A	.791	99.0
T1_B & T2_B	.63	98.6

Die Erhebung der Daten mit Manual (T1_A & T2_A) deuten mit $\kappa=.866$ & Proz.Ü.=95.2% auf eine sehr gute Übereinstimmung der Beobachterin T.P. im Vergleich der ersten und zweiten Phase hin. Die Übereinstimmung der Beurteilungen der Videos welche mit Manual verglichen wurden, und jenen welche mit zusätzlicher Einschulung erhoben wurden, ist mit $\kappa=.902$ & Proz.Ü.=95.3% gering noch höher. Es kann auf eine sehr gute Intra-Rater-Reliabilität geschlossen werden. Die Prozentuale Übereinstimmung ist in beiden Fällen mit Ergebnisses <95% ebenfalls sehr hoch. Die Kategorie *Keine Joint Attention* (K2) erreichte eine höhere Beobachtungsübereinstimmung bei Erfassung der Merkmale anhand der Vorgaben des Manuals (T1_A & T2_A).

Die Erhebung des *Joint Attention Following* im ersten Kodierdurchgang, sowie die Kategorie *Switching Joint Attention* des zweiten Kodierdurchgangs, erzielten eine höhere Übereinstimmung bei der Kodierung der Videos, welche in der ersten Erhebungsphase mit Manual und in der zweiten mit zusätzlichem Beobachtungstraining kodiert wurden (T1_B & T2_B).

7.1.2 Interrater-Reliabilität

In den folgenden Unterkapiteln wird auf die Übereinstimmung der Ergebnisse zwischen den Beobachterinnen T.P. und H.S. eingegangen. Im ersten Schritt wird auf die Übereinstimmung der Erhebung mit zusätzlichem Beobachtungstraining beider Beobachtenden eingegangen. Danach wird die Übereinstimmung der BeobachterInnen zwischen der Erhebung mit Manual seitens T.P. und der Erhebung mit Beobachtungstraining seitens H.S. verglichen. Die Interpretation des Cohens Kappa bezieht sich auf die von Döring et al. (2016) empfohlenen Grenzwerte, die Interpretation der Prozentualen Übereinstimmung richtet sich ebenfalls nach der Empfehlung von Döring et al. (2016), welche beide in Kap. 6.1.3. nachzulesen sind.

7.1.2.1 Übereinstimmung: Beobachtungseinschulung

Tabelle 6 gibt die Häufigkeitsverteilungen der Beobachtungsdimension *Joint Attention*, sowie die berechnete Differenz zwischen beiden Beobachterinnen, T.P. & H.S., nach erfolgtem Beobachtungstraining wieder.

Tabelle 6: Häufigkeiten und Anteilswerte der Beobachtungskategorien: Beurteilung mit Training (T2_B & H)

T2_B & H	T2_B		H		T2_B & H		
	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Differenz	Differenz der relativen Häufigkeit (%)	Relative Differenz (%)
Kodierdurchgang 1: Joint Attention Following							
Joint Attention Following	2555	93.7	2637	97.4	82	3.7	3.1
Switching Joint Attention	149	5.5	64	2.4	85	3.1	57.1
Keine Joint Attention	24	0.9	6	0.2	18	0.7	75.0
Kodierdurchgang 2: Switching Joint Attention							
swJA	1160	28.3	831	20.1	329	8.0	28.4
Kodierdurchgang 2: Keine Joint Attention							
kJA	57	1.4	29	0.7	28	0.7	49.1

Kodierdurchgang 1: Joint Attention Following

Die Erhebungsweise von H.S. (H), unterscheidet sich nach erfolgter Beobachtungseinschulung beider Beobachterinnen von der Erhebungsart von T.P. (T2_B) zum einen dadurch, dass sie sich häufiger für eine Aufrechterhaltung der *Joint Attention* (H: 97.4%; T2_B: 93.7%) entschied. Der Unterschied wirkt hier mit einer relativen Differenz von 3.1% gering, bei den damit in Verbindung stehenden anderen Kategorien, bei welchen ein geringeres Auftreten insgesamt gegeben ist, wirkt sich dieser Unterscheid stärker aus. Die Kategorie *Switching Joint Attention* zeigt eine relative Differenz zwischen den Erhebungshäufigkeiten der BeobachterInnen von 57.1% (H: 2.4%; T2_B: 5.5%). Besonders extrem ist der Unterschied in der Erhebung der Kategorie *Keine Joint Attention*, welche zusätzlich nur eine niedrige Gesamterhebungsrate aufweist (H: N=6; T2_B: N=24) und bei welcher sich die Beurteilerin T.P. in Relation gesehen um 75% häufiger für eine Protokollierung von *Keine Joint Attention* entschied, als die Beobachterin H.S.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Bei der Erhebung der Kategorie *Switching Joint Attention* ist nach beidseitig erfolgtem Beobachtungstraining ein relativer Unterschied von 28.4% in der Häufigkeit der Kodierung gegeben (H: 20.1%; T2_B: 28.3%). Der Unterschied zwischen den Häufigkeiten der Erhebung der Kategorie *Keine Joint Attention* nach der Beobachtungseinschulung von H.S. und T.P. beträgt 49.1% (H:0.7%; T2_B:1.4%).

Tabelle 7 gibt einen Überblick über die Interrater-Reliabilität zwischen den zwei BeobachterInnen T.P. und H.S. nach erfolgtem Beobachtungstraining durch Fr. Mag. Hirschmann. Zusätzlich wurde für die Erhebung der Daten auch das Manual verwendet.

Tabelle 7: Interrater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T2_B & H (beide mit Manual und zusätzlichem Beobachtungstraining kodiert)

Video	K1_JA		K2_swJA		K2-kJA	
	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %
NR11	.121	94.7	.533	82.4	-.005	96.3
NR13	.492	98.0	.462	83.8	-	100
NR15	.088	87.6	.63	87.0	-	88.4
NR17	-.014	97.2	.481	84.0	-	99.6
NR19	-	96.9	.501	84.4	-	98.8
PH1	.342	95.2	.657	89.3	-	99.6
PH3	-.023	94.7	.395	80.8	.161	90.4
PH5	-	97.0	.436	89.5	-	100
PH7	-	97.8	.301	82.2	-.004	99.3
PH9	.333	91.3	.567	82.4	-.006	98.8
PH11	.269	91.0	.491	75.7	-.005	97.8
PH13	.275	87.6	.443	75.2	-	98.0
PH15	.456	90.4	.621	82.0	.241	95.8
PH17	.183	95.1	.482	81.0	.535	98.0
PH19	.392	93.2	.58	85.5	.149	96.5
Insgesamt	.279	93.9	.533	82.7	.129	97.3

Kodierdurchgang 1: Joint Attention Following

Wie in der Tabelle 7 ersichtlich, ist die BeobachterInnenübereinstimmung der Erhebung des ersten Kodierdurchgangs dem Kappa-Koeffizienten zufolge sehr gering ($\kappa=.279$), die Proz.Ü 93.9% deutet auf eine gute Übereinstimmung hin. Die Ergebnisse des Cohens Kappa der einzelnen Videos sind insgesamt niedrig, nur bei zwei Erfassungen der Videos (NR13 & PH15) konnten mittelmäßige Übereinstimmungen zwischen den BeobachterInnen mit $\kappa=.492$ & $\kappa=.456$ ermittelt werden. Die Prozentuale Übereinstimmung erreicht durchgehend hohe Werte zwischen 87.6% und 98%.

Negative Cohens Kappa Werte deuten auf eine unzureichende Übereinstimmung zwischen den BeobachterInnen hin. Dies ist bei zwei Videos NR17 und PH3 der Fall. Wie

in Tabelle 8 erkenntlich, konnten in Bezug auf einige Videos keine Kappa-Werte berechnet werden (Video NR19, PH5, PH7). Der Grund dafür liegt in der durchgehend alleinigen Kodierung der Unterkategorie *Joint Attention Following* seitens einer Beobachterin, während die zweite Beobachterin auch Events der anderen Unterkategorien *Switching Joint Attention* und *Keine Joint Attention* im ersten Kodierdurchgang protokollierte (Details siehe in den Kontingenztabellen im Anhang XI).

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Fehlende Kappa-Werte sowie negative Kappa-Koeffizienten sind, wie in Tabelle 7 ersichtlich, häufig in der Kategorie *Keine Joint Attention* vertreten. Für mögliche Gründe dafür, wird auf die Erklärung unter Tabelle 2 verwiesen. Die berechenbaren Kappa-Koeffizienten dieser Kategorie zeigen geringe bis mittelmäßige Übereinstimmungen ($\kappa=.149$ bis $\kappa=.535$). In der Kategorie *Switching Joint Attention* reichen die Kappa-Werte von $\kappa=.301$ bis $\kappa=.657$. Dreizehn von Fünfzehn Werten liegen über der Grenze von $\kappa=.4$ und können als mittelmäßige BeobachterInnenübereinstimmung interpretiert werden. Die durchschnittliche Übereinstimmung der Kategorie *Switching Joint Attention* ist mittelmäßig ($\kappa=.533$), jene der Kategorie *Keine Joint Attention* dagegen mangelhaft ($\kappa=.129$).

Auffällig ist in Tabelle 8, dass die Prozentuale Übereinstimmung bei K2_kJA sehr hohe Werte zwischen 88.4% und 100% erreichte. Die Prozentuale Übereinstimmung ist somit in diesen Kategorien bedeutend höher als in K2_swJA, welche weitaus bessere Kappa-Werte erzielte.

7.1.2.2 Übereinstimmung: Manual und Beobachtungseinschulung

Die Ergebnisse dieses Kapitels überschneiden sich mit der Berechnung der BeobachterInnenübereinstimmung der Masterarbeit von H.Schulz (in Vorbereitung), da die Erhebungsdaten von denselben Videos zur Kalkulation der Interrater-Reliabilität herangezogen wurden.

Tabelle 8 zeigt die Häufigkeitsverteilungen der Beobachtungsdimension *Joint Attention*, sowie die berechnete Differenz zwischen beiden Beobachterinnen im Vergleich der Ergebnisse bei Kodierung mit Manual seitens der Beobachterin T.P. (T1_B) und bei Kodierung mit zusätzlichem Beobachtungstraining der zweiten Beurteilerin (H).

Tabelle 8: Häufigkeiten und Anteilswerte der Beobachtungskategorien: Beurteilung mit Manual (T1_B) im Vergleich zu Beurteilung mit Beobachtungstraining (H)

T1_B &	T1_B		H		T1_B & H		
	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Häufigkeit	Relative Häufigkeit (%)	Absolute Differenz	Differenz der relativen Häufigkeit (%)	Relative Differenz (%)
Kodierdurchgang 1: Joint Attention Following							
Joint Attention Following	2471	90.6	2637	97.4	166	6.8	6.3
Switching Joint Attention	237	8.7	64	2.4	173	6.3	73.0
Keine Joint Attention	19	0.7	6	0.2	13	0.5	68.4
Kodierdurchgang 2: Switching Joint Attention							
swJA	1199	29.2	831	20.1	368	9.1	30.7
Kodierdurchgang 2: Keine Joint Attention							
kJA	101	2.5	29	0.7	72	1.8	71.3

Kodierdurchgang 1: Joint Attention Following

Bei der Datenerhebung mit Manual entschied sich die Beobachterin T.P. im Gegensatz zur Beobachterin H.S., welche die Erhebung nach dem Training durchführte, in Relation gesehen um 73% häufiger für eine Kodierung der Kategorie *Switching Joint Attention* und um 68.4% häufiger für das Erheben eines *Keine Joint Attention*-Events. Damit in Verbindung steht eine im Vergleich zur Beobachterin H.S. geringere Erhebung der Kategorie *Joint Attention Following* (T2_B: 90.6%; H: 97.4%).

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Die Beurteilerin T.P. kodierte nach den Vorgaben des Manuals um 30.7% häufiger ein Event der Kategorie *Switching Joint Attention*, als die Beobachterin H.S. mit Informationen des Beobachtungstrainings. Ein sehr großer Unterschied ist in der Erfassung der Kategorie *Keine Joint Attention* wahrzunehmen, in welcher T.P. um 71.3% öfter die Erhebung eines Events vornahm.

Die Ergebnisse der Tabelle 9 beruhen auf der Übereinstimmung der Beobachterinnen H.S. und T.P. Die Beobachterin H.S. hatte zum Zeitpunkt der Datenerhebung bereits eine Einschulung durch Fr. Mag. Hirschmann, im Gegensatz dazu kodierte T.P. anhand der Vorgaben des Manuals. Gründe für fehlende Kappa-Werte in der Tabelle, auf welche bereits genauer in Kapitel 7.1.1.1 eingegangen wurden, sind auch hier zutreffend und werden nicht mehr näher erläutert.

Tabelle 9: Interrater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T1_B & H (T1_B mit Manual kodiert; H mit zusätzlichem Beobachtungstraining)

	K1_JA		K2_swJA		K2-kJA	
	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %	Kappa	Proz. Ü. %
NR11	.121	94.7	.533	82.4	-.005	96.3
NR13	.492	98.0	.462	83.8	-	100
NR15	.088	87.6	.63	87.0	-	88.4
NR17	-.014	97.2	.481	84.0	-	99.6
NR19	-	96.9	.501	84.4	-	98.8
PH1	.342	95.2	.657	89.3	-	99.6
PH3	-.023	94.7	.395	80.8	.161	90.4
PH5	-	97.0	.436	89.5	-	100.0
PH7	-	97.8	.301	82.2	-.004	99.3
PH9	.333	91.3	.567	82.4	-.006	98.8
PH11	.269	91.0	.491	75.7	-.005	97.8
PH13	.275	87.6	.443	75.2	-	98.0
PH15	.456	90.4	.621	82.0	.241	95.8
PH17	.183	95.1	.482	81.0	.535	98.0
PH19	.392	93.2	.58	85.5	.149	96.5
Insgesamt	.238	91.4	.516	81.8	.178	98.3

Kodierdurchgang 1: Joint Attention Following

Das Ergebnis $\kappa=.238$ gibt eine geringe Übereinstimmung der BeobachterInnen im Vergleich zwischen Erhebung mit und ohne Training an. Generell deutet Cohens Kappa zu den einzelnen Videos auf geringe Übereinstimmungen zwischen den Erhebungen der Beobachterinnen mit Werten $\kappa<.4$ hin. Lediglich bei zwei Videos konnte eine mittelmäßige Übereinstimmung erzielt werden (NR13: $\kappa=.492$ & PH15: $\kappa=.456$). Die Prozentuale Übereinstimmung ist, wie bereits in den Vergleichen zuvor, wieder mit hohen Werten $<87.6\%$ vertreten und deutet auf eine sehr gute Übereinstimmung zwischen den Erhebungen hin.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Die Übereinstimmung zwischen den Beobachterinnen bei der Erhebung der Kategorie *Switching Joint Attention* ist mit Kappa Werten von $\kappa=.301$ (Video PH7) bis $\kappa=.657$ (PH1) gering bis mittelmäßig. Der durchschnittliche Kappa Koeffizient ($\kappa=.516$), gerechnet über alle Videos dieser Gruppe, deutet auf eine mittelmäßige Übereinstimmung zwischen den Beurteilenden hin. Die Kategorie *Keine Joint Attention* zeigt, mit der Ausnahme des Videos PH17 ($\kappa=.535$), geringe oder keine Übereinstimmungswerte der Kodierungen der einzelnen Videos, insgesamt steht der Wert $\kappa=.178$ für eine mangelhafte Übereinstimmung beim Vergleich der Kodierung dieser Kategorie mit und ohne Training. Die Prozentuale Übereinstimmung ist in der Kategorie *Keine Joint Attention* mit beinahe allen Werten $>95\%$ wiederum sehr hoch, in der Kategorie *Switching Joint Attention* erlangt die Prozentuale Übereinstimmungen Werte zwischen 75% und 87%.

In Tabelle 10 wird die BeobachterInnenübereinstimmung der Beobachtungskategorie *Joint Attention* zusammengefasst dargestellt.

Tabelle 10: Interrater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung der Beobachtungsdimension *Joint Attention* insgesamt & der Kategorien JA, swJA, kJA

	Kappa	Prozentuale Übereinstimmung (%)
Insgesamt (K1 & K2)		
T1_B & H	.782	89.9
T2_B & H	.803	90.9
K1_JA		
T1_B & H	.238	91.4
T2_B & H	.279	93.9
K2_swJA		
T1_B & H	.516	81.8
T2_B & H	.533	82.7
K2_kJA		
T1_B & H	.178	98.3
T2_B & H	.129	97.3

Die Interrater-Reliabilität insgesamt ist bei Erhebung der Daten mit Beobachtungstraining beider BeobachterInnen (T2_B & H) mit $\kappa=.803$ insgesamt sehr hoch und deutet, auf eine sehr gute BeobachterInnenübereinstimmung hin.

Bei der Übereinstimmung zwischen den Erhebungsdaten, welche von einer Beobachterin nur anhand der Regeln des Manuals erfasst wurden, von der anderen Beobachterin jedoch mit zusätzlicher Einschulung (T1_B & H), konnte ebenfalls eine sehr gute

Übereinstimmung zwischen den BeobachterInnen $\kappa=.782$ festgestellt werden. Die Prozentuale Übereinstimmung (T1_B & H: 89.9%; T2_B & H:90.9%) weist bei beiden Vergleichen der Übereinstimmung ebenfalls auf eine hohe Kongruenz zwischen den beurteilenden Personen hin.

Die Erhebung des *Joint Attention Following* im ersten Kodierdurchgang sowie die Kategorie *Switching Joint Attention* des zweiten Kodierdurchgangs führten nach Absolvierung des Beobachtungstrainings zu einer geringfügig höheren BeobachterInnenübereinstimmung (T2_B & H), als die Übereinstimmung zwischen dem Vergleich Erhebung mit Manual und Erhebung mit Training (T1_B & H).

Die BeobachterInnenübereinstimmung der Kategorie *Keine Joint Attention* (K2) nahm nach dem Beobachtungstraining der Beobachterin T.P. im Vergleich zur ebenfalls nach dem Beobachtungstraining durchgeführten Erhebung von H.S. geringfügig ab. Die Unterschiede der Kappa-Koeffizienten und der Prozentualen Übereinstimmung führen in keiner Kategorie der Beobachtungsdimension zu einer Veränderung in der Interpretation der Übereinstimmung der BeobachterInnen.

7.1.2.3 Die zusammengefasste Prozentuale Übereinstimmung

Die *Zusammengefasste Prozentuale Übereinstimmung* dient zum Vergleich von Raten der Kodierungen der Beobachterinnen und ist ein Test auf Homogenität der Schwellenwerte. Das Übereinstimmungsmaß nimmt Bezug auf den *Effekt unterschiedlicher Schwellenwerte*, welcher in Kap. 2.1.4. näher erläutert wurde. Es wird untersucht ob beurteilende Personen das Verhalten gleich häufig erfassen und gibt das Verhältnis der Häufigkeit an, mit denen die Beurteilenden ein Merkmal oder eine Kategorie insgesamt registriert haben. Je geringer der Wert, das bedeutet, je weiter entfernt von 100%, desto mehr ist die Übereinstimmung durch den *Effekt der unterschiedlichen Grundwahrscheinlichkeiten* beeinflusst und desto höher die Wahrscheinlichkeit, dass die BeobachterInnen trotz derselben Ausgangsinformation, bei der Kodierung ihr Urteil bei unterschiedlicher Intensität oder unterschiedlicher Dauer des Merkmals abgeben. In Kontingenztabellen (siehe Kap. 2.1.2) wird diese Art des Unterschieds im Kodierverhalten sichtbar, wenn eine der Zellen in welcher die Nicht-Übereinstimmungen abgebildet werden, unbesetzt ist (Wirtz & Caspar, 2002).

Die *Zusammengefasste Prozentuale Übereinstimmung* wurde unter Verwendung folgender Formel berechnet:

$$\text{Proz. Ü(zusammen)} = \frac{\text{kleinere Anzahl von Beobachtungen}}{\text{größere Anzahl von Beobachtungen}} * 100\%$$

Konkret wurde die Anzahl der Erhebungen eines Merkmals der Beobachterin, welche das Merkmal weniger häufig kodierte, durch die Anzahl der Erhebungen der Beobachterin dividiert, welche das Merkmal häufiger erfasste. Anschließend wurde mit 100 multipliziert um eine prozentuale Aussage treffen zu können.

Tabelle 11 zeigt die Ergebnisse der *Zusammengefassten Prozentualen Übereinstimmung*, welche auf einen *Effekt unterschiedlicher Grundwahrscheinlichkeiten* zwischen der Erhebung von H.S. und T.P. nach beidseitiger Absolvierung der Durchführung des BeobachterInnentrainings prüft.

Tabelle 11: Zusammengefasste Prozentuale Übereinstimmung: T2_B & H (beide mit Manual und zusätzlichem Beobachtungstraining kodiert)

T2_B & H	Kappa	Proz. Ü. (%)	Zusammengefasste Proz. Ü. (%)
Kodierdurchgang 1_Joint Attention			
	.279	93.9	JAf → 96.3 swJA → 44.4 kJA → 25.0
Kodierdurchgang 2_Switching Joint Attention			
	.533	82.7	swJA → 71.5
Kodierdurchgang 2_Keine Joint Attention			
	.129	97.3	kJA → 28.7

Kodierdurchgang 1: Joint Attention Following

Die Erhebung der Kategorie *Joint Attention Following* im ersten Kodierdurchgang deutet mit einem Prozentwert von 96.3% auf eine ähnliche Rate der Kodierungen der BeurteilerInnen hin. Bei der Kategorie *Switching Joint Attention* ist mit 44.4% ein eher niedriger Prozentwert gegeben, was auf einen *Effekt unterschiedlicher Grundwahrscheinlichkeiten* der Beurteilenden schließen lässt. Noch deutlicher deutet die Kategorie *Keine Joint Attention*, welche mit 25% einen sehr niedrigen Wert erzielt, auf die Anwendung unterschiedlicher Schwellen der Beobachterinnen hin.

Kodierdurchgang 2: Switching Joint Attention & Keine Joint Attention

Die Ergebnisse der Kategorie *Keine Joint Attention* des zweiten Durchgangs, mit einem Prozentwert von 28.7%, deuten ebenfalls auf eine beim Auftreten eines Merkmals unterschiedliche Anwendung von Schwellen der BeobachterInnen hin. Bei der Kategorie *Switching Joint Attention* weist der Prozentwert von 71.5% daraufhin, dass sich die Grundwahrscheinlichkeiten der Beobachterinnen insgesamt weniger drastisch unterscheiden.

8 DISKUSSION

Die vorliegende Arbeit hatte zum Ziel einen ersten Einblick über die Zuverlässigkeit des Manuals und die Rolle des BeobachterInnentrainings bei der Erfassung der Beobachtungskategorie *Joint Attention* des Verhaltensinventars INTAKT zu geben.

Im Zuge der Untersuchung wurde, anhand bereits vorhandenem Videomaterial von 29 Mutter-Kind-Interaktionen, die Beobachtungskategorie *Joint Attention* händisch mittels *Time-Sampling-Verfahren* erhoben. Die Datenerfassung erfolgte in zwei Kodierdurchgängen: Der erste Kodierdurchgang konzentrierte sich unter der Verwendung des *Predominant-Activity-Samplings* auf die Erhebung der *Joint Attention*. Konkret wurde im Rahmen von 15-Sekunden-Intervallen beim überwiegendem Auftreten einer der Unterkategorien *Joint Attention Following*, *Switching Joint Attention*, *Keine Joint Attention* oder *Unkodierbar* zwischen einer der vier Kategorien gewählt.

Im zweiten Durchgang erfolgte die getrennte Erhebung der Kategorien *Switching Joint Attention* und *Keine Joint Attention* in 10-Sekunden-Intervallen auf Basis der *One-Zero-Sampling* Methode. Diese Erhebungsmethode erlaubte auch die gleichzeitige Erfassung beider Kategorien innerhalb eines Intervalls.

Für die Berechnung der *Intra-Rater-Reliabilität* wurde jedes Video doppelt kodiert. In der ersten Erhebungsphase wurden alle Videos von der Beobachterin T.P. mit Manual kodiert, in der zweiten Erhebungsphase erfolgte nach 14 Videos eine Beobachtungseinschulung, danach wurden die letzten 15 Videos wiederholt erhoben. Anschließend wurden die Erhebungsdaten miteinander verglichen.

Für die Berechnung der *Interrater-Reliabilität* wurde die Datenerhebung der Videos nach Absolvierung eines Beobachtungstrainings von zwei voneinander unabhängigen Beobachterinnen T.P. und H.S., auf den Grad der Übereinstimmung geprüft. Um Aussagen über den Unterschied zwischen der Kodierung mit Manual und der Kodierung mit Beobachtungseinschulung machen zu können, wurden die Ergebnisse der Beobachterin H.S. auch mit der Erhebung der Videos, welche ohne Training von T.P. erfasst wurden, verglichen.

Erst eine hohe Übereinstimmung zwischen den Ergebnissen der BeobachterInnen dient als guter Indikator für die genaue und stabile Erhebung eines Merkmals im Zuge einer Verhaltensbeobachtung. Eine hohe BeobachterInnenübereinstimmung ist ein

notweniges Kriterium für eine hohe Vertrauenswürdigkeit eines Verfahrens und den eingesetzten BeobachterInnen. Dabei sollen nach Döring et al. (2016) alle Kategorien einzeln berücksichtigt werden, um Fehlerquellen, wie ungenaue Abgrenzungen zwischen den Unterkategorien, zu identifizieren.

In Bezug auf das Verhaltensbeobachtungsinventar INTAKT konnten insgesamt sehr gute Kappa Werte in der Berechnung der *Intra-Rater-Reliabilität* (Vergleich Manual: $\kappa=.896$; Vergleich Manual und Training: $\kappa=.902$) und der *Interrater-Reliabilität* (Vergleich Manual und Training: $\kappa=.782$; Vergleich Training: $\kappa=.803$) festgestellt werden und deuten auf eine sehr gute BeobachterInnenübereinstimmung hin. Die Ergebnisse der *Prozentualen Übereinstimmung der Intra-Rater-Reliabilität* (Vergleich Manual: 95.2%; Vergleich Manual und Training: 95.3%) und der *Interrater-Reliabilität* (Vergleich Manual und Training: 89.9%; Vergleich Training: 90.9%) weisen ebenfalls auf eine hohe Übereinstimmung der Beobachtenden hin.

Bei genauerer Analyse wurden bei der Berechnung der *Interrater-Reliabilität* in einzelnen Kategorien der Beobachtungsdimension, trotz hoher Prozentualer Übereinstimmung, unzureichende Kappa-Koeffizienten berechnet. Mögliche Gründe dafür werden anschließend diskutiert:

Limitationen Cohens Kappa

Wie bereits in Kapitel 2.1.2. ausführlich dargestellt, kann das statistische Maß Cohens Kappa wesentlich durch Bedingungen, welche nicht auf das Kodierverhalten der BeurteilerInnen zurückzuführen ist, beeinflusst werden (Wirtz & Caspar, 2002). Konkret bedeutet das, dass obwohl Cohens Kappa theoretisch einen Wert zwischen -1 und +1 annehmen kann, $\kappa=1$ nur erreicht werden kann, wenn die Übereinstimmung zwischen den BeobachterInnen maximal ist und keine Verzerrungen aufgrund Beobachtungsfehlern auftreten (Bakeman & Quera, 2011).

Zusätzlich verhindern ungleiche Randverteilungen in der Kontingenztabelle das Erreichen des maximalen Kappa-Wertes 1 und führen zusätzlich zu einer Unterschätzung der Übereinstimmung (Wirtz & Caspar, 2002; Wirtz 2006). Ungleiche Randverteilungen können einerseits auftreten, wenn die BeobachterInnen sich in ihrem Kodierverhalten stark unterscheiden oder die Prävalenz des untersuchten Merkmals sehr niedrig oder hoch ist. Dies führt zu einer geringen Zellbesetzung in einzelnen Zellen der Kontingenztabelle, die zu einer Unterschätzung der Konkordanz beitragen können. Bei einer Auswahl von zwei Kodiermöglichkeiten und ungleichmäßiger Prävalenz des

Merkmals würde eine achtzigprozentige Wahrscheinlichkeit der Übereinstimmungen zwischen den BeobachterInnenübereinstimmung bereits mit einem Wert von $\kappa=.2$ gegeben sein (Bakeman & Quera, 2011, S. 66). Bakeman & Quera (2011) betonen die Wichtigkeit des Miteinbeziehens der Anzahl an Beurteilungskategorien in der Interpretation von Kappa, denn je geringer die Anzahl der Beurteilungskategorien, umso niedriger sind die zu erwartenden Kappa Werte.

In Hinblick auf die Ergebnisse der vorliegenden Arbeit sind die eben genannten Faktoren, welche Einfluss auf die Berechnung des Cohens Kappa nehmen können, nicht unbedeutend. Vor allem in der Interpretation der Übereinstimmungsergebnisse der Kategorie *Keine Joint Attention* (K2_kJA), in welcher trotz abgeschlossenem Training beider BeobachterInnen eine mangelhafte Übereinstimmung gegeben ist ($\kappa=.129$), sollten mögliche Einflüsse aufgrund dieser Faktoren in die Interpretation einfließen. Die Erhebung der Kategorie *Keine Joint Attention* bietet aufgrund des *One-Zero-Samplings* nur zwei Kodiermöglichkeiten (tritt auf/tritt nicht auf) und darüber hinaus zeigt das Merkmal *Keine Joint Attention* eine sehr niedrige Prävalenz. Bei Erhebung der Kategorie nach Abschluss des Beobachtungstrainings kodierte die Beobachterin T.P. in 1.4% der Fälle ein Auftreten von *Keine Joint Attention*, die Beobachterin H.S. erfasste im zweiten Kodierdurchgang bloß zu 0.7% das Event *Keine Joint Attention* (siehe Tabelle 6). Die Interpretation der Ergebnisse anhand der gängigen Grenzwerte (Kap.5.1.3., Döring et al., 2016) würden somit nach Bakeman und Quera (2011) zu einer weitgehenden Unterschätzung der Übereinstimmung führen. Auch die *Prozentuale Übereinstimmung* mit einem Wert von 97.3% in dieser Kategorie deutet daraufhin, dass das Cohens Kappa durch den Einfluss der seltenen Auftrittshäufigkeit des Verhaltens, sowie die geringe Anzahl der Codes, die Übereinstimmung der BeobachterInnen unterschätzt. Die bedeutend höhere BeobachterInnenübereinstimmung $\kappa=.533$ der Kategorie *Switching Joint Attention* (K2_swJA) nach stattgefundenem Beobachtungstraining beider Beurteilenden, könnte neben einer möglicherweise einfacheren Erfassung des Merkmals, auch mit einer generell höheren Prävalenz dieses Verhaltens zusammenhängen. Die Kodierung eines Events wurde von T.P. zu 28.3% vorgenommen, von der zweiten Beobachterin H.S. zu 20.1%

Zusammenfassend deuten die hohen Ergebnisse der Prozentualen Übereinstimmung mit Werten $>80\%$ in allen Berechnungen der *Interrater-Reliabilität* (K1_JA: 91.4%; K2_swJA: 81.8%; K2_kJA: 98.3%) ebenfalls auf eine Unterschätzung der Übereinstimmung der BeobachterInnen durch das Cohens Kappa hin. Die Diskrepanz zeigt sich

wieder besonders deutlich in der Kategorie *Keine Joint Attention* (Kodierdurchgang 2), da der Kappa-Wert $\kappa=.129$, nach Training beider BeobachterInnen, auf eine mangelnde Übereinstimmung, die *Prozentuale Übereinstimmung* mit 97.3% gleichzeitig auf eine sehr gute Übereinstimmung in der Kodierung hinweist (siehe Tab.8). Abschließend kann angemerkt werden, dass viele Anzeichen dafürsprechen, dass in der Kategorie *Keine Joint Attention* das Cohens Kappa zusätzlich durch eine Vielzahl beobachterinnenunabhängiger Faktoren beeinflusst wurde und eine Interpretation angelehnt an die gängigen Cutoffs (siehe Kap. 5.1.3; Döring et al., 2016), in diesem Fall wahrscheinlich zu einer Unterschätzung der Übereinstimmung zwischen den BeobachterInnen führt. Die in diesem Absatz genannten Faktoren, welche die Aussagekraft des Cohens Kappa limitieren, können ebenfalls zu einer Unterschätzung der BeobachterInnenübereinstimmung in der Kategorie *Keine Joint Attention* des ersten Kodierdurchgangs geführt haben, weitere statistische Analysen könnten helfen, nähere Information dazu zu generieren.

Intra-Rater-Reliabilität

Die *Intra-Rater-Reliabilität* der nach den Regeln des Manuals erhobenen Ergebnisse (T1_A & T2_A) deuten mit $\kappa=.896$ und $\text{Proz.Ü} = 95.2\%$ auf eine sehr gute Übereinstimmung der Beobachterin T.P. zwischen erster und zweiter Kodierung der Videos hin. Überraschend ist die Tatsache, dass die Übereinstimmung der Daten zu den Aufnahmen, welche in der ersten Phase nur mit Manual kodiert wurden, bei wiederholter Kodierung nach erfolgtem Training (T1_B & T2_B) mit $\kappa=.902$ und $\text{Proz.Ü} = 96.3\%$ noch höher ausfiel (Siehe Tab. 6). Dieses Ergebnis spricht gegen die Erwartung, dass das Beobachtungstraining zu einer Veränderung des Wissensstandes über das Beobachtungsverfahren und damit einhergehend zu einer Veränderung in der Kodierung führen würde. Diesem Ansatz zufolge sollte die Übereinstimmung der Ergebnisse bei einem Vergleich mit demselben Wissensstand (alle Videos wurden anhand der Vorgaben des Manuals kodiert) höher ausfallen, als bei einem Vergleich mit unterschiedlichem Wissensstand (ein Teil der Videos wurde mithilfe des Regelwerkes kodiert, der andere Teil mit „vertieftem“ Wissen des zusätzlichen Trainings). Anzumerken ist, dass der Unterschied zwischen der Übereinstimmung bei den Erhebungen mit Manual und der Übereinstimmung bei der Erhebung mit Manual im Vergleich zu der Erhebung mit Beobachtungstraining mit einer Differenz von $\kappa=.006$ nur schwindend gering ist. Das Ergebnis

deutet dennoch darauf hin, dass die Einschulung zu keiner wesentlichen Veränderung in der Erfassung der Beobachtungskategorie *Joint Attention* geführt hat.

Eine mögliche Erklärung für die ausbleibende Veränderung durch das BeobachterInnentraining wäre, wenn die Einschulung für die Beurteilerin als Bestätigung ihres Tuns empfunden wurde und deshalb kein Einfluss auf ihr Kodierverhalten erfolgte. Dies kann positiv ausgelegt werden, da dieser Ansicht nach das Manual selbst gut ausgearbeitet ist und eine Beobachtungseinschulung nicht zwingend notwendig für eine reliable Erhebung der Daten scheint. Eine weitere Erklärung könnte jedoch sein, dass die Beobachterin nach der Kodierung von 43 Videos anhand der Regeln des Manuals bereits einen „festgefahrenen Stil der Kodierung“ ausgebildet hatte und sich deshalb nicht mehr offen zeigte, neue Aspekte welche im Rahmen der Schulung auftauchten, wahrzunehmen und umzusetzen.

Ergänzend soll erwähnt werden, dass die Intra-Rater-Reliabilität Aussagen über die Stabilität und Konsistenz der Kodierungsweise einer beurteilenden Person macht. In Anbetracht dieser Tatsache sind Erinnerungseffekte ein bedeutender Faktor, welcher zu einer Überschätzung der Beobachtungsübereinstimmung beitragen kann, vor allem wenn die Kodierung innerhalb eines kurzen Zeitrahmes durchgeführt wurde (Wirtz & Caspar, 2002). Die Verhaltensbeobachtung ist stets mit der selektiven Wahrnehmung der Beobachtenden verbunden, die durch individuelle Sichtweisen, subjektiven Interessen und Erwartungen beeinflusst ist (Döring et al., 2016). Unter diesem Aspekt ist es nicht überraschend, dass die Intra-Rater-Reliabilität sehr hoch ausfiel, da dies vor allem bedeuten könnte, dass die beobachtende Person bei wiederholtem Kodieren der Videos dieselben Merkmale verstärkt wahrnahm.

Ein letzter Einflussfaktor, im Zuge der höheren Übereinstimmung zwischen der Erhebung mit und ohne Training, könnte der Zeitraum der Erhebung dargestellt haben. Der Zeitraum zwischen der Kodierung der Videos mit Manual T1_A & T2_A war länger (ca. zwei Monate), als der Zeitraum zwischen der Erhebung der Videogruppen T1_B & T2_B, bei welchen ein Vergleich zwischen Manual und Training gemacht wurde (ca. ein Monat). Für weitere Untersuchungen sollte eine Standardisierung der zeitlichen Variable vorgenommen werden, um Verzerrungen der Ergebnisse aufgrund Erinnerungseffekte auszuschließen.

Interrater-Reliabilität:

Die Interrater-Reliabilität der Gruppen T2_B und H, in denen die Videos mit zusätzlicher Information des Trainings kodiert wurden, sowie der Gruppen T1_B und H, in welcher nur die Videos der Gruppe H mit Einschulung, diejenigen der Gruppe T1_B nur mit Manual erhoben wurden, deuten beide auf eine sehr gute Beobachtungsübereinstimmung mit Kappa Werten über dem Cutoff $\kappa=.75$ hin. Die Übereinstimmung zwischen den BeobachterInnen nach dem Training ($\kappa=.803$; Proz.Ü.=90.9) ist höher als die Übereinstimmung zwischen dem Vergleich der Erhebung mit Manual und der Erhebung mit Einschulung ($\kappa=.782$; Proz.Ü.=89.9%), der Unterschied ist jedoch sehr schwach. Dies beruht unter anderem auf der Tatsache, dass die Beobachterin T.P. keine wesentliche Veränderung in ihrem Kodierverhalten nach dem Training, im Vergleich zu der Erhebung davor, zeigte. Bei Betrachtung der Ergebnisse der einzelnen Kategorien beider Kodierdurchgänge verbesserte sich die Übereinstimmung mit Training beider BeobachterInnen in den Kategorien *Joint Attention Following* (K1_JA) und *Switching Joint Attention* (K2_swJA) geringfügig. Die Konkordanz der Erhebung von *Joint Attention Following* stieg von $\kappa=.238$ (Vergleich einer Beobachterin mit und einer ohne Training) auf $\kappa=.279$ nach beidseitig erfolgtem INTAKT-Beobachtungstraining. Die Kategorie *Switching Joint Attention* zeigt eine ähnliche Entwicklung hin zu einer minimalen Verbesserung der Übereinstimmung der BeobachterInnen nach der Beobachtungseinschulung (T1_B & H: $\kappa=.516$; proz.Ü.=81.8% & T2_B & H: $\kappa=.533$; proz.Ü.=82.7%). Nach dem Training der Beobachterin T.P. kam es zu einer Verringerung der Konkordanz zwischen den Erhebungen der BeobachterInnen in der Kategorie *Keine Joint Attention* (T1_B & H: $\kappa=.178$; Proz.Ü.=98.3%; T2_B & H: $\kappa=.129$; Proz.Ü.=97.3%). Die Ergebnisse können Tabelle 10 entnommen werden. Interessant ist dabei, dass die Häufigkeitsverteilung (siehe Tab.6) zeigt, dass die Anzahl der Kodierung eines Events *Keine Joint Attention* durch die Beobachterin T.P. nach dem Beobachtungstraining abnahm und eigentlich ein Angleichen der Erhebungshäufigkeit zwischen den BeobachterInnen stattfand (T1_B: N=101; T2_B: N=57; H: N=29). Der geringere Kappa-Wert ist jedoch ein Hinweis, dass sich die BeobachterInnen uneinig über das Event *Keine Joint Attention* waren. Dies führt zu der Vermutung, dass die BeobachterInnen unterschiedliche Auffassungen über das Grundverständnis der Kategorie *Keine Joint Attention* hatten und aufgrund mangelnder Konsistenz Schwierigkeiten in einer einheitlichen Erhebung des Merkmals hatten. Es wird empfohlen, dieser Kategorie im Rahmen weiterer Übereinstimmungsverfahren mehr Aufmerksamkeit zu

widmen und bei Bedarf im Manual und Beobachtungstraining spezifischer darauf einzugehen.

Bei näherer Betrachtung der Ergebnisse der *Interrater-Reliabilität* fällt außerdem auf, dass es in einigen Bereichen eine hohe Diskrepanz zwischen den Cohens Kappa Werten der einzelnen Kategorien *Joint Attention Following*, *Joint Attention Switching* und *Keine Joint Attention* im Vergleich zur Übereinstimmung der Beobachtungskategorie *Joint Attention* insgesamt gibt (siehe Tabelle 11).

Der Einfluss von ungleichen Randverteilungen in der Kontingenztafel oder die Prävalenz des interessierenden Merkmals auf das Cohens Kappa wurden in diesem Kapitel sowie in Kap. 2.1.2. näher erläutert. Ungleiche Randverteilungen können entstehen, wenn eine der beurteilenden Personen prinzipiell häufiger oder geringer eine Kodierung vornimmt. Dieser Unterschied wirkt sich direkt auf den Koeffizienten zur Berechnung der Beobachtungsübereinstimmung aus. Heterogene Randverteilungen führen Wirtz und Caspar (2002) zufolge immer zu einer Verringerung des Cohens Kappa Wertes, da der Variationsbereich des Koeffizienten kleiner wird, je mehr die Randsummen der Kontingenztabellen voneinander abweichen. Für eine Verbesserung der Interpretation der *Interrater-Reliabilität* wurde deshalb die *Zusammengefasste Prozentuale Übereinstimmung* berechnet, welche den *Effekt der unterschiedlichen Grundwahrscheinlichkeiten* untersucht. Diesem Effekt zufolge ist die Kongruenz abhängig vom Kodierverhalten der beurteilenden Person: je konservativer diese kodiert, umso stärker muss ein Merkmal ausgeprägt sein oder umso länger muss ein Merkmal auftreten, damit eine Erhebung erfolgt (Wirtz & Caspar, 2002). Die Überprüfung unterschiedlicher Schwellenwerte ergab, dass die BeobachterInnen, trotz beidseitig erfolgtem BeobachterInnentraining, in der Erhebung der Daten unterschiedliche Schwellen für die Kodierung einzelner Verhaltensmerkmale anwandten. Vor allem die Kategorien *Switching Joint Attention* (44.4%) und *Keine Joint Attention* (25%) des ersten Kodierdurchgangs (PAS), sowie die Kategorie *Keine Joint Attention* (28.7%) des zweiten Kodierdurchgangs (OZS) lassen mit Werten, welche weit von 100% entfernt sind, auf unterschiedliche Schwellenanwendungen in der Kodierung seitens der BeobachterInnen schließen. Ebenso ist der Häufigkeitsverteilung (siehe Tabelle 6) zu entnehmen, dass es in der Erhebung der Daten unterschiedliche Tendenzen zwischen den Beobachtenden gab. Die Beobachterin T.P. protokollierte in allen Kategorien, ausgenommen der *Joint Attention Following* (K1_JA), zwischen 28.4% bis 75% häufiger das Auftreten eines Events, als die Beobachterin H.S. Um einen korrekten Umgang mit dieser

Differenz der Grundwahrscheinlichkeiten zwischen den BeurteilerInnen zu finden, ist es wichtig zu klären, weshalb individuelle Tendenzen in der Erfassung zustande gekommen sind und auf diese einzugehen. Da systematische Verzerrungen wie diese in der Beobachtung die Ergebnisse massiv beeinflussen können, raten Caspar und Wirtz (2002) mögliche Ursachen detailliert mit den BeobachterInnen zu besprechen. Wichtig wäre zu klären, inwieweit der *Effekt der mangelnden Konsistenz* zusätzlich eine Rolle in der Beeinträchtigung der Übereinstimmung einnimmt. Falls die Ursache zwischen den abweichenden Erfassungen der Merkmale der Beobachtenden durch Beurteilungsfehler & -verzerrungen, wie dem Milde-Härte-Effekt, dem Observer-Drift, dem Ankereffekt, oder konservativer oder progressiver Urteilsbildungstendenzen bedingt ist, sollten diesbezüglich Maßnahmen zu Vermeidung weiterer Fehler gesetzt werden (Kap.2.3). Ebenso könnte der Wunsch der BeobachterInnen, den Erwartungen der ForscherInnen gerecht zu werden, Einfluss auf das Kodierverhalten genommen haben (Wirtz & Caspar, 2002).

Die Beantwortung der Ausgangsfragestellung dieser Arbeit, ob für die Beobachtungsdimension *Joint Attention* eine ausreichende Übereinstimmung zwischen der Datenerhebung mit Manual im Vergleich zur Kodierung mit zusätzlichem Beobachtungstraining gegeben ist, lässt sich in Anbetracht der Vielzahl an möglichen beobachtungsunabhängigen Einflüssen nicht eindeutig beantworten. Es kann festgehalten werden, dass es im Zuge des Beobachtungstrainings zu keiner aussagekräftigen Veränderung in der Datenerhebung gekommen ist. Dies kann als positiver Aspekt im Sinne der Ökonomie hinsichtlich der Anwendung des Beobachtungsinventars in der Praxis betrachtet werden, da das Ergebnis darauf hindeutet, dass eine Erhebung mit Manual ausreichend die Beobachtungsdimension *Joint Attention* zu erfassen vermag. Andererseits sind die Ergebnisse der Interrater-Reliabilität, welche für die notwendige Objektivität und Reliabilität des Verfahrens essentiell sind, in einzelnen Kategorien von unzureichender Aussagekraft. Die Ursache für niedrige Übereinstimmungen im Cohens Kappa zwischen den BeobachterInnen kann, wie in diesem Teil der Arbeit bereits ausführlich diskutiert, in beobachterunabhängigen Faktoren begründet sein, das Ausschließen anderer Gründe ist jedoch nicht gegeben und sollte in weiteren Untersuchungen geprüft werden.

Anregungen und Ausblick

Es wird empfohlen eine weitere Untersuchung mit mehr BeobachterInnen durchzuführen, da die Repräsentativität dieser Arbeit, aufgrund der niedrigen Anzahl der BeobachterInnen (N=2) stark eingeschränkt ist, insbesondere da individuelle Beobachtungsfehler der BeobachterInnen angenommen werden können, die zu einer Verzerrung der Ergebnisse beigetragen haben. Für eine klare Beantwortung der Fragestellung, ob ein Beobachtungstraining zu einer Veränderung der Übereinstimmung zwischen den BeurteilerInnen hinsichtlich Zuverlässigkeit und Genauigkeit der Erfassung der Merkmale führt, oder anhand der Datenerhebung mit Manual ausreichend objektive und reliable Ergebnisse erreicht werden, wird ebenfalls empfohlen weitere Untersuchungen unter Einbezug mehrerer BeobachterInnen vorzunehmen. Bei einer Erhebung durch mehrere BeobachterInnen könnte inspiziert werden, ob sich die Differenzen relativieren, andernfalls könnte eine zusätzliche Einführung der BeobachterInnen zum Thema Beobachtungsfehler und -verzerrungen hilfreich sein, um eine höhere Aussagekraft der Reliabilität zu erzielen (Döring et al., 2016; Rentzsch & Schürz, 2009).

Konkrete Maßnahmen könnten zur Verbesserung der BeobachterInnenübereinstimmung der Kategorie *Keine Joint Attention* gesetzt werden, indem im Manual und Beobachtungstraining vertiefend auf diese Kategorie eingegangen wird. Die im INTAKT-Manual enthaltene Definition ist eher kurz gehalten und könnte den Beobachtenden zu viel Raum für subjektive Interpretation des Verhaltensmerkmals bieten, sodass die BeobachterInnen für sich eigens definierte Regeln in der Erhebung des Merkmals anwenden (Fisseni, 1997). Zusätzlich wurde in der Beobachtungseinschulung für die Probekodierung eine Videosequenz gewählt, welche kein Auftreten des Events *Keine Joint Attention* zeigte. Eine kurze Demonstration eines Videoausschnittes im Zuge des BeobachterInnentrainings, in welcher ein Auftreten des Merkmals *Keine Joint Attention* deutlich wird, könnte zu einer Verbesserung der Objektivität und Reliabilität der Kategorie beitragen.

ABSTRACT DEUTSCH

Die Gütekriterien Objektivität und Reliabilität werden bei wissenschaftlichen Verhaltensbeobachtungen unter anderem dadurch beurteilt, ob die beobachtenden Personen zu ähnlichen Urteilen in der Erhebung gelangen. Die BeobachterInnenübereinstimmung gilt als essentielle Voraussetzung hinsichtlich Zuverlässigkeit und Messgenauigkeit bei der Erprobung eines neuen Verfahrens.

In Bezug dazu stellt sich die Frage nach einem effektiven und ökonomischen Training der BeobachterInnen. In vielen Fällen genügt ein Manual mit einem gut ausgearbeiteten Regelwerk, um die beurteilenden Personen mit dem Verfahren und dessen Anwendung vertraut zu machen. Bei der Beobachtung komplexer Verhaltensprozesse kann jedoch ein zusätzliches Beobachtungstraining nötig sein, um die Qualität der Erhebung zu sichern.

Das standardisierte Videobeobachtungsinstrument INTAKT ist ein Verfahren, das die Beziehungs- und Interaktionsqualität anhand der Kategorien *Feinfühligkeit*, *Rückmeldung* und *Joint Attention* zwischen Bezugsperson und Kind erhebt. Anhand der Beobachtungsdimension *Joint Attention* wurde ein Vergleich der Übereinstimmung der Beobachtenden bei Erhebung der Daten mit dem INTAKT-Manual und bei Erhebung der Daten mit zusätzlichem Beobachtungstraining vorgenommen. Zu diesem Zweck wurden 29 bereits aufgezeichnete Videos mit Manual kodiert. Anschließend wurde die Hälfte der Aufzeichnungen wiederholt mit Manual erfasst, die wiederholte Kodierung der zweiten Hälfte erfolgte nach einem Beobachtungstraining. Es folgte die Berechnung der *Interrater-Reliabilität* und der *Intra-Rater-Reliabilität* mittels *Cohens Kappa* (κ) und *Prozentualer Übereinstimmung* (Proz. Ü.).

Das Beobachtungstraining zeigte keine signifikanten Veränderungen in der Erhebung der Beobachtungskategorie *Joint Attention* verglichen mit der Erhebung mit Manual. Die Übereinstimmung der *Intra-Rater-Reliabilität* der Beobachtungsdimension *Joint Attention* zwischen Erhebung mit Manual und Erhebung mit Beobachtungstraining war sehr gut ($\kappa=.902$; Proz. Ü.= 95,3%). Die Übereinstimmung der *Interrater-Reliabilität* der Beobachtungsdimension *Joint Attention* war ebenfalls sehr hoch, sowohl nach absolviertem Beobachtungstraining beider BeobachterInnen ($\kappa=.803$; Proz. Ü.=90.9%), als

auch beim Vergleich zwischen Kodierung mit Manual einer Beobachterin und Kodierung mit zusätzlichem Training der anderen Beobachterin ($\kappa=.782$; Proz. Ü.=89.9%). Die *Interrater-Reliabilität* zeigt jedoch bei beiden Vergleichen geringe Übereinstimmung in den Kategorien *Keine Joint Attention* und *Joint Attention Following*. Limitationen und die Möglichkeit über den Einfluss externer Faktoren (z.B. Prävalenz eines Merkmals) auf das Resultat der BeobachterInnenübereinstimmung werden diskutiert und sollten in weiteren Untersuchungen erneut überprüft werden.

ABSTRACT ENGLISH

Objectivity and Reliability are commonly reviewed in scientific behavioral observation by evaluating, if the observers succeed in making consistent rating decisions. Therefore, observer agreement is an essential requirement for trustworthiness and accuracy in measurement when testing a new instrument.

Regarding this subject, the question about an effective and economic observer training emerges. In many cases a training manual with a good developed set of rules and regulation is sufficient to familiarize the observers with the instrument and its underlying method. However, an additional observer training can be necessary to assure the quality of elicitation of data when observing a complex behavioral process.

The standardized video-observational instrument INTAKT assesses the quality of interaction between the caregiver and the child by observing the categories *Sensitivity*, *Feedback* and *Joint Attention*. Based on the observational category *Joint Attention* the observer agreement when recording the data with the INTAKT-manual and when recording the data after an additional training was compared. Therefore, observational data from 29 videos has been collected with the help of the manual. Then half of the videos were again coded using the manual. The recording of the other half was repeated after successfully completing an observer training. For the calculation of the *Intra-Rater-Reliability* and the *Interrater-Reliability* the statistical methods *Cohens Kappa* (κ) and *Percentual Agreement* (*per.A.*) were used.

The observer training did not lead to a significant change in the results regarding the category *Joint Attention* compared to the results recorded with the manual. The observer agreement when calculating the *Intra-Rater-Reliability* of the dimension *Joint Attention* was very good ($\kappa=.902$; *perc.A.*= 95,3%) when comparing the results recorded with manual and those recorded after the training. The *Interrater-Reliability* of the *Joint Attention* was also very high comparing the results when both observers recorded with the information of the training ($\kappa=.803$; *perc.A.*=90.9%), as well as the results when one observer recorded with manual and the other with the information of the training ($\kappa=.782$; *perc. A.*=89.9%). However, the *Interrater-Reliability* showed low agreements in the categories *Keine Joint Attention* and *Joint Attention Following* in both designs. Limitations and the possibility of the influence of extern factors on the

result of the observer agreement (e.g. prevalence) are discussed and should be further investigated in additional studies.

LITERATURVERZEICHNIS

- Ainsworth, M. D. S., Bell, S. M., & Stayton, D. J. (1974). Infant-mother attachment and social development: 'Socialisation' as a product of reciprocal responsiveness to signals. In M. P. M. Richards (Hrsg.), *The integration of a child into a social world* (S. 99-135). Cambridge: Cambridge University Press.
- Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge: Cambridge University Press.
- Bakeman, R., McArthur, D., Quera, V., & Robinson, B. (1997). Detecting sequential patterns and determining their reliability with fallible observers. *Psychological Methods*, 2(4), 357-370. doi: 10.1037//1082-989x.2.4.357
- Berk, L., & Schönplug, U. (2011). *Entwicklungspsychologie*. (5.Aufl.). München: Pearson Deutschland.
- Bischof-Köhler, D. (2011). *Soziale Entwicklung in Kindheit und Jugend*. Stuttgart: Kohlhammer Verlag.
- Bodenmann G. (2006). Beobachtungsmethoden. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (S.151-159). Göttingen: Hogrefe.
- Bortz, J., & Lienert, A.L. (2008). *Kurzgefasste Statistik für die klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben* (3.Aufl.) Heidelberg: Springer.
- Büttner, O., & Florack, A. (2009). In C. Baumgarth, M. Eisend & H. Evanschitzky (Hrsg.), *Empirische Mastertechniken. Eine anwendungsorientierte Einführung für die Marketing- und Managementforschung* (S.167-202). Wiesbaden: Gabler.
- Carpenter, M., & Liebal, K. (2011). Joint Attention, Communication and Knowing Together in Infancy. In A. Seemann (Hrsg.), *Joint Attention. New Developments in*

- Psychology, Philosophy of Mind, and Social Neuroscience* (S.159-182). Cambridge: MIT Press.
- Dinkelaker, J., & Herrle, M. (2009). *Erziehungswissenschaftliche Videographie – Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Döring, N., Bortz, J., & Pöschl, S. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5.Aufl.). Heidelberg: Springer.
- Fassnacht, G. (1995). *Systematische Verhaltensbeobachtung. Eine Einführung in die Methodologie und Praxis* (2.Aufl.). München: Reinhardt.
- Fieguth, G. (1977). Beobachtertraining. In U. Mees & H. Selg (Hrsg.), *Verhaltensbeobachtung und Verhaltensmodifikation* (S.78-86). Stuttgart: Klett.
- Fisseni, H. (1997). *Lehrbuch der psychologischen Diagnostik* (2. Aufl.). Göttingen: Hogrefe.
- Fisseni, H., & Fennekels, G. (1995). *Das Assessment-Center*. Göttingen: Verlag für Angewandte Psychologie.
- Gauvain, M. (2001). *The social context of cognitive development*. New York: Guilford Press.
- Girard, J., & Cohn, J. (2016). A Primer on Observational Measurement. *Assessment*, 23(4), 404-413. doi: 10.1177/1073191116635807
- Greve, W. & Wentura, D. (1997). *Wissenschaftliche Beobachtung. Eine Einführung*. Weinheim: Beltz.
- Grouven, U., Bender, R., Ziegler, A., & Lange, S. (2007). Der Kappa-Koeffizient. *DMW - Deutsche Medizinische Wochenschrift*, 132, e65-e68. doi: 10.1055/s-2007-959046

- Heikamp, T. (2014). Beobachtung. In G. Endruweit, G. Trommsdorff, & N. Burzan (Hrsg.), *Wörterbuch der Soziologie* (3.Aufl. S.48-51). Konstanz: UVK.
- Hirschmann, N., Kastner-Koller, U., Deimann, P., Aigner, N., & Svecz, T. (2011). IN-TAKT: A new instrument for assessing the quality of mother-child interactions. *Psychological Test and Assessment Modeling*, 53(3), 295-311.
- Hirschmann, N., Aigner, N., Deimann, P., & Kastner-Koller, U. (2017). *INTAKT. Ein Video-Beobachtungsinstrument zur Erfassung der Mutter-Kind-Interaktion-Manual* (4.Aufl.). Unveröffentlichtes Manuskript, Wien.
- Huber, P. (2013). *Zusammenhang zwischen mütterlicher Feinfühligkeit und kindlicher Entwicklung: ein Beitrag zur Validierung des Video-Beobachtungssystems IN-TAKT*. Diplomarbeit, Universität Wien.
- Jacob, A. (2016). *Interaktionsbeobachtung von Eltern und Kind*. (2. Aufl.). Stuttgart: Kohlhammer.
- König, A. (2013). Videographie. In M. Stamm & D. Edelmann (Hrsg.), *Handbuch frühkindliche Bildungsforschung* (S.817-829). Wiesbaden: Springer.
- Mandrekar, J.N. (2011). Measures of Interrater Agreement. *Journal Of Thoracic Oncology*, 6(1), 6-7. doi:10.1097/jto.0b013e318200f983
- McHugh, M. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 276-282. doi: 10.11613/bm.2012.031
- Mees, U. (1977). Methodologische Probleme der Verhaltensbeobachtung in der natürlichen Umgebung: I. Zuverlässigkeit und Generalisierbarkeit von Beobachtungsdaten. In U. Mees & H. Selg (Hrsg.), *Verhaltensbeobachtung und Verhaltensmodifikation* (S. 43-63). Stuttgart: Klett.
- Mees, U. (1977). Methodologische Probleme der Verhaltensbeobachtung in der natürlichen Umgebung: II. Beobachter und Beobachtete als mögliche Fehlerquellen

- von Beobachtungsdaten. In U. Mees & H. Selg (Hrsg.), *Verhaltensbeobachtung und Verhaltensmodifikation* (S. 66-74). Stuttgart: Klett.
- Mees, U. (1977). Zur Validität von Verhaltensbeobachtungen. In U. Mees & H. Selg (Hrsg.), *Verhaltensbeobachtung und Verhaltensmodifikation* (S. 88-93). Stuttgart: Klett.
- Laatz, W. (1993). *Empirische Methoden. Ein Lehrbuch für Sozialwissenschaftler*. Thun: Harri Deutsch.
- Medistat. Medizinische Statistik. (2018). Abgerufen am 02. Februar 2018 von <https://www.medistat.de/glossar/uebereinstimmung/cohens-kappa-koeffizient/>
- Moosbrugger, H. & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kevala (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Aufl., S.8-25). Heidelberg: Springer.
- Obermann, C. (2018). *Assessment Center. Entwicklung, Durchführung, Trends*. (6. Aufl.). Wiesbaden: Springer Gabler.
- Papousek, M. (2000). Einsatz von Video in der Eltern-Säuglings-Beratung und Psychotherapie. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 49, 611-627.
- Rentzsch, K., & Schütz, A. (2009). *Psychologische Diagnostik. Grundlagen und Anwendungsperspektiven*. Stuttgart: Kohlhammer.
- Reischer, N. (2013). *Zusammenhang zwischen mütterlicher Rückmeldung und kindlicher Entwicklung: ein Beitrag zur Validierung des Video-Beobachtungssystems IN-TAKT*. Diplomarbeit, Universität Wien.
- Renner, K. H. (2013). Verhaltensbeobachtung Beobachtungsfehler und Beobachtertraining. In W. Sages (Hrsg.), *Management-Diagnostik* (4. Aufl., S.656-663). Göttingen: Hogrefe.

- Roch, S., Woehr, D., Mishra, V., & Kieszczyńska, U. (2011). Rater training revisited: An updated meta-analytic review of frame-of-reference training. *Journal Of Occupational And Organizational Psychology*, 85(2), 370-395. doi: 10.1111/j.2044-8325.2011.02045.x
- Saxon, T. F., Colombo, J., Robinson, E. L. & Frick, J. E. (2000). Dyadic interaction profiles in infancy and preschool intelligence. *Journal of School Psychology*, 38, 9-25.
- Schölmerich, A. (2011). Verhaltensbeobachtung. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (4. Aufl., S. 768-789). Bern: Huber.
- Schmücker, G. & Buchheim, A. (2002). Mutter-Kind-Interaktion und Bindung in den ersten Lebensjahren. In B. Strauß, A. Buchheim & H. Kächele (Hrsg.), *Klinische Bindungsforschung. Theorien- Methoden- Ergebnisse* (S. 173-190). Stuttgart: Schattauer.
- Seidel, T., & Prenzel, M. (2010). Beobachtungsverfahren: Vom Datenmaterial zur Datenanalyse. In: H. Holling and B. Schmitz (Hrsg.), *Handbuch Statistik, Methoden und Evaluation* (S. 139-152). Göttingen: Hogrefe.
- Spinath, F.M., & Becker, N. (2011). Verhaltensbeobachtung. In L.F. Hornke, M. Amelang & M. Kersting (Hrsg.), *Methoden der psychologischen Diagnostik* (S.326-367). Göttingen: Hogrefe.
- Thiel-Bonney, C. (2002). Beratung von Eltern mit Säuglingen und Kleinkindern. *Psychotherapeut*, 47(6), 381-384. doi: 10.1007/s00278-002-0265-2
- Thiel, T. (2003). Film- und Videotechnik in der Psychologie. Eine erkenntnistheoretische Analyse mit Jean Piaget, Anwendungsbeispiele aus der Kleinkind-Forschung und ein historischer Rückblick auf Kurt Lewin und Arnold Gesell. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (3.Aufl., S.649-708). Bern: Huber.

- Thiel, T. (2011). Film- und Videotechnik in der Psychologie. Eine Entwicklungsgeschichte aus erkenntnistheoretisch-methodischer Perspektive. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (4.Aufl., S.792-818). Bern: Huber.
- Thornton, G., & Zorich, S. (1980). Training to improve observer accuracy. *Journal Of Applied Psychology*, 65(3), 351-354. <http://dx.doi.org/10.1037/0021-9010.65.3.351>
- Tomasello, M. (1995). Joint Attention and Social Cognititon. In C. Moore & P. Dunham (Hrsg.), *Joint Attention. Its Origin and Role in Development* (S. 103-130). New York: Psychology Press.
- Trudewind, C., Unzner, L. & Schneider K. (1997). Die Entwicklung der Leistungsmotivation. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (2. Aufl., S. 587-622). Bern: Huber.
- Ulber, D., & Imhof, M. (2014). *Beobachtung in der Frühpädagogik. Theoretische Grundlagen, Methoden, Anwendung*. Stuttgart: Kohlhammer.
- Vogler, J. (2015). *Zusammenhang zwischen der Joint Attention und kindlicher Entwicklung: ein Beitrag zur Validierung des Video-Beobachtungssystems INTAKT*. Diplomarbeit, Universität Wien.
- Wicki, W. (2015). *Entwicklungspsychologie* (2.Aufl.). München [u.a.]: Reinhardt.
- Wirtz, M. (2006). Methoden zur Bestimmung der Beurteilerübereinstimmung. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Klinischen Diagnostik*. (S. 369-382). Göttingen: Hogrefe.
- Wirtz, M., & Kutschmann, M. (2007). Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Maße. *Die Rehabilitation*, 46(6), 370-377. doi: 10.1055/s-2007-976535

Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.

Woehr, D., & Huffcutt, A. (1994). Rater training for performance appraisal: A quantitative review. *Journal Of Occupational And Organizational Psychology*, 67(3), 189-205. doi: 10.1111/j.2044-8325.1994.tb00562

TABELLENVERZEICHNIS

Tabelle 1: Beurteilung mit Manual: Häufigkeiten und Anteilswerte der Beobachtungskategorien	61
Tabelle 2: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T1_A & T2_A (beide mit Manual kodiert)	62
Tabelle 3: Häufigkeiten und Anteilswerte der Beobachtungskategorien: Beurteilung mit Manual (T1_B) im Vergleich zu Beurteilung mit Training (T2_B)	65
Tabelle 4: Intra-Rater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T1_B & T2_B (T1_B mit Manual kodiert; T2_B mit zusätzlichem Beobachtungstraining)	66
Tabelle 5: Intra-Rater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung der Beobachtungsdimension Joint Attention insgesamt & der Kategorien JA, swJA, kJA	68
Tabelle 6: Häufigkeiten und Anteilswerte der Beobachtungskategorien: Beurteilung mit Training (T2_B & H)	69
Tabelle 7: Interrater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T2_B & H (beide mit Manual und zusätzlichem Beobachtungstraining kodiert)	71
Tabelle 8: Häufigkeiten und Anteilswerte der Beobachtungskategorien: Beurteilung mit Manual (T1_B) im Vergleich zu Beurteilung mit Beobachtungstraining (H)	73
Tabelle 9: Interrater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung von den Videos der Gruppen T1_B & H (T1_B mit Manual kodiert; H mit zusätzlichem Beobachtungstraining)	74
Tabelle 10: Interrater-Reliabilität: Cohens-Kappa und Prozentuale Übereinstimmung der Beobachtungsdimension Joint Attention insgesamt & der Kategorien JA, swJA, kJA	75
Tabelle 11: Zusammengefasste Prozentuale Übereinstimmung: T2_B & H (beide mit Manual und zusätzlichem Beobachtungstraining kodiert)	77

ABBILDUNGSVERZEICHNIS

Abbildung 1. Erhebungsphasen der Untersuchung.....55

ABKÜRZUNGSVERZEICHNIS

PAS	→ Predominant-Activity-Sampling
OZS	→ One-Zero-Sampling
K1_JA	→ Kodierdurchgang1_Joint Attention (PAS)
K2_swJA	→ Kodierdurchgang2_ Switching Joint Attention (OZS)
K2_kJA	→ Kodierdurchgang2_ keine Joint Attention (OZS)
T1_A	→ Erhebungsphase 1/ Videos der Gruppe A kodiert von T.P. mit Manual
T1_B	→ Erhebungsphase 1/ Videos der Gruppe B kodiert von T.P. mit Manual
T2_A	→ Erhebungsphase 1/ Videos der Gruppe A kodiert von T.P. mit Manual
T2_B	→ Erhebungsphase 1/ Videos der Gruppe B kodiert von T.P. mit Manual und Beobachtungstraining
H	→ Videos kodiert von H.S. mit Manual und Beobachtungstraining
Proz.Ü.	→ Prozentuale Übereinstimmung

ANHANGSVERZEICHNIS

<u>I.</u>	Übersicht der Videoaufzeichnungen	102
<u>II.</u>	Berechnung der Verteilungsformen: Normalverteilung, Kurtosis, Schiefe.....	103
<u>III.</u>	Intra-Rater-Reliabilität Kontingenztabelle: Gesamt Joint Attention & Gesamt Kategorien Joint Attention Following, Switching Joint Attention und Keine Joint Attention	103
<u>IV.</u>	Interrater-Reliabilität Kontingenztabelle: Gesamt Joint Attention & Gesamt Kategorien Joint Attention Following, Switching Joint Attention und Keine Joint Attention	105
<u>V.</u>	Intra-Rater-Reliabilität K1_ JA: Kontingenztabelle pro Video (T1_A & T2_A).....	106
<u>VI.</u>	Intra-Rater-Reliabilität K1_ JA: Kontingenztabelle pro Video (T1_B & T2_B).....	107
<u>VII.</u>	Intra-Rater-Reliabilität K2_ swJA: Kontingenztabelle pro Video (T1_A & T2_A)	109
<u>VIII.</u>	Intra-Rater-Reliabilität K2_ swJA: Kontingenztabelle pro Video (T1_B & T2_B)	110
<u>IX.</u>	Intra-Rater-Reliabilität K2_ kJA: Kontingenztabelle pro Video (T1_A & T2_A).....	112
<u>X.</u>	Intra-Rater-Reliabilität K2_ kJA: Kontingenztabelle pro Video (T1_B & T2_B).....	113
<u>XI.</u>	Interrater-Reliabilität K1_ JA: Kontingenztabelle pro Video (T2_B & H).....	115
<u>XII.</u>	Interrater-Reliabilität K1_ JA: Kontingenztabelle pro Video (T1_B & H)	117
<u>XIII.</u>	Interrater-Reliabilität K2_ swJA: Kontingenztabelle pro Video (T2_B & H)	118
<u>XIV.</u>	Interrater-Reliabilität K2_ swJA: Kontingenztabelle pro Video (T1_B & H)	120
<u>XV.</u>	Interrater-Reliabilität K2_ kJA: Kontingenztabelle pro Video (T2_B & H).....	121
<u>XVI.</u>	Interrater-Reliabilität K2_ kJA: Kontingenztabelle pro Video (T1_B & H).....	123
<u>XVII.</u>	Beobachtungsbogen Joint Attention (erste Seite)	124

ANHANG

I. Übersicht der Videoaufzeichnungen

Videocode	Dauer (min:sek)	T1_A & T1_B T.Pilshofer Ma- nual Phase1	T2_A T.Pilshofer Manual Phase2	T2_B T.Pilshofer Einschulung	H H.Schulz Einschulung	Video auf- gezeichnet von:
JV2	57:16	X	X			J.Vogler
JV4	37:26	X	X			J.Vogler
JV6	62:25	X	X			J.Vogler
JV8	54:19	X	X			J.Vogler
JV10	37:21	X	X			J.Vogler
JV14	31:16	X	X			J.Vogler
JV16	35:27	X	X			J.Vogler
JV18	30:30	X	X			J.Vogler
JV20	47:21	X	X			J.Vogler
NR1	63:16	X	X			N.Reischer
NR3	72:52	X	X			N.Reischer
NR5	52:57	X	X			N.Reischer
NR7	54:37	X	X			N.Reischer
NR9	56:50	X	X			N.Reischer
NR11	59:24	X		X	X	N.Reischer
NR13	51:59	X		X	X	N.Reischer
NR15	34:27	X		X	X	N.Reischer
NR17	45:53	X		X	X	N.Reischer
NR19	69:10	X		X	X	N.Reischer
PH1	37:33	X		X	X	P.Huber
PH3	40:12	X		X	X	P.Huber
PH5	33:45	X		X	X	P.Huber
PH7	45:15	X		X	X	P.Huber
PH9	69:10	X		X	X	P.Huber
PH11	61:00	X		X	X	P.Huber
PH13	27:36	X		X	X	P.Huber
PH15	47:54	X		X	X	P.Huber
PH17	41:59	X		X	X	P.Huber
PH19	52:18	X		X	X	P.Huber
Durch- schnitt	47:48					

II. Berechnung der Verteilungsformen: Normalverteilung, Kurtosis, Schiefe

	Kolmogorov-Smirnov			Shapiro-Wilk			Schiefe	Kurtosis
	Stat.	df	Sig.	Stat.	df	Sig.		
K1_JA_T1_A	.528	2685	.000	.314	2685	.000	3.664	13.517
K1_JA_T2_A	.531	2686	.000	.292	2686	.000	3.929	15.809
K1_JA_T1_B	.529	2727	.000	.334	2727	.000	3.252	10.504
K1_JA_T2_B	.535	2731	.000	.259	2731	.000	4.325	19.703
K2SW_T1_A	.484	4034	.000	.508	4034	.000	1.373	-0.116
K2SW_T2_A	.490	4029	.000	.495	4029	.000	1.468	.154
K2SW_T1_B	.447	4100	.000	.571	4100	.000	3.252	10.504
K2SW_T2_B	.452	4103	.000	.564	4103	.000	4.323	19.676
K2kJA_T1_A	.539	4035	.000	.153	4035	.000	5.696	30.463
K2kJA_T2_A	.537	4030	.000	.123	4030	.000	6.711	43.052
K2kJA_T1_B	.538	4103	.000	.139	4103	.000	6.138	35.694
K2kJA_T2_B	.533	4102	.000	.092	4102	.000	8.308	67.062
K1_JA_H	.537	2707	.000	.142	2707	.000	6.851	51.374
K2_swJA_H	.491	4131	.000	.491	4131	.000	1.491	0.225
K2_kJA_H	.526	4131	.000	.055	4131	.000	11.813	137.623

Stat. → Statistik
df → Freiheitsgrade
Sig. → Signifikanz

III. Intra-Rater-Reliabilität Kontingenztabelle: Gesamt Joint Attention & Gesamt Kategorien
Joint Attention Following, Switching Joint Attention und Keine Joint Attention

JA_A_T1Gesamt * JA_A_T2Gesamt Kreuztabelle						JA_B_T1Gesamt * JA_B_T2Gesamt Kreuztabelle						
Anzahl	JA_A_T2Gesamt					Gesamt	Anzahl	JA_B_T2Gesamt				Gesamt
	1,00	2,00	3,00	Gesamt				1,00	2,00	3,00	Gesamt	
JA_A_T1Gesamt	1,00	2387	65	0	2452	JA_B_T1Gesamt	1,00	2441	18	11	2470	
	2,00	84	99	1	184		2,00	108	129	0	237	
	3,00	8	2	38	48		3,00	4	2	13	19	
Gesamt		2479	166	39	2684	Gesamt		2553	149	24	2726	

K2SW_T1_A * K2SW_T2_A Kreuzta- belle				K2SW_T1_B * K2SW_T2_B Kreuztab- elle					
Anzahl				Anzahl					
		K2SW_T2_A				K2SW_T2_B			
		,00	1,00	Gesamt			Gesamt		
K2SW_T1_	,00	3019	130	3149	K2SW_T1_	,00	2763	134	2897
A	1,00	183	692	875	B	1,00	175	1024	1199
Gesamt		3202	822	4024	Gesamt		2938	1158	4096
K2kJA_T1_A * K2kJA_T2_A Kreuztab- elle				K2kJA_T2_B * K2kJA_T1_B Kreuztabelle					
Anzahl				Anzahl					
		K2kJA_T2_A				K2kJA_T1_B			
		,00	1,00	Gesamt			Gesamt		
K2kJA_T1_	,00	3908	5	3913	K2kJA_T2_	,00	3989	51	4040
A	1,00	35	78	113	B	1,00	6	50	56
Gesamt		3943	83	4026	Gesamt		3995	101	4096
JA_K1_2_T1_A * JA_K1_2_T2_A Kreuztabelle									
Anzahl									
		JA_K1_2_T2_A							
		,00	1,00	2,00	3,00	Gesamt			
JA_K1_2_T1_A	,00	6927	135	0	0	7062			
	1,00	218	3157	65	0	3440			
	2,00	0	84	99	1	184			
	3,00	0	8	2	38	48			
Gesamt		7145	3384	166	39	10734			
JA_K1u2_T1_B * JA_K1u2_T2_B Kreuztabelle									
Anzahl									
		JA_K1u2_T2_B							
		,00	1,00	2,00	3,00	Gesamt			
JA_K1u2_T1_B	,00	6752	140	0	0	6892			
	1,00	226	3515	18	11	3770			
	2,00	0	108	129	0	237			
	3,00	0	4	2	13	19			
Gesamt		6978	3767	149	24	10918			

IV. Interrater-Reliabilität Kontingenztabelle: Gesamt Joint Attention & Gesamt Kategorien
 Joint Attention Following, Switching Joint Attention und Keine Joint Attention

NeuJA_B_HGesamt * NeuJA_B_T2Gesamt Kreuztabelle						JA_H_1 * K1_JA_T1_B Kreuztabelle					
Anzahl						Anzahl					
	NeuJA_B_T2Gesamt				Ge- samt		K1_JA_T1_B				Gesamt
	1,00	2,00	3,00			1,00	2,00	3,00			
NeuJA_B_1,00	2475	125	25		2625	K1_JA_H_1 1,00	2424	187	19	2630	
HGesamt 2,00	27	38	0		65	2,00	21	42	0	63	
3,00	4	1	1		6	3,00	5	0	0	5	
Gesamt	2506	164	26		2696	Gesamt	2450	229	19	2698	
K2SW_H * K2SW_T2_B Kreuztabelle						K2_swJA_H_B * K2_swJA_T1_B Kreuztabelle					
Anzahl						Anzahl					
	K2SW_T2_B			Gesamt		K2_swJA_T1_B			Gesamt		
	,00	1,00			,00	1,00					
K2SW_H ,00	2753	520		3273	K2_swJA_H_B ,00	2713	559		3272		
1,00	190	640		830	1,00	188	641		829		
Gesamt	2943	1160		4103	Gesamt	2901	1200		4101		
K2kJA_T2_B * K2kJA_H Kreuztabelle						K2_swJA_H_B * K2_swJA_T1_B Kreuztabelle					
Anzahl						Anzahl					
	K2kJA_H			Ge- samt		K2_swJA_T1_B			Gesamt		
	,00	1,00			,00	1,00					
K2kJA_T2_B ,00	4024	21		4045	K2_swJA_H_B ,00	2713	559		3272		
1,00	49	8		57	1,00	188	641		829		
Gesamt	4073	29		4102	Gesamt	2901	1200		4101		
JA_K1u2_H * JA_K1u2_T2_B Kreuztabelle											
Anzahl											
	JA_K1u2_T2_B					Gesamt					
	,00	1,00	2,00	3,00							
JA_K1u2_H ,00	6736	613	0	0		7349					
1,00	210	3150	109	23		3492					
2,00	0	29	35	0		64					
3,00	0	5	0	1		6					
Gesamt	6946	3797	144	24		10911					

JA_K1u2_T1_B * JA_K1u2_H Kreuztabelle						
Anzahl						
		JA_K1u2_H				
		,00	1,00	2,00	3,00	Gesamt
JA_K1u2_T1_B	,00	6690	213	0	0	6903
	1,00	656	3068	21	5	3750
	2,00	0	187	42	0	229
	3,00	0	19	0	0	19
Gesamt		7346	3487	63	5	10901

V. Intra-Rater-Reliabilität K1_JA: Kontingenztabellen pro Video (T1_A & T2_A)

<p>JV2_T1 * JV2_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV2_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV2_T1</th> <th>1,00</th> <td>205</td> <td>3</td> <td>208</td> </tr> <tr> <th>2,00</th> <td>13</td> <td>4</td> <td>17</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>218</td> <td>7</td> <td>225</td> </tr> </tbody> </table>			JV2_T2					1,00	2,00	Gesamt	JV2_T1	1,00	205	3	208	2,00	13	4	17	Gesamt		218	7	225	<p>JV4_T1 * JV4_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV4_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV4_T1</th> <th>1,00</th> <td>119</td> <td>4</td> <td>123</td> </tr> <tr> <th>2,00</th> <td>9</td> <td>10</td> <td>19</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>128</td> <td>14</td> <td>142</td> </tr> </tbody> </table>			JV4_T2					1,00	2,00	Gesamt	JV4_T1	1,00	119	4	123	2,00	9	10	19	Gesamt		128	14	142										
		JV2_T2																																																									
		1,00	2,00	Gesamt																																																							
JV2_T1	1,00	205	3	208																																																							
	2,00	13	4	17																																																							
Gesamt		218	7	225																																																							
		JV4_T2																																																									
		1,00	2,00	Gesamt																																																							
JV4_T1	1,00	119	4	123																																																							
	2,00	9	10	19																																																							
Gesamt		128	14	142																																																							
<p>JV6_T1 * JV6_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV6_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV6_T1</th> <th>1,00</th> <td>216</td> <td>3</td> <td>219</td> </tr> <tr> <th>2,00</th> <td>11</td> <td>12</td> <td>23</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>227</td> <td>15</td> <td>242</td> </tr> </tbody> </table>			JV6_T2					1,00	2,00	Gesamt	JV6_T1	1,00	216	3	219	2,00	11	12	23	Gesamt		227	15	242	<p>JV8_T1 * JV8_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="3">JV8_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>3,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="3">JV8_T1</th> <th>1,00</th> <td>140</td> <td>0</td> <td>0</td> <td>140</td> </tr> <tr> <th>2,00</th> <td>1</td> <td>1</td> <td>0</td> <td>2</td> </tr> <tr> <th>3,00</th> <td>7</td> <td>0</td> <td>36</td> <td>43</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>148</td> <td>1</td> <td>36</td> <td>185</td> </tr> </tbody> </table>			JV8_T2						1,00	2,00	3,00	Gesamt	JV8_T1	1,00	140	0	0	140	2,00	1	1	0	2	3,00	7	0	36	43	Gesamt		148	1	36	185
		JV6_T2																																																									
		1,00	2,00	Gesamt																																																							
JV6_T1	1,00	216	3	219																																																							
	2,00	11	12	23																																																							
Gesamt		227	15	242																																																							
		JV8_T2																																																									
		1,00	2,00	3,00	Gesamt																																																						
JV8_T1	1,00	140	0	0	140																																																						
	2,00	1	1	0	2																																																						
	3,00	7	0	36	43																																																						
Gesamt		148	1	36	185																																																						
<p>JV10_T1 * JV10_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV10_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV10_T1</th> <th>1,00</th> <td>128</td> <td>5</td> <td>133</td> </tr> <tr> <th>2,00</th> <td>3</td> <td>3</td> <td>6</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>131</td> <td>8</td> <td>139</td> </tr> </tbody> </table>			JV10_T2					1,00	2,00	Gesamt	JV10_T1	1,00	128	5	133	2,00	3	3	6	Gesamt		131	8	139	<p>JV14_T1 * JV14_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV14_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV14_T1</th> <th>1,00</th> <td>102</td> <td>4</td> <td>106</td> </tr> <tr> <th>2,00</th> <td>5</td> <td>5</td> <td>10</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>107</td> <td>9</td> <td>116</td> </tr> </tbody> </table>			JV14_T2					1,00	2,00	Gesamt	JV14_T1	1,00	102	4	106	2,00	5	5	10	Gesamt		107	9	116										
		JV10_T2																																																									
		1,00	2,00	Gesamt																																																							
JV10_T1	1,00	128	5	133																																																							
	2,00	3	3	6																																																							
Gesamt		131	8	139																																																							
		JV14_T2																																																									
		1,00	2,00	Gesamt																																																							
JV14_T1	1,00	102	4	106																																																							
	2,00	5	5	10																																																							
Gesamt		107	9	116																																																							
<p>JV16_T1 * JV16_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV16_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV16_T1</th> <th>1,00</th> <td>131</td> <td>0</td> <td>131</td> </tr> <tr> <th>2,00</th> <td>2</td> <td>2</td> <td>4</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>133</td> <td>2</td> <td>135</td> </tr> </tbody> </table>			JV16_T2					1,00	2,00	Gesamt	JV16_T1	1,00	131	0	131	2,00	2	2	4	Gesamt		133	2	135	<p>JV18_T1 * JV18_T2 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2"></th> <th colspan="2">JV18_T2</th> <th></th> </tr> <tr> <th colspan="2"></th> <th>1,00</th> <th>2,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th rowspan="2">JV18_T1</th> <th>1,00</th> <td>106</td> <td>3</td> <td>109</td> </tr> <tr> <th>2,00</th> <td>6</td> <td>4</td> <td>10</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>112</td> <td>7</td> <td>119</td> </tr> </tbody> </table>			JV18_T2					1,00	2,00	Gesamt	JV18_T1	1,00	106	3	109	2,00	6	4	10	Gesamt		112	7	119										
		JV16_T2																																																									
		1,00	2,00	Gesamt																																																							
JV16_T1	1,00	131	0	131																																																							
	2,00	2	2	4																																																							
Gesamt		133	2	135																																																							
		JV18_T2																																																									
		1,00	2,00	Gesamt																																																							
JV18_T1	1,00	106	3	109																																																							
	2,00	6	4	10																																																							
Gesamt		112	7	119																																																							

JV20_T1 * JV20_T2 Kreuztabelle					NR1_T1 * NR1_T2 Kreuztabelle					
Anzahl					Anzahl					
					NR1_T2					
					1,00 2,00 3,00 Gesamt					
JV20_T	1,00	169	7	176	NR1_T	1,00	224	7	0	231
1	2,00	3	5	8	1	2,00	7	7	1	15
Gesamt		172	12	184	Gesamt	3,00	1	2	2	5
					Gesamt		232	16	3	251

NR3_T1 * NR3_T2 Kreuztabelle					NR5_T1 * NR5_T2 Kreuztabelle				
Anzahl					Anzahl				
					NR3_T2				
					1,00 2,00 Gesamt				
NR3_T	1,00	287	1	288	NR5_T	1,00	165	10	175
1	2,00	2	0	2	1	2,00	12	25	37
Gesamt		289	1	290	Gesamt		177	35	212

NR7_T1 * NR7_T2 Kreuztabelle					NR9_T1 * NR9_T2 Kreuztabelle				
Anzahl					Anzahl				
					NR7_T2				
					1,00 2,00 Gesamt				
NR7_T1	1,00	177	13	190	NR9_T	1,00	218	5	223
	2,00	9	18	27	1	2,00	1	3	4
Gesamt		186	31	217	Gesamt		219	8	227

VI. Intra-Rater-Reliabilität K1_JA: Kontingenztabellen pro Video (T1_B & T2_B)

NR11_T1 * NR11_T2_E Kreuztabelle					NR13_T1 * NR13_T2_E Kreuztabelle				
Anzahl					Anzahl				
					NR13_T2_E				
					1,00 2,00 Gesamt				
NR11_T1	1,00	223	6	229	NR13_T1	1,00	190	0	190
	2,00	2	4	6		2,00	7	6	13
Gesamt		225	10	235	Gesamt		197	6	203

<p>NR15_T1 * NR15_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="3">NR15_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> <th>3,00</th> </tr> </thead> <tbody> <tr> <th rowspan="3">NR15_T1</th> <th>1,00</th> <td>111</td> <td>2</td> <td>1</td> <td>114</td> </tr> <tr> <th>2,00</th> <td>8</td> <td>6</td> <td>0</td> <td>14</td> </tr> <tr> <th>3,00</th> <td>2</td> <td>0</td> <td>8</td> <td>10</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>121</td> <td>8</td> <td>9</td> <td>138</td> </tr> </tbody> </table>			NR15_T2_E			Gesamt	1,00	2,00	3,00	NR15_T1	1,00	111	2	1	114	2,00	8	6	0	14	3,00	2	0	8	10	Gesamt		121	8	9	138	<p>NR17_T1 * NR17_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">NR17_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">NR17_T1</th> <th>1,00</th> <td>174</td> <td>0</td> <td>174</td> </tr> <tr> <th>2,00</th> <td>3</td> <td>5</td> <td>8</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>177</td> <td>5</td> <td>182</td> </tr> </tbody> </table>			NR17_T2_E		Gesamt	1,00	2,00	NR17_T1	1,00	174	0	174	2,00	3	5	8	Gesamt		177	5	182					
			NR15_T2_E				Gesamt																																																			
		1,00	2,00	3,00																																																						
NR15_T1	1,00	111	2	1	114																																																					
	2,00	8	6	0	14																																																					
	3,00	2	0	8	10																																																					
Gesamt		121	8	9	138																																																					
		NR17_T2_E		Gesamt																																																						
		1,00	2,00																																																							
NR17_T1	1,00	174	0	174																																																						
	2,00	3	5	8																																																						
Gesamt		177	5	182																																																						
<p>NR19_T1 * NR19_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">NR19_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">NR19_T1</th> <th>1,00</th> <td>154</td> <td>0</td> <td>154</td> </tr> <tr> <th>2,00</th> <td>4</td> <td>5</td> <td>9</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>158</td> <td>5</td> <td>163</td> </tr> </tbody> </table>			NR19_T2_E		Gesamt	1,00	2,00	NR19_T1	1,00	154	0	154	2,00	4	5	9	Gesamt		158	5	163	<p>PH1_T1 * PH1_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">PH1_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">PH1_T1</th> <th>1,00</th> <td>141</td> <td>0</td> <td>141</td> </tr> <tr> <th>2,00</th> <td>3</td> <td>4</td> <td>7</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>144</td> <td>4</td> <td>148</td> </tr> </tbody> </table>			PH1_T2_E		Gesamt	1,00	2,00	PH1_T1	1,00	141	0	141	2,00	3	4	7	Gesamt		144	4	148															
			NR19_T2_E			Gesamt																																																				
		1,00	2,00																																																							
NR19_T1	1,00	154	0	154																																																						
	2,00	4	5	9																																																						
Gesamt		158	5	163																																																						
		PH1_T2_E		Gesamt																																																						
		1,00	2,00																																																							
PH1_T1	1,00	141	0	141																																																						
	2,00	3	4	7																																																						
Gesamt		144	4	148																																																						
<p>PH3_T1 * PH3_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="3">PH3_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> <th>3,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">PH3_T1</th> <th>1,00</th> <td>152</td> <td>1</td> <td>4</td> <td>157</td> </tr> <tr> <th>2,00</th> <td>1</td> <td>2</td> <td>0</td> <td>3</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>153</td> <td>3</td> <td>4</td> <td>160</td> </tr> </tbody> </table>			PH3_T2_E			Gesamt	1,00	2,00	3,00	PH3_T1	1,00	152	1	4	157	2,00	1	2	0	3	Gesamt		153	3	4	160	<p>PH5_T1 * PH5_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">PH5_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">PH5_T1</th> <th>1,00</th> <td>129</td> <td>0</td> <td>129</td> </tr> <tr> <th>2,00</th> <td>0</td> <td>4</td> <td>4</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>129</td> <td>4</td> <td>133</td> </tr> </tbody> </table>			PH5_T2_E		Gesamt	1,00	2,00	PH5_T1	1,00	129	0	129	2,00	0	4	4	Gesamt		129	4	133										
			PH3_T2_E				Gesamt																																																			
		1,00	2,00	3,00																																																						
PH3_T1	1,00	152	1	4	157																																																					
	2,00	1	2	0	3																																																					
Gesamt		153	3	4	160																																																					
		PH5_T2_E		Gesamt																																																						
		1,00	2,00																																																							
PH5_T1	1,00	129	0	129																																																						
	2,00	0	4	4																																																						
Gesamt		129	4	133																																																						
<p>PH7_T1 * PH7_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">PH7_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">PH7_T1</th> <th>1,00</th> <td>172</td> <td>1</td> <td>173</td> </tr> <tr> <th>2,00</th> <td>4</td> <td>3</td> <td>7</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>176</td> <td>4</td> <td>180</td> </tr> </tbody> </table>			PH7_T2_E		Gesamt	1,00	2,00	PH7_T1	1,00	172	1	173	2,00	4	3	7	Gesamt		176	4	180	<p>PH9_T1 * PH9_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">PH9_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">PH9_T1</th> <th>1,00</th> <td>244</td> <td>6</td> <td>250</td> </tr> <tr> <th>2,00</th> <td>5</td> <td>22</td> <td>27</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>249</td> <td>28</td> <td>277</td> </tr> </tbody> </table>			PH9_T2_E		Gesamt	1,00	2,00	PH9_T1	1,00	244	6	250	2,00	5	22	27	Gesamt		249	28	277															
			PH7_T2_E			Gesamt																																																				
		1,00	2,00																																																							
PH7_T1	1,00	172	1	173																																																						
	2,00	4	3	7																																																						
Gesamt		176	4	180																																																						
		PH9_T2_E		Gesamt																																																						
		1,00	2,00																																																							
PH9_T1	1,00	244	6	250																																																						
	2,00	5	22	27																																																						
Gesamt		249	28	277																																																						
<p>PH11_T1 * PH11_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="3">PH11_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> <th>3,00</th> </tr> </thead> <tbody> <tr> <th rowspan="3">PH11_T1</th> <th>1,00</th> <td>209</td> <td>1</td> <td>0</td> <td>210</td> </tr> <tr> <th>2,00</th> <td>17</td> <td>14</td> <td>0</td> <td>31</td> </tr> <tr> <th>3,00</th> <td>0</td> <td>0</td> <td>3</td> <td>3</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>226</td> <td>15</td> <td>3</td> <td>244</td> </tr> </tbody> </table>			PH11_T2_E			Gesamt	1,00	2,00	3,00	PH11_T1	1,00	209	1	0	210	2,00	17	14	0	31	3,00	0	0	3	3	Gesamt		226	15	3	244	<p>PH13_T1 * PH13_T2_E Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th> <th colspan="3">PH13_T2_E</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>1,00</th> <th>2,00</th> <th>3,00</th> </tr> </thead> <tbody> <tr> <th rowspan="2">PH13_T1</th> <th>1,00</th> <td>85</td> <td>0</td> <td>2</td> <td>87</td> </tr> <tr> <th>2,00</th> <td>9</td> <td>13</td> <td>0</td> <td>22</td> </tr> <tr> <th colspan="2">Gesamt</th> <td>94</td> <td>13</td> <td>2</td> <td>109</td> </tr> </tbody> </table>			PH13_T2_E			Gesamt	1,00	2,00	3,00	PH13_T1	1,00	85	0	2	87	2,00	9	13	0	22	Gesamt		94	13	2	109
			PH11_T2_E				Gesamt																																																			
		1,00	2,00	3,00																																																						
PH11_T1	1,00	209	1	0	210																																																					
	2,00	17	14	0	31																																																					
	3,00	0	0	3	3																																																					
Gesamt		226	15	3	244																																																					
		PH13_T2_E			Gesamt																																																					
		1,00	2,00	3,00																																																						
PH13_T1	1,00	85	0	2	87																																																					
	2,00	9	13	0	22																																																					
Gesamt		94	13	2	109																																																					

PH15_T1 * PH15_T2_E Kreuztabelle						PH17_T1 * PH17_T2_E Kreuztabelle					
Anzahl						Anzahl					
		PH15_T2_E			Ge- samt			PH17_T2_E			Ge- samt
		1,00	2,00	3,00				1,00	2,00	3,00	
PH15_T1	1,00	141	1	2	144	PH17_T1	1,00	140	0	1	141
	2,00	15	21	0	36		2,00	16	6	0	22
	3,00	1	0	1	2		3,00	1	0	0	1
Gesamt		157	22	3	182	Gesamt		157	6	1	164

PH19_T1 * PH19_T2_E Kreuztabelle					
Anzahl					
		PH19_T2_E			Ge- samt
		1,00	2,00	3,00	
PH19_T1	1,00	175	0	1	176
	2,00	14	14	0	28
	3,00	0	2	1	3
Gesamt		189	16	2	207

VII. Intra-Rater-Reliabilität K2_swJA: Kontingenztabellen pro Video (T1_A & T2_A)

NR9_T2 * NR9_T1 Kreuztabelle					NR7_T2 * NR7_T1 Kreuztabelle				
Anzahl					Anzahl				
		NR9_T1		Gesamt			NR7_T1		Gesamt
		,00	1,00				,00	1,00	
NR9_T2	,00	272	9	281	NR7_T2	,00	184	18	202
	1,00	11	48	59		1,00	16	107	123
Gesamt		283	57	340	Gesamt		200	125	325

NR5_T2 * NR5_T1 Kreuztabelle					NR3_T2 * NR3_T1 Kreuztabelle				
Anzahl					Anzahl				
		NR5_T1		Gesamt			NR3_T1		Gesamt
		,00	1,00				,00	1,00	
NR5_T2	,00	174	22	196	NR3_T2	,00	383	12	395
	1,00	10	112	122		1,00	10	29	39
Gesamt		184	134	318	Gesamt		393	41	434

NR1_T2 * NR1_T1 Kreuztabelle					JV20_T2 * JV20_T1 Kreuztabelle				
Anzahl					Anzahl				
		NR1_T1		Gesamt			JV20_T1		Gesamt
		,00	1,00				,00	1,00	
NR1_T2	,00	292	15	307	JV20_T2	,00	222	11	233
	1,00	15	54	69		1,00	11	32	43
Gesamt		307	69	376	Gesamt		233	43	276

<p align="center">JV18_T2 * JV18_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV18_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV18_T2 ,00</th> <td>115</td> <td>6</td> <td>121</td> </tr> <tr> <th>1,00</th> <td>7</td> <td>50</td> <td>57</td> </tr> <tr> <th>Gesamt</th> <td>122</td> <td>56</td> <td>178</td> </tr> </tbody> </table>					JV18_T1		Gesamt	,00	1,00	JV18_T2 ,00	115	6	121	1,00	7	50	57	Gesamt	122	56	178	<p align="center">JV16_T2 * JV16_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV16_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV16_T2 ,00</th> <td>164</td> <td>10</td> <td>174</td> </tr> <tr> <th>1,00</th> <td>9</td> <td>18</td> <td>27</td> </tr> <tr> <th>Gesamt</th> <td>173</td> <td>28</td> <td>201</td> </tr> </tbody> </table>					JV16_T1		Gesamt	,00	1,00	JV16_T2 ,00	164	10	174	1,00	9	18	27	Gesamt	173	28	201
	JV18_T1		Gesamt																																								
	,00	1,00																																									
JV18_T2 ,00	115	6	121																																								
1,00	7	50	57																																								
Gesamt	122	56	178																																								
	JV16_T1		Gesamt																																								
	,00	1,00																																									
JV16_T2 ,00	164	10	174																																								
1,00	9	18	27																																								
Gesamt	173	28	201																																								
<p align="center">JV14_T2 * JV14_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV14_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV14_T2 ,00</th> <td>124</td> <td>9</td> <td>133</td> </tr> <tr> <th>1,00</th> <td>2</td> <td>38</td> <td>40</td> </tr> <tr> <th>Gesamt</th> <td>126</td> <td>47</td> <td>173</td> </tr> </tbody> </table>					JV14_T1		Gesamt	,00	1,00	JV14_T2 ,00	124	9	133	1,00	2	38	40	Gesamt	126	47	173	<p align="center">JV10_T2 * JV10_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV10_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV10_T2 ,00</th> <td>155</td> <td>8</td> <td>163</td> </tr> <tr> <th>1,00</th> <td>3</td> <td>42</td> <td>45</td> </tr> <tr> <th>Gesamt</th> <td>158</td> <td>50</td> <td>208</td> </tr> </tbody> </table>					JV10_T1		Gesamt	,00	1,00	JV10_T2 ,00	155	8	163	1,00	3	42	45	Gesamt	158	50	208
	JV14_T1		Gesamt																																								
	,00	1,00																																									
JV14_T2 ,00	124	9	133																																								
1,00	2	38	40																																								
Gesamt	126	47	173																																								
	JV10_T1		Gesamt																																								
	,00	1,00																																									
JV10_T2 ,00	155	8	163																																								
1,00	3	42	45																																								
Gesamt	158	50	208																																								
<p align="center">JV8_T2 * JV8_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV8_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV8_T2 ,00</th> <td>243</td> <td>6</td> <td>249</td> </tr> <tr> <th>1,00</th> <td>3</td> <td>27</td> <td>30</td> </tr> <tr> <th>Gesamt</th> <td>246</td> <td>33</td> <td>279</td> </tr> </tbody> </table>					JV8_T1		Gesamt	,00	1,00	JV8_T2 ,00	243	6	249	1,00	3	27	30	Gesamt	246	33	279	<p align="center">JV6_T2 * JV6_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV6_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV6_T2 ,00</th> <td>275</td> <td>14</td> <td>289</td> </tr> <tr> <th>1,00</th> <td>5</td> <td>70</td> <td>75</td> </tr> <tr> <th>Gesamt</th> <td>280</td> <td>84</td> <td>364</td> </tr> </tbody> </table>					JV6_T1		Gesamt	,00	1,00	JV6_T2 ,00	275	14	289	1,00	5	70	75	Gesamt	280	84	364
	JV8_T1		Gesamt																																								
	,00	1,00																																									
JV8_T2 ,00	243	6	249																																								
1,00	3	27	30																																								
Gesamt	246	33	279																																								
	JV6_T1		Gesamt																																								
	,00	1,00																																									
JV6_T2 ,00	275	14	289																																								
1,00	5	70	75																																								
Gesamt	280	84	364																																								
<p align="center">JV4_T2 * JV4_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV4_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV4_T2 ,00</th> <td>156</td> <td>12</td> <td>168</td> </tr> <tr> <th>1,00</th> <td>10</td> <td>36</td> <td>46</td> </tr> <tr> <th>Gesamt</th> <td>166</td> <td>48</td> <td>214</td> </tr> </tbody> </table>					JV4_T1		Gesamt	,00	1,00	JV4_T2 ,00	156	12	168	1,00	10	36	46	Gesamt	166	48	214	<p align="center">JV2_T2 * JV2_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV2_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV2_T2 ,00</th> <td>260</td> <td>31</td> <td>291</td> </tr> <tr> <th>1,00</th> <td>18</td> <td>29</td> <td>47</td> </tr> <tr> <th>Gesamt</th> <td>278</td> <td>60</td> <td>338</td> </tr> </tbody> </table>					JV2_T1		Gesamt	,00	1,00	JV2_T2 ,00	260	31	291	1,00	18	29	47	Gesamt	278	60	338
	JV4_T1		Gesamt																																								
	,00	1,00																																									
JV4_T2 ,00	156	12	168																																								
1,00	10	36	46																																								
Gesamt	166	48	214																																								
	JV2_T1		Gesamt																																								
	,00	1,00																																									
JV2_T2 ,00	260	31	291																																								
1,00	18	29	47																																								
Gesamt	278	60	338																																								

VIII. Intra-Rater-Reliabilität K2_swJA: Kontingenztabellen pro Video (T1_B & T2_B)

<p align="center">PH19_T2_E * PH19_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH19_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH19_T2_E ,00</th> <td>207</td> <td>20</td> <td>227</td> </tr> <tr> <th>1,00</th> <td>6</td> <td>77</td> <td>83</td> </tr> <tr> <th>Gesamt</th> <td>213</td> <td>97</td> <td>310</td> </tr> </tbody> </table>					PH19_T1		Gesamt	,00	1,00	PH19_T2_E ,00	207	20	227	1,00	6	77	83	Gesamt	213	97	310	<p align="center">PH17_T2_E * PH17_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH17_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH17_T2_E ,00</th> <td>154</td> <td>18</td> <td>172</td> </tr> <tr> <th>1,00</th> <td>2</td> <td>72</td> <td>74</td> </tr> <tr> <th>Gesamt</th> <td>156</td> <td>90</td> <td>246</td> </tr> </tbody> </table>					PH17_T1		Gesamt	,00	1,00	PH17_T2_E ,00	154	18	172	1,00	2	72	74	Gesamt	156	90	246
	PH19_T1		Gesamt																																								
	,00	1,00																																									
PH19_T2_E ,00	207	20	227																																								
1,00	6	77	83																																								
Gesamt	213	97	310																																								
	PH17_T1		Gesamt																																								
	,00	1,00																																									
PH17_T2_E ,00	154	18	172																																								
1,00	2	72	74																																								
Gesamt	156	90	246																																								

<p>PH15_T2_E * PH15_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH15_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH15_T2_E ,00</th> <td>152</td> <td>13</td> <td>165</td> </tr> <tr> <th>1,00</th> <td>16</td> <td>101</td> <td>117</td> </tr> <tr> <th>Gesamt</th> <td>168</td> <td>114</td> <td>282</td> </tr> </tbody> </table>		PH15_T1		Gesamt	,00	1,00	PH15_T2_E ,00	152	13	165	1,00	16	101	117	Gesamt	168	114	282	<p>PH13_T2_E * PH13_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH13_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH13_T2_E ,00</th> <td>90</td> <td>10</td> <td>100</td> </tr> <tr> <th>1,00</th> <td>7</td> <td>57</td> <td>64</td> </tr> <tr> <th>Gesamt</th> <td>97</td> <td>67</td> <td>164</td> </tr> </tbody> </table>		PH13_T1		Gesamt	,00	1,00	PH13_T2_E ,00	90	10	100	1,00	7	57	64	Gesamt	97	67	164
		PH15_T1			Gesamt																																
	,00	1,00																																			
PH15_T2_E ,00	152	13	165																																		
1,00	16	101	117																																		
Gesamt	168	114	282																																		
	PH13_T1		Gesamt																																		
	,00	1,00																																			
PH13_T2_E ,00	90	10	100																																		
1,00	7	57	64																																		
Gesamt	97	67	164																																		
<p>PH11_T2_E * PH11_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH11_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH11_T2_E ,00</th> <td>202</td> <td>9</td> <td>211</td> </tr> <tr> <th>1,00</th> <td>20</td> <td>135</td> <td>155</td> </tr> <tr> <th>Gesamt</th> <td>222</td> <td>144</td> <td>366</td> </tr> </tbody> </table>		PH11_T1		Gesamt	,00	1,00	PH11_T2_E ,00	202	9	211	1,00	20	135	155	Gesamt	222	144	366	<p>PH9_T2_E * PH9_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH9_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH9_T2_E ,00</th> <td>265</td> <td>14</td> <td>279</td> </tr> <tr> <th>1,00</th> <td>12</td> <td>124</td> <td>136</td> </tr> <tr> <th>Gesamt</th> <td>277</td> <td>138</td> <td>415</td> </tr> </tbody> </table>		PH9_T1		Gesamt	,00	1,00	PH9_T2_E ,00	265	14	279	1,00	12	124	136	Gesamt	277	138	415
		PH11_T1			Gesamt																																
	,00	1,00																																			
PH11_T2_E ,00	202	9	211																																		
1,00	20	135	155																																		
Gesamt	222	144	366																																		
	PH9_T1		Gesamt																																		
	,00	1,00																																			
PH9_T2_E ,00	265	14	279																																		
1,00	12	124	136																																		
Gesamt	277	138	415																																		
<p>PH7_T2_E * PH7_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH7_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH7_T2_E ,00</th> <td>205</td> <td>10</td> <td>215</td> </tr> <tr> <th>1,00</th> <td>6</td> <td>49</td> <td>55</td> </tr> <tr> <th>Gesamt</th> <td>211</td> <td>59</td> <td>270</td> </tr> </tbody> </table>		PH7_T1		Gesamt	,00	1,00	PH7_T2_E ,00	205	10	215	1,00	6	49	55	Gesamt	211	59	270	<p>PH5_T2_E * PH5_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH5_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH5_T2_E ,00</th> <td>169</td> <td>4</td> <td>173</td> </tr> <tr> <th>1,00</th> <td>4</td> <td>23</td> <td>27</td> </tr> <tr> <th>Gesamt</th> <td>173</td> <td>27</td> <td>200</td> </tr> </tbody> </table>		PH5_T1		Gesamt	,00	1,00	PH5_T2_E ,00	169	4	173	1,00	4	23	27	Gesamt	173	27	200
		PH7_T1			Gesamt																																
	,00	1,00																																			
PH7_T2_E ,00	205	10	215																																		
1,00	6	49	55																																		
Gesamt	211	59	270																																		
	PH5_T1		Gesamt																																		
	,00	1,00																																			
PH5_T2_E ,00	169	4	173																																		
1,00	4	23	27																																		
Gesamt	173	27	200																																		
<p>PH3_T2_E * PH3_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH3_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH3_T2_E ,00</th> <td>172</td> <td>8</td> <td>180</td> </tr> <tr> <th>1,00</th> <td>6</td> <td>54</td> <td>60</td> </tr> <tr> <th>Gesamt</th> <td>178</td> <td>62</td> <td>240</td> </tr> </tbody> </table>		PH3_T1		Gesamt	,00	1,00	PH3_T2_E ,00	172	8	180	1,00	6	54	60	Gesamt	178	62	240	<p>PH1_T2_E * PH1_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH1_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH1_T2_E ,00</th> <td>163</td> <td>10</td> <td>173</td> </tr> <tr> <th>1,00</th> <td>3</td> <td>47</td> <td>50</td> </tr> <tr> <th>Gesamt</th> <td>166</td> <td>57</td> <td>223</td> </tr> </tbody> </table>		PH1_T1		Gesamt	,00	1,00	PH1_T2_E ,00	163	10	173	1,00	3	47	50	Gesamt	166	57	223
		PH3_T1			Gesamt																																
	,00	1,00																																			
PH3_T2_E ,00	172	8	180																																		
1,00	6	54	60																																		
Gesamt	178	62	240																																		
	PH1_T1		Gesamt																																		
	,00	1,00																																			
PH1_T2_E ,00	163	10	173																																		
1,00	3	47	50																																		
Gesamt	166	57	223																																		
<p>NR19_T2_E * NR19_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">NR19_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>NR19_T2_E ,00</th> <td>177</td> <td>12</td> <td>189</td> </tr> <tr> <th>1,00</th> <td>7</td> <td>48</td> <td>55</td> </tr> <tr> <th>Gesamt</th> <td>184</td> <td>60</td> <td>244</td> </tr> </tbody> </table>		NR19_T1		Gesamt	,00	1,00	NR19_T2_E ,00	177	12	189	1,00	7	48	55	Gesamt	184	60	244	<p>NR17_T2_E * NR17_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">NR17_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>NR17_T2_E ,00</th> <td>202</td> <td>11</td> <td>213</td> </tr> <tr> <th>1,00</th> <td>5</td> <td>56</td> <td>61</td> </tr> <tr> <th>Gesamt</th> <td>207</td> <td>67</td> <td>274</td> </tr> </tbody> </table>		NR17_T1		Gesamt	,00	1,00	NR17_T2_E ,00	202	11	213	1,00	5	56	61	Gesamt	207	67	274
		NR19_T1			Gesamt																																
	,00	1,00																																			
NR19_T2_E ,00	177	12	189																																		
1,00	7	48	55																																		
Gesamt	184	60	244																																		
	NR17_T1		Gesamt																																		
	,00	1,00																																			
NR17_T2_E ,00	202	11	213																																		
1,00	5	56	61																																		
Gesamt	207	67	274																																		

NR15_T2_E * NR15_T1 Kreuztabelle				NR13_T2_E * NR13_T1 Kreuztabelle					
Anzahl				Anzahl					
		NR15_T1				NR13_T1			
		,00	1,00	Gesamt			,00	1,00	Gesamt
NR15_T2_E	,00	144	7	151	NR13_T2_E	,00	228	9	237
	1,00	4	52	56		1,00	6	63	69
Gesamt		148	59	207	Gesamt		234	72	306

NR11_T2_E * NR11_T1 Kreuztabelle				
Anzahl				
		NR11_T1		
		,00	1,00	Gesamt
NR11_T2_E	,00	249	6	255
	1,00	16	82	98
Gesamt		265	88	353

IX. Intra-Rater-Reliabilität K2_ kJA: Kontingenztabelle pro Video (T1_A & T2_A)

NR9_T2 * NR9_T1 Kreuztabelle				NR7_T2 * NR7_T1 Kreuztabelle				
Anzahl				Anzahl				
		NR9_T1				NR7_T1		
		,00	Gesamt			,00	1,00	Gesamt
NR9_T2	,00	340	340	NR7_T2	,00	322	3	325
	Gesamt	340	340		1,00	0	1	1
Gesamt		340	340	Gesamt		322	4	326

NR5_T2 * NR5_T1 Kreuztabelle				NR3_T2 * NR3_T1 Kreuztabelle					
Anzahl				Anzahl					
		NR5_T1				NR3_T1			
		,00	1,00	Gesamt			,00	1,00	Gesamt
NR5_T2	,00	317	0	317	NR3_T2	,00	433	1	434
	1,00	0	1	1		Gesamt	433	1	434
Gesamt		317	1	318	Gesamt		433	1	434

NR1_T2 * NR1_T1 Kreuztabelle				JV20_T2 * JV20_T1 Kreuztabelle				
Anzahl				Anzahl				
		NR1_T1				JV20_T1		
		,00	1,00	Gesamt			,00	Gesamt
NR1_T2	,00	365	1	366	JV20_T2	,00	276	276
	1,00	0	10	10		Gesamt	276	276
Gesamt		365	11	376	Gesamt		276	276

<p>JV18_T2 * JV18_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV18_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV18_T2 ,00</th> <td>177</td> <td>1</td> <td>178</td> </tr> <tr> <th>Gesamt</th> <td>177</td> <td>1</td> <td>178</td> </tr> </tbody> </table>		JV18_T1		Gesamt	,00	1,00	JV18_T2 ,00	177	1	178	Gesamt	177	1	178	<p>JV16_T2 * JV16_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV16_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV16_T2 ,00</th> <td>196</td> <td>3</td> <td>199</td> </tr> <tr> <th>1,00</th> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <th>Gesamt</th> <td>197</td> <td>5</td> <td>202</td> </tr> </tbody> </table>		JV16_T1		Gesamt	,00	1,00	JV16_T2 ,00	196	3	199	1,00	1	2	3	Gesamt	197	5	202				
		JV18_T1			Gesamt																																
	,00	1,00																																			
JV18_T2 ,00	177	1	178																																		
Gesamt	177	1	178																																		
	JV16_T1		Gesamt																																		
	,00	1,00																																			
JV16_T2 ,00	196	3	199																																		
1,00	1	2	3																																		
Gesamt	197	5	202																																		
<p>JV14_T2 * JV14_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV14_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV14_T2 ,00</th> <td>172</td> <td>1</td> <td>173</td> </tr> <tr> <th>Gesamt</th> <td>172</td> <td>1</td> <td>173</td> </tr> </tbody> </table>		JV14_T1		Gesamt	,00	1,00	JV14_T2 ,00	172	1	173	Gesamt	172	1	173	<p>JV10_T2 * JV10_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV10_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV10_T2 ,00</th> <td>205</td> <td>0</td> <td>205</td> </tr> <tr> <th>1,00</th> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <th>Gesamt</th> <td>206</td> <td>2</td> <td>208</td> </tr> </tbody> </table>		JV10_T1		Gesamt	,00	1,00	JV10_T2 ,00	205	0	205	1,00	1	2	3	Gesamt	206	2	208				
		JV14_T1			Gesamt																																
	,00	1,00																																			
JV14_T2 ,00	172	1	173																																		
Gesamt	172	1	173																																		
	JV10_T1		Gesamt																																		
	,00	1,00																																			
JV10_T2 ,00	205	0	205																																		
1,00	1	2	3																																		
Gesamt	206	2	208																																		
<p>JV8_T2 * JV8_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV8_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV8_T2 ,00</th> <td>197</td> <td>19</td> <td>216</td> </tr> <tr> <th>1,00</th> <td>3</td> <td>60</td> <td>63</td> </tr> <tr> <th>Gesamt</th> <td>200</td> <td>79</td> <td>279</td> </tr> </tbody> </table>		JV8_T1		Gesamt	,00	1,00	JV8_T2 ,00	197	19	216	1,00	3	60	63	Gesamt	200	79	279	<p>JV6_T2 * JV6_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV6_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>Gesamt</th> </tr> </thead> <tbody> <tr> <th>JV6_T2 ,00</th> <td>364</td> <td>364</td> </tr> <tr> <th>Gesamt</th> <td>364</td> <td>364</td> </tr> </tbody> </table>		JV6_T1		Gesamt	,00	Gesamt	JV6_T2 ,00	364	364	Gesamt	364	364						
		JV8_T1			Gesamt																																
	,00	1,00																																			
JV8_T2 ,00	197	19	216																																		
1,00	3	60	63																																		
Gesamt	200	79	279																																		
	JV6_T1		Gesamt																																		
	,00	Gesamt																																			
JV6_T2 ,00	364	364																																			
Gesamt	364	364																																			
<p>JV4_T2 * JV4_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV4_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV4_T2 ,00</th> <td>211</td> <td>1</td> <td>212</td> </tr> <tr> <th>1,00</th> <td>1</td> <td>1</td> <td>2</td> </tr> <tr> <th>Gesamt</th> <td>212</td> <td>2</td> <td>214</td> </tr> </tbody> </table>		JV4_T1		Gesamt	,00	1,00	JV4_T2 ,00	211	1	212	1,00	1	1	2	Gesamt	212	2	214	<p>JV2_T2 * JV2_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">JV2_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>JV2_T2 ,00</th> <td>332</td> <td>5</td> <td>337</td> </tr> <tr> <th>1,00</th> <td>0</td> <td>1</td> <td>1</td> </tr> <tr> <th>Gesamt</th> <td>332</td> <td>6</td> <td>338</td> </tr> </tbody> </table>		JV2_T1		Gesamt	,00	1,00	JV2_T2 ,00	332	5	337	1,00	0	1	1	Gesamt	332	6	338
		JV4_T1			Gesamt																																
	,00	1,00																																			
JV4_T2 ,00	211	1	212																																		
1,00	1	1	2																																		
Gesamt	212	2	214																																		
	JV2_T1		Gesamt																																		
	,00	1,00																																			
JV2_T2 ,00	332	5	337																																		
1,00	0	1	1																																		
Gesamt	332	6	338																																		

X. Intra-Rater-Reliabilität K2_ kJA: Kontingenztabellen pro Video (T1_B & T2_B)

<p>PH19_T2_E * PH19_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH19_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH19_T2_E ,00</th> <td>298</td> <td>0</td> <td>298</td> </tr> <tr> <th>1,00</th> <td>7</td> <td>5</td> <td>12</td> </tr> <tr> <th>Gesamt</th> <td>305</td> <td>5</td> <td>310</td> </tr> </tbody> </table>		PH19_T1		Gesamt	,00	1,00	PH19_T2_E ,00	298	0	298	1,00	7	5	12	Gesamt	305	5	310	<p>PH17_T2_E * PH17_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH17_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH17_T2_E ,00</th> <td>241</td> <td>0</td> <td>241</td> </tr> <tr> <th>1,00</th> <td>1</td> <td>4</td> <td>5</td> </tr> <tr> <th>Gesamt</th> <td>242</td> <td>4</td> <td>246</td> </tr> </tbody> </table>		PH17_T1		Gesamt	,00	1,00	PH17_T2_E ,00	241	0	241	1,00	1	4	5	Gesamt	242	4	246
		PH19_T1			Gesamt																																
	,00	1,00																																			
PH19_T2_E ,00	298	0	298																																		
1,00	7	5	12																																		
Gesamt	305	5	310																																		
	PH17_T1		Gesamt																																		
	,00	1,00																																			
PH17_T2_E ,00	241	0	241																																		
1,00	1	4	5																																		
Gesamt	242	4	246																																		

<p>PH15_T2_E * PH15_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH15_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH15_T2_E ,00</th> <td>268</td> <td>0</td> <td>268</td> </tr> <tr> <th>1,00</th> <td>10</td> <td>4</td> <td>14</td> </tr> <tr> <th>Gesamt</th> <td>278</td> <td>4</td> <td>282</td> </tr> </tbody> </table>		PH15_T1		Gesamt	,00	1,00	PH15_T2_E ,00	268	0	268	1,00	10	4	14	Gesamt	278	4	282	<p>PH13_T2_E * PH13_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH13_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH13_T2_E ,00</th> <td>160</td> <td>0</td> <td>160</td> </tr> <tr> <th>1,00</th> <td>3</td> <td>1</td> <td>4</td> </tr> <tr> <th>Gesamt</th> <td>163</td> <td>1</td> <td>164</td> </tr> </tbody> </table>		PH13_T1		Gesamt	,00	1,00	PH13_T2_E ,00	160	0	160	1,00	3	1	4	Gesamt	163	1	164
		PH15_T1			Gesamt																																
	,00	1,00																																			
PH15_T2_E ,00	268	0	268																																		
1,00	10	4	14																																		
Gesamt	278	4	282																																		
	PH13_T1		Gesamt																																		
	,00	1,00																																			
PH13_T2_E ,00	160	0	160																																		
1,00	3	1	4																																		
Gesamt	163	1	164																																		
<p>PH11_T2_E * PH11_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH11_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH11_T2_E ,00</th> <td>359</td> <td>0</td> <td>359</td> </tr> <tr> <th>1,00</th> <td>2</td> <td>5</td> <td>7</td> </tr> <tr> <th>Gesamt</th> <td>361</td> <td>5</td> <td>366</td> </tr> </tbody> </table>		PH11_T1		Gesamt	,00	1,00	PH11_T2_E ,00	359	0	359	1,00	2	5	7	Gesamt	361	5	366	<p>PH9_T2_E * PH9_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH9_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH9_T2_E ,00</th> <td>412</td> <td>0</td> <td>412</td> </tr> <tr> <th>1,00</th> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <th>Gesamt</th> <td>413</td> <td>2</td> <td>415</td> </tr> </tbody> </table>		PH9_T1		Gesamt	,00	1,00	PH9_T2_E ,00	412	0	412	1,00	1	2	3	Gesamt	413	2	415
		PH11_T1			Gesamt																																
	,00	1,00																																			
PH11_T2_E ,00	359	0	359																																		
1,00	2	5	7																																		
Gesamt	361	5	366																																		
	PH9_T1		Gesamt																																		
	,00	1,00																																			
PH9_T2_E ,00	412	0	412																																		
1,00	1	2	3																																		
Gesamt	413	2	415																																		
<p>PH7_T2_E * PH7_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH7_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH7_T2_E ,00</th> <td>269</td> <td></td> <td>269</td> </tr> <tr> <th>1,00</th> <td>1</td> <td></td> <td>1</td> </tr> <tr> <th>Gesamt</th> <td>270</td> <td></td> <td>270</td> </tr> </tbody> </table>		PH7_T1		Gesamt	,00	1,00	PH7_T2_E ,00	269		269	1,00	1		1	Gesamt	270		270	<p>PH5_T2_E * PH5_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH5_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH5_T2_E ,00</th> <td>200</td> <td></td> <td>200</td> </tr> <tr> <th>Gesamt</th> <td>200</td> <td></td> <td>200</td> </tr> </tbody> </table>		PH5_T1		Gesamt	,00	1,00	PH5_T2_E ,00	200		200	Gesamt	200		200				
		PH7_T1			Gesamt																																
	,00	1,00																																			
PH7_T2_E ,00	269		269																																		
1,00	1		1																																		
Gesamt	270		270																																		
	PH5_T1		Gesamt																																		
	,00	1,00																																			
PH5_T2_E ,00	200		200																																		
Gesamt	200		200																																		
<p>PH3_T2_E * PH3_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH3_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH3_T2_E ,00</th> <td>221</td> <td>1</td> <td>222</td> </tr> <tr> <th>1,00</th> <td>12</td> <td>6</td> <td>18</td> </tr> <tr> <th>Gesamt</th> <td>233</td> <td>7</td> <td>240</td> </tr> </tbody> </table>		PH3_T1		Gesamt	,00	1,00	PH3_T2_E ,00	221	1	222	1,00	12	6	18	Gesamt	233	7	240	<p>PH1_T2_E * PH1_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">PH1_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>PH1_T2_E ,00</th> <td>222</td> <td></td> <td>222</td> </tr> <tr> <th>1,00</th> <td>1</td> <td></td> <td>1</td> </tr> <tr> <th>Gesamt</th> <td>223</td> <td></td> <td>223</td> </tr> </tbody> </table>		PH1_T1		Gesamt	,00	1,00	PH1_T2_E ,00	222		222	1,00	1		1	Gesamt	223		223
		PH3_T1			Gesamt																																
	,00	1,00																																			
PH3_T2_E ,00	221	1	222																																		
1,00	12	6	18																																		
Gesamt	233	7	240																																		
	PH1_T1		Gesamt																																		
	,00	1,00																																			
PH1_T2_E ,00	222		222																																		
1,00	1		1																																		
Gesamt	223		223																																		
<p>NR19_T2_E * NR19_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">NR19_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>NR19_T2_E ,00</th> <td>244</td> <td></td> <td>244</td> </tr> <tr> <th>Gesamt</th> <td>244</td> <td></td> <td>244</td> </tr> </tbody> </table>		NR19_T1		Gesamt	,00	1,00	NR19_T2_E ,00	244		244	Gesamt	244		244	<p>NR17_T2_E * NR17_T1 Kreuztabelle</p> <p>Anzahl</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">NR17_T1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th>,00</th> <th>1,00</th> </tr> </thead> <tbody> <tr> <th>NR17_T2_E ,00</th> <td>274</td> <td></td> <td>274</td> </tr> <tr> <th>Gesamt</th> <td>274</td> <td></td> <td>274</td> </tr> </tbody> </table>		NR17_T1		Gesamt	,00	1,00	NR17_T2_E ,00	274		274	Gesamt	274		274								
		NR19_T1			Gesamt																																
	,00	1,00																																			
NR19_T2_E ,00	244		244																																		
Gesamt	244		244																																		
	NR17_T1		Gesamt																																		
	,00	1,00																																			
NR17_T2_E ,00	274		274																																		
Gesamt	274		274																																		

NR15_T2_E * NR15_T1 Kreuztabelle				NR13_T2_E * NR13_T1 Kreuztabelle				
Anzahl				Anzahl				
		NR15_T1			NR13_T1			
		,00	1,00	Gesamt	,00	Gesamt		
NR15_T2_E	,00	180	3	183	NR13_T2_E	,00	306	306
	1,00	4	20	24	Gesamt		306	306
Gesamt		184	23	207				

NR11_T2_E * NR11_T1 Kreuztabelle				
Anzahl				
		NR11_T1		Gesamt
		,00	1,00	Gesamt
NR11_T2_E	,00	341	0	341
	1,00	6	6	12
Gesamt		347	6	353

XI. Interrater-Reliabilität K1_JA: Kontingenztabellen pro Video (T2_B & H)

PH19_T2_E * PH19 Kreuztabelle				PH17_T2_E * PH17 Kreuztabelle						
Anzahl				Anzahl						
		PH19			PH17					
		1,00	2,00	Gesamt	1,00	2,00	3,00	Gesamt		
PH19_T2_E	1,00	188	1	189	PH17_T2_E	1,00	154	2	0	156
	2,00	11	5	16		2,00	6	0	0	6
	3,00	2	0	2		3,00	0	0	1	1
Gesamt		201	6	207	Gesamt		160	2	1	163

PH15_T2_E * PH15 Kreuztabelle				PH13_T2_E * PH13 Kreuztabelle					
Anzahl				Anzahl					
		PH15			PH13				
		1,00	2,00	Gesamt	1,00	2,00	Gesamt		
PH15_T2_E	1,00	160	3	163	PH13_T2_E	1,00	89	1	90
	2,00	12	9	21		2,00	10	3	13
	3,00	3	0	3		3,00	2	0	2
Gesamt		175	12	187	Gesamt		101	4	105

PH11_T2_E * PH11 Kreuztabelle					PH9_T2_E * PH9 Kreuztabelle					
Anzahl					Anzahl					
		PH11					PH9			
		1,00	2,00	Gesamt			1,00	2,00	Gesamt	
PH11_T2_E	1,00	216	9	225	PH9_T2_E	1,00	246	3	249	
	2,00	10	5	15		2,00	21	7	28	
	3,00	3	0	3		Gesamt				
Gesamt				229	14	243				
PH7_T2_E * PH7 Kreuztabelle					PH5_T2_E * PH5 Kreuztabelle					
Anzahl					Anzahl					
		PH7					PH5			
		1,00	Gesamt				1,00	Gesamt		
PH7_T2_E	1,00	176		176	PH5_T2_E	1,00	129		129	
	2,00	4		4		2,00	4		4	
Gesamt				180	180	Gesamt				
PH3_T2_E * PH3 Kreuztabelle					PH1_T2_E * PH1 Kreuztabelle					
Anzahl					Anzahl					
		PH3					PH1			
		1,00	3,00	Gesamt			1,00	2,00	3,00	Gesamt
PH3_T2_E	1,00	144	3	147	PH1_T2_E	1,00	138	4	1	143
	2,00	1	0	1		2,00	2	2	0	4
	3,00	4	0	4		Gesamt				
Gesamt				149	3	152				
NR19_T2_E * NR19 Kreuztabelle					NR17_T2_E * NR17 Kreuztabelle					
Anzahl					Anzahl					
		NR19					NR17			
		1,00	Gesamt				1,00	2,00	Gesamt	
NR19_T2_E	1,00	158		158	NR17_T2_E	1,00	174		2	176
	2,00	5		5		2,00	3		0	3
Gesamt				163	163	Gesamt				
NR15_T2_E * NR15 Kreuztabelle					NR13_T2_E * NR13 Kreuztabelle					
Anzahl					Anzahl					
		NR15					NR13			
		1,00	2,00	Gesamt			1,00	2,00	Gesamt	
NR15_T2_E	1,00	119	1	120	NR13_T2_E	1,00	196		0	196
	2,00	7	1	8		2,00	4		2	6
	3,00	9	0	9		Gesamt				
Gesamt				135	2	137				

NR11_T2_E * NR11 Kreuztabelle				
Anzahl				
		NR11		
		1,00	2,00	Gesamt
NR11_T2_E	1,00	215	3	218
	2,00	9	1	10
Gesamt		224	4	228

XII. Interrater-Reliabilität K1_JA: Kontingenztabelle pro Video (T1_B & H)

JV2_T2 * JV2 Kreuztabelle				
Anzahl				
		JV2		
		1,00	Gesamt	
JV2_T2	1,00	218	218	
	2,00	7	7	
Gesamt		225	225	

JV4_T2 * JV4 Kreuztabelle				
Anzahl				
		JV4		
		1,00	Gesamt	
JV4_T2	1,00	128	128	
	2,00	14	14	
Gesamt		142	142	

JV6_T2 * JV6 Kreuztabelle				
Anzahl				
		JV6		
		1,00	2,00	Gesamt
JV6_T2	1,00	226	1	227
	2,00	12	3	15
Gesamt		238	4	242

JV8_T2 * JV8 Kreuztabelle					
Anzahl					
		JV8			Ge-
		1,00	2,00	3,00	samt
JV8_T2	1,00	142	1	3	146
	2,00	1	0	0	1
	3,00	37	0	0	37
Gesamt		180	1	3	184

JV10_T2 * JV10 Kreuztabelle				
Anzahl				
		JV10		
		1,00	Gesamt	
JV10_T2	1,00	129	129	
	2,00	8	8	
Gesamt		137	137	

JV14_T2 * JV14 Kreuztabelle				
Anzahl				
		JV14		
		1,00	2,00	Gesamt
JV14_T2	1,00	104	0	104
	2,00	7	2	9
Gesamt		111	2	113

JV16_T2 * JV16 Kreuztabelle				
Anzahl				
		JV16		
		1,00	3,00	Gesamt
JV16_T2	1,00	128	1	129
	2,00	2	0	2
Gesamt		130	1	131

JV18_T2 * JV18 Kreuztabelle				
Anzahl				
		JV18		
		1,00	2,00	Gesamt
JV18_T2	1,00	110	1	111
	2,00	6	1	7
Gesamt		116	2	118

JV20_T2 * JV20 Kreuztabelle				
Anzahl				
		JV20		
		1,00	2,00	Gesamt
JV20_T2	1,00	169	0	169
	2,00	8	4	12
Gesamt		177	4	181

NR1_T2 * NR1 Kreuztabelle			
Anzahl			
		NR1	
		1,00	Gesamt
NR1_T2	1,00	231	231
	2,00	16	16
	3,00	3	3
Gesamt		250	250

NR3_T2 * NR3 Kreuztabelle			
Anzahl			
		NR3	
		1,00	Gesamt
NR3_T2	1,00	289	289
	2,00	1	1
Gesamt		290	290

NR5_T2 * NR5 Kreuztabelle				
Anzahl				
		NR5		
		1,00	2,00	Gesamt
NR5_T2	1,00	177	0	177
	2,00	31	4	35
Gesamt		208	4	212

NR7_T2 * NR7 Kreuztabelle				
Anzahl				
		NR7		
		1,00	2,00	Gesamt
NR7_T2	1,00	186	1	187
	2,00	30	1	31
Gesamt		216	2	218

NR9_T2 * NR9 Kreuztabelle				
Anzahl				
		NR9		
		1,00	2,00	Gesamt
NR9_T2	1,00	214	0	214
	2,00	5	3	8
Gesamt		219	3	222

XIII. Interrater-Reliabilität K2_swJA: Kontingenztabellen pro Video (T2_B & H)

NR11_T2_E * NR11 Kreuztabelle				
Anzahl				
		NR11		
		,00	1,00	Gesamt
NR11_T2_E	,00	234	21	255
	1,00	41	57	98
Gesamt		275	78	353

NR13_T2_E * NR13 Kreuztabelle				
Anzahl				
		NR13		
		,00	1,00	Gesamt
NR13_T2_E	,00	227	11	238
	1,00	39	31	70
Gesamt		266	42	308

NR15_T2_E * NR15 Kreuztabelle				
Anzahl				
		NR15		
		,00	1,00	Gesamt
NR15_T2_E	,00	147	4	151
	1,00	23	33	56
Gesamt		170	37	207

NR17_T2_E * NR17 Kreuztabelle				
Anzahl				
		NR17		
		,00	1,00	Gesamt
NR17_T2_E	,00	200	13	213
	1,00	31	30	61
Gesamt		231	43	274

<p style="text-align: center;">NR19_T2_E * NR19 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">NR19</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">NR19_T2_E ,00</td> <td style="text-align: center;">177</td> <td style="text-align: center;">12</td> <td style="text-align: center;">189</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">26</td> <td style="text-align: center;">28</td> <td style="text-align: center;">54</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">203</td> <td style="text-align: center;">40</td> <td style="text-align: center;">243</td> </tr> </tbody> </table>		NR19		Gesamt	,00	1,00	NR19_T2_E ,00	177	12	189	1,00	26	28	54	Gesamt	203	40	243	<p style="text-align: center;">PH1_T2_E * PH1 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH1</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH1_T2_E ,00</td> <td style="text-align: center;">169</td> <td style="text-align: center;">5</td> <td style="text-align: center;">174</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">19</td> <td style="text-align: center;">31</td> <td style="text-align: center;">50</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">188</td> <td style="text-align: center;">36</td> <td style="text-align: center;">224</td> </tr> </tbody> </table>		PH1		Gesamt	,00	1,00	PH1_T2_E ,00	169	5	174	1,00	19	31	50	Gesamt	188	36	224
		NR19			Gesamt																																
	,00	1,00																																			
NR19_T2_E ,00	177	12	189																																		
1,00	26	28	54																																		
Gesamt	203	40	243																																		
	PH1		Gesamt																																		
	,00	1,00																																			
PH1_T2_E ,00	169	5	174																																		
1,00	19	31	50																																		
Gesamt	188	36	224																																		
<p style="text-align: center;">PH3_T2_E * PH3 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH3</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH3_T2_E ,00</td> <td style="text-align: center;">171</td> <td style="text-align: center;">9</td> <td style="text-align: center;">180</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">37</td> <td style="text-align: center;">23</td> <td style="text-align: center;">60</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">208</td> <td style="text-align: center;">32</td> <td style="text-align: center;">240</td> </tr> </tbody> </table>		PH3		Gesamt	,00	1,00	PH3_T2_E ,00	171	9	180	1,00	37	23	60	Gesamt	208	32	240	<p style="text-align: center;">PH5_T2_E * PH5 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH5</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH5_T2_E ,00</td> <td style="text-align: center;">169</td> <td style="text-align: center;">4</td> <td style="text-align: center;">173</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">17</td> <td style="text-align: center;">10</td> <td style="text-align: center;">27</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">186</td> <td style="text-align: center;">14</td> <td style="text-align: center;">200</td> </tr> </tbody> </table>		PH5		Gesamt	,00	1,00	PH5_T2_E ,00	169	4	173	1,00	17	10	27	Gesamt	186	14	200
		PH3			Gesamt																																
	,00	1,00																																			
PH3_T2_E ,00	171	9	180																																		
1,00	37	23	60																																		
Gesamt	208	32	240																																		
	PH5		Gesamt																																		
	,00	1,00																																			
PH5_T2_E ,00	169	4	173																																		
1,00	17	10	27																																		
Gesamt	186	14	200																																		
<p style="text-align: center;">PH7_T2_E * PH7 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH7</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH7_T2_E ,00</td> <td style="text-align: center;">207</td> <td style="text-align: center;">8</td> <td style="text-align: center;">215</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">40</td> <td style="text-align: center;">15</td> <td style="text-align: center;">55</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">247</td> <td style="text-align: center;">23</td> <td style="text-align: center;">270</td> </tr> </tbody> </table>		PH7		Gesamt	,00	1,00	PH7_T2_E ,00	207	8	215	1,00	40	15	55	Gesamt	247	23	270	<p style="text-align: center;">PH9_T2_E * PH9 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH9</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH9_T2_E ,00</td> <td style="text-align: center;">263</td> <td style="text-align: center;">16</td> <td style="text-align: center;">279</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">57</td> <td style="text-align: center;">79</td> <td style="text-align: center;">136</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">320</td> <td style="text-align: center;">95</td> <td style="text-align: center;">415</td> </tr> </tbody> </table>		PH9		Gesamt	,00	1,00	PH9_T2_E ,00	263	16	279	1,00	57	79	136	Gesamt	320	95	415
		PH7			Gesamt																																
	,00	1,00																																			
PH7_T2_E ,00	207	8	215																																		
1,00	40	15	55																																		
Gesamt	247	23	270																																		
	PH9		Gesamt																																		
	,00	1,00																																			
PH9_T2_E ,00	263	16	279																																		
1,00	57	79	136																																		
Gesamt	320	95	415																																		
<p style="text-align: center;">PH11_T2_E * PH11 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH11</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH11_T2_E ,00</td> <td style="text-align: center;">179</td> <td style="text-align: center;">32</td> <td style="text-align: center;">211</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">57</td> <td style="text-align: center;">98</td> <td style="text-align: center;">155</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">236</td> <td style="text-align: center;">130</td> <td style="text-align: center;">366</td> </tr> </tbody> </table>		PH11		Gesamt	,00	1,00	PH11_T2_E ,00	179	32	211	1,00	57	98	155	Gesamt	236	130	366	<p style="text-align: center;">PH13_T2_E * PH13 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH13</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH13_T2_E ,00</td> <td style="text-align: center;">91</td> <td style="text-align: center;">10</td> <td style="text-align: center;">101</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">31</td> <td style="text-align: center;">33</td> <td style="text-align: center;">64</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">122</td> <td style="text-align: center;">43</td> <td style="text-align: center;">165</td> </tr> </tbody> </table>		PH13		Gesamt	,00	1,00	PH13_T2_E ,00	91	10	101	1,00	31	33	64	Gesamt	122	43	165
		PH11			Gesamt																																
	,00	1,00																																			
PH11_T2_E ,00	179	32	211																																		
1,00	57	98	155																																		
Gesamt	236	130	366																																		
	PH13		Gesamt																																		
	,00	1,00																																			
PH13_T2_E ,00	91	10	101																																		
1,00	31	33	64																																		
Gesamt	122	43	165																																		
<p style="text-align: center;">PH15_T2_E * PH15 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH15</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH15_T2_E ,00</td> <td style="text-align: center;">149</td> <td style="text-align: center;">16</td> <td style="text-align: center;">165</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">35</td> <td style="text-align: center;">83</td> <td style="text-align: center;">118</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">184</td> <td style="text-align: center;">99</td> <td style="text-align: center;">283</td> </tr> </tbody> </table>		PH15		Gesamt	,00	1,00	PH15_T2_E ,00	149	16	165	1,00	35	83	118	Gesamt	184	99	283	<p style="text-align: center;">PH17_T2_E * PH17 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">PH17</th> <th rowspan="2">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">PH17_T2_E ,00</td> <td style="text-align: center;">150</td> <td style="text-align: center;">22</td> <td style="text-align: center;">172</td> </tr> <tr> <td style="text-align: center;">1,00</td> <td style="text-align: center;">30</td> <td style="text-align: center;">44</td> <td style="text-align: center;">74</td> </tr> <tr> <td style="text-align: center;">Gesamt</td> <td style="text-align: center;">180</td> <td style="text-align: center;">66</td> <td style="text-align: center;">246</td> </tr> </tbody> </table>		PH17		Gesamt	,00	1,00	PH17_T2_E ,00	150	22	172	1,00	30	44	74	Gesamt	180	66	246
		PH15			Gesamt																																
	,00	1,00																																			
PH15_T2_E ,00	149	16	165																																		
1,00	35	83	118																																		
Gesamt	184	99	283																																		
	PH17		Gesamt																																		
	,00	1,00																																			
PH17_T2_E ,00	150	22	172																																		
1,00	30	44	74																																		
Gesamt	180	66	246																																		

PH19_T2_E * PH19 Kreuztabelle				
Anzahl		PH19		Gesamt
		,00	1,00	
PH19_T2_E	,00	220	7	227
	1,00	38	45	83
Gesamt		258	52	310

XIV. Interrater-Reliabilität K2_swJA: Kontingenztabelle pro Video (T1_B & H)

JV2_T2 * JV2 Kreuztabelle				JV4_T2 * JV4 Kreuztabelle					
Anzahl		JV2		Gesamt	Anzahl		JV4		Gesamt
		,00	1,00			,00	1,00		
JV2_T2	,00	274	17	291	JV4_T2	,00	141	27	168
	1,00	25	22	47		1,00	22	24	46
Gesamt		299	39	338	Gesamt		163	51	214
JV6_T2 * JV6 Kreuztabelle				JV8_T2 * JV8 Kreuztabelle					
Anzahl		JV6		Gesamt	Anzahl		JV8		Gesamt
		,00	1,00			,00	1,00		
JV6_T2	,00	272	18	290	JV8_T2	,00	240	10	250
	1,00	23	52	75		1,00	15	16	31
Gesamt		295	70	365	Gesamt		255	26	281
JV10_T2 * JV10 Kreuztabelle				JV14_T2 * JV14 Kreuztabelle					
Anzahl		JV10		Gesamt	Anzahl		JV14		Gesamt
		,00	1,00			,00	1,00		
JV10_T2	,00	153	10	163	JV14_T2	,00	125	8	133
	1,00	20	25	45		1,00	11	29	40
Gesamt		173	35	208	Gesamt		136	37	173
JV16_T2 * JV16 Kreuztabelle				JV18_T2 * JV18 Kreuztabelle					
Anzahl		JV16		Gesamt	Anzahl		JV18		Gesamt
		,00	1,00			2,00	,00	1,00	
JV16_T2	,00	170	4	174	JV18_T2	,00	113	8	121
	1,00	16	8	25		1,00	25	32	57
Gesamt		186	12	199	Gesamt		138	40	178

JV20_T2 * JV20 Kreuztabelle				NR1_T2 * NR1 Kreuztabelle					
Anzahl				Anzahl					
		JV20			NR1		Ge-		
		,00	1,00	Gesamt		,00	1,00	Ge-	
								gesamt	
JV20_T2	,00	217	16	233	NR1_T2	,00	288	17	305
	1,00	14	29	43		1,00	38	31	69
Gesamt		231	45	276	Gesamt		326	48	374
NR3_T2 * NR3 Kreuztabelle				NR5_T2 * NR5 Kreuztabelle					
Anzahl				Anzahl					
		NR3			NR5				
		,00	1,00	Gesamt		,00	1,00	Gesamt	
NR3_T2	,00	389	6	395	NR5_T2	,00	189	7	196
	1,00	31	8	39		1,00	68	54	122
Gesamt		420	14	434	Gesamt		257	61	318
NR7_T2 * NR7 Kreuztabelle				NR9_T2 * NR9 Kreuztabelle					
Anzahl				Anzahl					
		NR7			NR9				
		,00	1,00	Gesamt		,00	1,00	Gesamt	
NR7_T2	,00	185	19	204	NR9_T2	,00	260	21	281
	1,00	60	63	123		1,00	33	26	59
Gesamt		245	82	327	Gesamt		293	47	340

XV. Interrater-Reliabilität K2_kJA: Kontingenztabelle pro Video (T2_B & H)

NR11_T1 * NR11 Kreuztabelle				NR13_T1 * NR13 Kreuztabelle				
Anzahl				Anzahl				
		NR11			NR13			
		,00	1,00	Gesamt	,00	Gesamt		
NR11_T1	,00	346	1	347	NR13_T1	,00	306	306
	1,00	6	0	6	Gesamt		306	306
Gesamt		352	1	353				
NR15_T1 * NR15 Kreuztabelle				NR17_T1 * NR17 Kreuztabelle				
Anzahl				Anzahl				
		NR15			NR17			
		,00	Gesamt		,00	1,00	Gesamt	
NR15_T1	,00	184	184	NR17_T1	,00	275	1	276
	1,00	23	23	Gesamt		275	1	276
Gesamt		207	207					

NR19_T1 * NR19 Kreuztabelle				PH1_T1 * PH1 Kreuztabelle					
Anzahl				Anzahl					
		NR19				PH1			
		,00	1,00	Gesamt		,00	Gesamt		
NR19_T1	,00	240	3	243	PH1_T1	,00	223	223	
	1,00					1,00			
Gesamt		240	3	243	Gesamt		223	223	
PH7_T1 * PH7 Kreuztabelle				PH9_T1 * PH9 Kreuztabelle					
Anzahl				Anzahl					
		PH7				PH9			
		,00	1,00	Gesamt		,00	1,00	Gesamt	
PH7_T1	,00	269	1	270	PH9_T1	,00	411	2	413
	1,00					1,00	2	0	2
Gesamt		269	1	270	Gesamt		413	2	415
PH11_T1 * PH11 Kreuztabelle				PH13_T1 * PH13 Kreuztabelle					
Anzahl				Anzahl					
		PH11				PH13			
		,00	1,00	Gesamt		,00	Gesamt		
PH11_T1	,00	360	1	361	PH13_T1	,00	163	163	
	1,00	5	0	5		1,00	1	1	
Gesamt		365	1	366	Gesamt		164	164	
PH15_T1 * PH15 Kreuztabelle				PH17_T1 * PH17 Kreuztabelle					
Anzahl				Anzahl					
		PH15				PH17			
		,00	1,00	Gesamt		,00	1,00	Gesamt	
PH15_T1	,00	280	0	280	PH17_T1	,00	239	3	242
	1,00	2	2	4		1,00	1	3	4
Gesamt		282	2	284	Gesamt		240	6	246
PH19_T1 * PH19 Kreuztabelle									
Anzahl									
		PH19							
		,00	1,00	Gesamt					
PH19_T1	,00	305	0	305					
	1,00	4	1	5					
Gesamt		309	1	310					

XVI. Interrater-Reliabilität K2_kJA: Kontingenztabelle pro Video (T1_B & H)

<p style="text-align: center;">JV4_T2 * JV4 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV4</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV4_T2 ,00</th> <td style="text-align: center;">211</td> <td style="text-align: center;">1</td> <td style="text-align: center;">212</td> </tr> <tr> <th style="text-align: center;">1,00</th> <td style="text-align: center;">2</td> <td style="text-align: center;">0</td> <td style="text-align: center;">2</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">213</td> <td style="text-align: center;">1</td> <td style="text-align: center;">214</td> </tr> </tbody> </table>		JV4		Gesamt	,00	1,00	JV4_T2 ,00	211	1	212	1,00	2	0	2	Gesamt	213	1	214	<p style="text-align: center;">JV6_T2 * JV6 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV6</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV6_T2 ,00</th> <td style="text-align: center;">364</td> <td style="text-align: center;">1</td> <td style="text-align: center;">365</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">364</td> <td style="text-align: center;">1</td> <td style="text-align: center;">365</td> </tr> </tbody> </table>		JV6		Gesamt	,00	1,00	JV6_T2 ,00	364	1	365	Gesamt	364	1	365				
		JV4			Gesamt																																
	,00	1,00																																			
JV4_T2 ,00	211	1	212																																		
1,00	2	0	2																																		
Gesamt	213	1	214																																		
	JV6		Gesamt																																		
	,00	1,00																																			
JV6_T2 ,00	364	1	365																																		
Gesamt	364	1	365																																		
<p style="text-align: center;">JV8_T2 * JV8 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV8</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV8_T2 ,00</th> <td style="text-align: center;">212</td> <td style="text-align: center;">5</td> <td style="text-align: center;">217</td> </tr> <tr> <th style="text-align: center;">1,00</th> <td style="text-align: center;">64</td> <td style="text-align: center;">0</td> <td style="text-align: center;">64</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">276</td> <td style="text-align: center;">5</td> <td style="text-align: center;">281</td> </tr> </tbody> </table>		JV8		Gesamt	,00	1,00	JV8_T2 ,00	212	5	217	1,00	64	0	64	Gesamt	276	5	281	<p style="text-align: center;">JV10_T2 * JV10 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV10</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV10_T2 ,00</th> <td style="text-align: center;">204</td> <td style="text-align: center;">1</td> <td style="text-align: center;">205</td> </tr> <tr> <th style="text-align: center;">1,00</th> <td style="text-align: center;">3</td> <td style="text-align: center;">0</td> <td style="text-align: center;">3</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">207</td> <td style="text-align: center;">1</td> <td style="text-align: center;">208</td> </tr> </tbody> </table>		JV10		Gesamt	,00	1,00	JV10_T2 ,00	204	1	205	1,00	3	0	3	Gesamt	207	1	208
		JV8			Gesamt																																
	,00	1,00																																			
JV8_T2 ,00	212	5	217																																		
1,00	64	0	64																																		
Gesamt	276	5	281																																		
	JV10		Gesamt																																		
	,00	1,00																																			
JV10_T2 ,00	204	1	205																																		
1,00	3	0	3																																		
Gesamt	207	1	208																																		
<p style="text-align: center;">JV14_T2 * JV14 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV14</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV14_T2 ,00</th> <td style="text-align: center;">173</td> <td style="text-align: center;">0</td> <td style="text-align: center;">173</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">173</td> <td style="text-align: center;">0</td> <td style="text-align: center;">173</td> </tr> </tbody> </table>		JV14		Gesamt	,00	1,00	JV14_T2 ,00	173	0	173	Gesamt	173	0	173	<p style="text-align: center;">JV16_T2 * JV16 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV16</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV16_T2 ,00</th> <td style="text-align: center;">196</td> <td style="text-align: center;">3</td> <td style="text-align: center;">199</td> </tr> <tr> <th style="text-align: center;">1,00</th> <td style="text-align: center;">3</td> <td style="text-align: center;">0</td> <td style="text-align: center;">3</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">199</td> <td style="text-align: center;">3</td> <td style="text-align: center;">202</td> </tr> </tbody> </table>		JV16		Gesamt	,00	1,00	JV16_T2 ,00	196	3	199	1,00	3	0	3	Gesamt	199	3	202				
		JV14			Gesamt																																
	,00	1,00																																			
JV14_T2 ,00	173	0	173																																		
Gesamt	173	0	173																																		
	JV16		Gesamt																																		
	,00	1,00																																			
JV16_T2 ,00	196	3	199																																		
1,00	3	0	3																																		
Gesamt	199	3	202																																		
<p style="text-align: center;">JV18_T2 * JV18 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV18</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV18_T2 ,00</th> <td style="text-align: center;">177</td> <td style="text-align: center;">1</td> <td style="text-align: center;">178</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">177</td> <td style="text-align: center;">1</td> <td style="text-align: center;">178</td> </tr> </tbody> </table>		JV18		Gesamt	,00	1,00	JV18_T2 ,00	177	1	178	Gesamt	177	1	178	<p style="text-align: center;">JV20_T2 * JV20 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">JV20</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">JV20_T2 ,00</th> <td style="text-align: center;">276</td> <td style="text-align: center;">0</td> <td style="text-align: center;">276</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">276</td> <td style="text-align: center;">0</td> <td style="text-align: center;">276</td> </tr> </tbody> </table>		JV20		Gesamt	,00	1,00	JV20_T2 ,00	276	0	276	Gesamt	276	0	276								
		JV18			Gesamt																																
	,00	1,00																																			
JV18_T2 ,00	177	1	178																																		
Gesamt	177	1	178																																		
	JV20		Gesamt																																		
	,00	1,00																																			
JV20_T2 ,00	276	0	276																																		
Gesamt	276	0	276																																		
<p style="text-align: center;">NR3_T2 * NR3 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">NR3</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">NR3_T2 ,00</th> <td style="text-align: center;">433</td> <td style="text-align: center;">0</td> <td style="text-align: center;">433</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">433</td> <td style="text-align: center;">0</td> <td style="text-align: center;">433</td> </tr> </tbody> </table>		NR3		Gesamt	,00	1,00	NR3_T2 ,00	433	0	433	Gesamt	433	0	433	<p style="text-align: center;">NR5_T2 * NR5 Kreuztabelle</p> <p>Anzahl</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2" style="text-align: center;">NR5</th> <th rowspan="2" style="text-align: center;">Gesamt</th> </tr> <tr> <th style="text-align: center;">,00</th> <th style="text-align: center;">1,00</th> </tr> </thead> <tbody> <tr> <th style="text-align: center;">NR5_T2 ,00</th> <td style="text-align: center;">316</td> <td style="text-align: center;">1</td> <td style="text-align: center;">317</td> </tr> <tr> <th style="text-align: center;">1,00</th> <td style="text-align: center;">0</td> <td style="text-align: center;">1</td> <td style="text-align: center;">1</td> </tr> <tr> <th style="text-align: center;">Gesamt</th> <td style="text-align: center;">316</td> <td style="text-align: center;">2</td> <td style="text-align: center;">318</td> </tr> </tbody> </table>		NR5		Gesamt	,00	1,00	NR5_T2 ,00	316	1	317	1,00	0	1	1	Gesamt	316	2	318				
		NR3			Gesamt																																
	,00	1,00																																			
NR3_T2 ,00	433	0	433																																		
Gesamt	433	0	433																																		
	NR5		Gesamt																																		
	,00	1,00																																			
NR5_T2 ,00	316	1	317																																		
1,00	0	1	1																																		
Gesamt	316	2	318																																		

NR7_T2 * NR7 Kreuztabelle					NR9_T2 * NR9 Kreuztabelle				
Anzahl					Anzahl				
		NR7					NR9		
		,00	1,00	Gesamt			,00	1,00	Gesamt
NR7_T2	,00	325	1	326	NR9_T2	,00	336	4	340
	1,00	1	0	1		Gesamt	336	4	340
Gesamt		326	1	327					

XVII. Beobachtungsbogen Joint Attention (erste Seite)

BeobachterIn: _____

Videonr: _____

Datum: _____

Protokollbogen

Joint Attention	00:00:00-00:00:59				00:01:00-00:01:59				00:02:00-00:02:59				00:03:00-00:03:59				00:04:00-00:04:59				00:05:00-00:05:59			
	0-14	15-29	30-44	45-59	0-14	15-29	30-44	45-59	0-14	15-29	30-44	45-59	0-14	15-29	30-44	45-59	0-14	15-29	30-44	45-59	0-14	15-29	30-44	45-59
JA Following																								
JA Switching																								
Keine JA																								
unkodierbar – JA																								

Joint Attention	00:00:00-00:00:59					00:01:00-00:01:59					00:02:00-00:02:59					00:03:00-00:03:59					00:04:00-00:04:59					00:05:00-00:05:59									
	0	10	20	30	40	50	0	10	20	30	40	50	0	10	20	30	40	50	0	10	20	30	40	50	0	10	20	30	40	50	0	10	20	30	40
JA Switching																																			
Keine JA																																			