



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„INTAKT: Vergleich der Kategorien Feinfühligkeit und Rückmeldung in Bezug auf die händische und computer-gestützte Kodierung“

verfasst von / submitted by

Pavlata Georg Bsc

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

Wien, 2019 / Vienna 2019

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

A 066 840

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

Masterstudium Psychologie UG2002

Betreut von / Supervisor:

Ass.-Prof. Dr. Ursula Kastner-Koller

Danksagung

An dieser Stelle möchte ich mich bei all jenen Menschen bedanken, die es mir ermöglicht haben, diese Masterarbeit zu verfassen.

Mein wissenschaftlicher Dank geht an Frau Ass.-Prof. Dr. Kastner Koller und Frau Ass.-Prof. Dr. Deimann für die Bereitstellung des Masterarbeitsthemas sowie die persönliche und geduldige Betreuung.

Ein herzliches Dankeschön richtet sich auch an Mag. Dr. Hirschmann für die Einschulung in das Beobachtungssystem INTAKT.

Mein persönlich Dank geht aus tiefstem Herzen an meine Familie. Meine Eltern hatten mit ihrer unermüdlichen Informationslust zum Status dieser Arbeit einen großen Anteil an meiner kontinuierlichen Tätigkeit daran, meine Frau Malou und meine Tochter Emma waren in letzter Konsequenz der Treibstoff für die Vollen-
dung. Ohne eure Unterstützung wäre ich nicht so weit gekommen.

Danke!

Inhaltsverzeichnis

Einleitung	1
Theoretischer Teil	3
1 Eltern-Kind-Bindung	3
2 Eltern-Kind-Interaktion	6
2.1 Feinfühligkeit	7
2.2 Rückmeldung	8
3 Wissenschaftliche Verhaltensbeobachtung	12
3.1 Formen der Verhaltensbeobachtung	13
3.2 Quantifizierung in der Verhaltensbeobachtung	13
3.3 Beobachtungssysteme: Kodierung	14
3.4 Verhaltensbeobachtung mittels Videotechnik	16
3.5 Beobachtungsfehler in der wissenschaftlichen Verhaltensbeobachtung	18
3.5.1 Fehler zu Lasten des/der BeobachterIn	18
3.5.2 Fehler zu Lasten der Beobachtung	20
3.5.3 Fehler zu Lasten äußerer Bedingungen	20
4 Gütekriterien der Verhaltensbeobachtung	21
4.1 Objektivität	21
4.2 Validität	21
4.3 Reliabilität	22
4.4 Ökonomie	23
5 Papier-Bleistift oder computergestützte Kodierung	24

Empirischer Teil	26
6 Hintergrund der Untersuchung	26
6.1 Forschungsfrage	27
7 Untersuchungsinstrument	28
7.1 Beobachtungssystem INTAKT	28
7.1.1 Beobachtungssituation	28
7.1.2 Beobachtungskategorien	28
7.1.2.1 Feinfühligkeit	29
7.1.2.2 Rückmeldung	30
7.1.2.3 Joint-Attention	32
7.1.3 INTERACT	33
8 Untersuchungsdesign	34
8.1 INTAKT: Einschulung und Probekodierung	34
8.2 Stichprobe	34
8.3 Händische Kodierung: Adaption und Ablauf	35
8.3.1 Skala Feinfühligkeit	35
8.3.2 Skala Rückmeldung	35
8.3.3 Ablauf	37
8.4 Retest-Reliabilität	37
9 Koeffizienten zur Überprüfung der BeobachterInnenübereinstimmung	40
9.1 Intraklassenkorrelation	40
9.2 Kappa-Koeffizient nach Cohen (κ)	43
10 Ergebnisse der Untersuchung	45
10.1 Feinfühligkeit: Vergleich händischer mit computerbasierter Kodierung	45
10.1.1 Korrelationskoeffizienten (r_{PM} , r_{IC} , κ_w) Feinfühligkeit . . .	48

10.2 Rückmeldung: Vergleich händischer mit computerbasierter Ko-	
dierung	51
10.2.1 Korrelationskoeffizienten (r_{PM} , κ_w) Rückmeldung	55
10.2.2 Kappa-Koeffizient nach Cohen (κ) pro Video	56
11 Diskussion	59
11.1 Feinfühligkeit	60
11.1.1 Übereinstimmungsvalidität	60
11.1.2 BeobachterInnenübereinstimmung	60
11.2 Rückmeldung	61
11.2.1 Übereinstimmungsvalidität	61
11.2.2 BeobachterInnenübereinstimmung	62
11.3 Anmerkungen und Verbesserungsvorschläge	64
Abstract (Deutsch)	66
Abstract (Englisch)	67
Literaturverzeichnis	68
Abbildungsverzeichnis	75
Tabellenverzeichnis	76
Anhang	77

Einleitung

Der Einsatz der Interaktionsbeobachtung wird vorrangig durch ein Anwendungsfach der Psychologie bestimmt: der pädagogischen Psychologie. Hier ist sie überwiegend in zwei Gebieten häufiger Gebrauchsgegenstand, in der Entwicklungsdiagnostik und der Familiendiagnostik. In der Familiendiagnostik wird sie bei der Erfassung von Paarinteraktionen und Familienbeziehungen mit dem Ziel der Bereitstellung von Informationen für den Interventionsprozess oder der Operationalisierung von theoretischen Konstrukten verwendet.

In der Entwicklungsdiagnostik steht der Informationsgewinn für den entwicklungsdiagnostischen Prozess im Vordergrund. Vor allem die Beobachtung in der natürlichen Situation bietet Vorteile wie höherer ökologischer Validität gegenüber der standardisierten Testsituation (Deimann & Kastner-Koller, 2007) und das Problem der Diagnostik in den ersten Lebensjahren kann umgangen werden. Die Komplexität der Zusammenhänge der unterschiedlichen Beurteilungskriterien wie der Responsivität, Anregung, Lenkung aber auch die individuellen Aspekte der elterlichen und kindlichen Verhaltensmerkmale bzw. die damit einhergehenden Problematik bei der Beobachtung haben zu einer großen Anzahl an unterschiedlichen Beobachtungssystemen mit unterschiedlichen Fokussierungen geführt. In Jakob (2016) findet sich eine Auflistung verschiedenster Interaktionsbeobachtungsverfahren. Auch wenn die angeführten Verfahren keinen vollständigen Blick über dieses Segment geben, so fällt doch auf, dass bei einer großen Zahl die Gütekriterien wenig ausgearbeitet sind. Dieser Mangel an standardisierten Verfahren (Deimann & Kastner-Koller, 2007; Sarimski, 2006) ist natürlich ein Nachteil, ebenso dass viele Verfahren unökonomisch sind (Schneewind, 2006).

Genau dieser Thematik wird versucht sich mittels des INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) anzunehmen, handelt es sich hierbei doch um ein standardisiertes Video-Beobachtungssystem zur Erfassung und Beurteilung der Interaktionsqualität zwischen der primären Bezugsperson und dem Kind. Eingeschätzt wird das Verhalten der Bezugsperson anhand von drei Beobachtungskategorien: *Feinfühligkeit*, *Rückmeldung* und *Joint Attention*. Die

Beobachtungssituation teilt sich in zwei Szenarien, bestehend aus einer Bastelaufgabe und einer Spielsituation. In Bezug auf die Gütekriterien Reliabilität, Objektivität und Validität liefert das INTAKT-Beobachtungssystem gute Ergebnisse (Hirschmann, Kastner-Koller, Deimann, Aigner & Svecz, 2011), momentan wird zur Kodierung der drei Verhaltenskategorien jedoch die Hilfe einer Software benötigt. Im Sinne der Ökonomie entwickelte Zouzoula (2015) eine Kodiermöglichkeit im Form der Papier-Bleistift-Methode für die Kategorien *Feinfühligkeit* und *Rückmeldung*. Die händische Kodiermöglichkeit bildet auch den zentralen Aspekt dieser Arbeit, denn: Die Ergebnisse des INTAKT-Verfahrens sind laut Hirschmann et al. (2011) in allen drei Hauptgütekriterien gut. Jedoch muss bei einer neuen Kodierform wie der händischen Kodierung nach Zouzoula (2015) ein Vergleich mit der computergestützten Kodierung erfolgen, um sicher zu gehen, dass keine Abweichungen entstehen. Diese Gegenüberstellung soll in dieser Arbeit unter dem Aspekt der Kriteriumsvalidität und der BeobachterInnenübereinstimmung erfolgen.

Da die Bindung zwischen Eltern und Kind unmittelbar mit der Interaktion der selbigen verwoben ist, werden im ersten Kapitel kurze Definitionen zu den für den Autor wichtigsten Begrifflichkeiten gegeben.

Theoretischer Teil

1 Eltern-Kind-Bindung

Eine Definition zum Wort Bindung, wie sie in der Bindungsforschung gesehen wird, findet sich bei Zimmermann (2007, S. 326): „Unter Bindung versteht man eine lang andauernde, emotionale Beziehung zu vertrauten Personen, die Schutz und Unterstützung bieten ...“ Diese Konkretisierung ist wichtig, um eine Abgrenzung zum Wort Beziehung zu schaffen. Beziehung wird durch starke emotionale Reaktionen definiert wohingegen Bindung als ein theoretisches Konstrukt verstanden wird.

Bindungen sind überlebensnotwendig. Soziale Bindungen aufzubauen und aufrecht zu erhalten lässt uns Sicherheit erfahren. Laut Grossmann und Grossmann (2004) hat eine kontinuierliche und feinfühliges Fürsorge für die seelische Gesundheit eines sich noch zu entwickelnden Kindes eine essentielle Bedeutung. Ebenso muss auf die biologische Funktionalität von Bindung in Form von Sicherheit und Stressregulation hingewiesen werden.

Bindung entwickelt sich und kann nicht als etwas Angeborenes betrachtet werden. Jedoch kommen Kinder mit einem gewissen Verhaltensrepertoire auf die Welt. Weinen und Lachen, Rufen und Klammern helfen bei der aktiven Suche nach einer Bindungsperson. Nicht nur diese Verhaltensweisen sind angeboren, auch der Protest beim Verlassen fällt in diese Kategorie. Diese Bindungsverhaltensweisen gewährleisten die Entstehung einer gewissen Nähe zur Bindungsperson welche dann wiederum Fürsorge weckt und im weiteren Sinne Schutz (in Form der Bindungsperson) für das Kind bietet. Das Bindungsverhalten wird nur beim Wunsch nach Nähe oder in belastenden Situationen gezeigt, wohingegen die Bindung etwas Kontinuierliches darstellt.

Die aktiven Handlungen sind ausschlaggebend für die Verhaltensbeobachtung, erlauben diese Bindungsverhaltensweisen es doch erst, Bindung „sichtbar“ zu machen. Unter Bindungsverhalten versteht man im Allgemeinen jede Art von ziel-

orientierten Verhalten, welches Nähe zur Bindungsperson schafft.

Ainsworth, Blehar, Waters und Wall (1978) konnte aufzeigen, dass verschiedene Arten von Bindungsverhaltensweisen existieren. Durch die „Fremde Situation“ (Ainsworth et al., 1978) konnte sie das Bindungs- und Explorationsverhalten beim Kind aktivieren und folgende Bindungsklassifikationen herausarbeiten: sicher gebundene Kinder, unsicher-vermeidende Kinder sowie unsicher-ambivalente Kinder. Durch Main und Solomon (1986) konnte eine weitere Kategorie definiert werden, die desorganisierte Bindung.

Sicher-gebundene Kinder haben gelernt, dass ihre Bindungsperson stets verfügbar ist. Dadurch können sie die Anwesenheit der Mutter als sichere Basis nutzen und von ihr aus die Umgebung explorieren. Unsicher-vermeidende Kinder hingegen reagieren scheinbar unbeeindruckt, wenn sie allein gelassen werden und reagieren kaum auf die Anwesenheit der Bezugsperson. Es kann auch vorkommen, dass sie Fremde gegenüber der Mutter vorziehen. Bei unsicher-ambivalenten Kindern steht die Ängstlichkeit und Abhängigkeit von der Bezugsperson im Vordergrund. Sie befinden sich im Spannungsverhältnis von Nähe und Ärger auf die Mutter. Unter die vierte Kategorie (desorganisierte Kinder) fallen alle Kinder, die sich in keine der bereits genannten Bindungsklassifikationen einordnen lassen. Hier konnten bizarre, stereotype Verhaltensweisen wie Unterbrechen einer Bewegung, Erstarren, rhythmisch-monotone Bewegungen oder völlige Emotionslosigkeit identifiziert werden.

Die erforschten Implikationen sind, vor allem für das sichere Bindungsmuster, für die Entwicklung des Kindes weitreichend. Sie erstrecken sich von der modellierenden Rolle der sicheren Bindung für zukünftiges antisoziales Verhalten im frühen Schulalter (Kim, Kochanska, Boldt, Nordling & O’Bleness, 2014) im Spezifischen, bis hin zum Zusammenhang mit der positiven Entwicklung des Kindes in den Sparten Emotionsregulation, Selbständigkeit und sozialer Kompetenzen (Sroufe, 2005) in einem weiteren Sinne. Doch welche Faktoren bedingen ein sicheres Bindungsmuster? Laut Grossmann und Grossmann (2004) existieren Zusammenhänge bei den Verhaltensweisen der Mutter mit der Bindungsqualität des

Kindes. Je feinfühlicher auf die Bindungs- sowie Explorationswünsche des Kindes durch die Mutter reagiert wird, desto angemessener und ausgewogener sind die Reaktionen in der „Fremden Situation“. Neu ist das natürlich nicht, postuliert dies schon Ainsworth et al. (1978). Genau diese Verhaltensweisen in der Interaktionen zwischen Bezugsperson und Kind, die Einfluss auf die Bindungsqualität haben, werden im nächsten Abschnitt behandelt.

2 Eltern-Kind-Interaktion

Im vorherigen Kapitel wurden die grundlegenden Aspekte des kindlichen Bindungsverhaltens kurz dargestellt. Dieses theoretische Konstrukt der Bindung kann nur durch eine Beziehung zwischen Eltern und Kind entstehen. Doch auch Beziehung in diesem Kontext kann, wie Dunitz-Scheer, Scheer, Stadler und Kaimbacher (2011) erklären, nur aus Interaktion und Kommunikation hervorgehen. In diesem Kapitel soll eine Definition der eben genannten Begriffe Kommunikation, Interaktion und Beziehung erfolgen und weiters auf den Zusammenhang der Interaktionsverhaltensweisen zu der kindlichen Entwicklung eingegangen werden.

Dunitz-Scheer et al. (2011, S. 965) schreiben, dass „die Voraussetzung für Beziehung eine zeitlich gewachsene Sequenz von Interaktionen zwischen stabilen Interaktionspartnern“ ist. Zuerst entsteht aber durch verbale und nonverbale Signale Kommunikation. Aus der Kommunikation heraus wird durch Verhaltensweisen sowie deren Wirkung auf das Gegenüber eine Interaktion gestaltet. Interaktion ist also das Wechselspiel von Verhaltensweisen zwischen mindestens zwei Individuen. Ein Zweig der Diagnostik beschäftigt sich mit genau diesem Bereich, die Interaktionsdiagnostik. Es soll das sichtbare Verhalten zweier Menschen zueinander erkannt, verstanden, wissenschaftlich nachvollziehbar benannt und evaluiert sowie kategorisiert und klassifiziert werden (vgl. Dunitz-Scheer et al., 2011, S. 968).

Hier kommen wir wieder auf den zum ersten Mal von Ainsworth et al. (1978) postulierten Zusammenhang von zwischen Bindungssicherheit und der Qualität des mütterlichen Interaktionsverhaltens zurück. Sie konnte in ihrer Studie zeigen, dass Kinder von Müttern, die im Alltag eher feinfühlig waren, in der „Fremden Situation“ häufiger sicheres Verhalten zeigten als Kinder von Müttern, die sich im Alltag weniger feinfühlig zeigten. Die Qualität des mütterlichen Verhaltens wurde durch die mütterliche Feinfühligkeit definiert. Dies ist aber nicht das einzige Interaktionsverhalten von Bezugspersonen mit Auswirkung auf die Bindungssicherheit. De Wolff und van Ijzendoorn (1997) identifizierten weitere Aspekte wie:

- gegenseitige, wechselseitige und fruchtbare Interaktion
- positiver, zielgerichteter Austausch von Mutter und Kind, die Fähigkeit der Mutter die Aufmerksamkeit des Kindes zu lenken, zu halten und ihre Empfänglichkeit für kindliche Zeichen
- mütterliche positive und negative Aufmerksamkeit in Richtung des Kleinkindes sowie der Grad wechselseitiger Interaktion
- jegliche Art von Aktion der Mutter in Richtung des Kleinkindes
- emotionale Unterstützung in Form der „sicheren Basis“

In den Grundzügen sind hier bereits die drei Hauptkategorien des Beobachtungssystems INTAKT (Hirschmann et al., 2013) erkennbar. Im weiteren Verlauf werden die Kategorien *Feinfühligkeit* und *Rückmeldung* des INTAKT näher beleuchtet sowie deren postulierte Auswirkung auf die kindliche Entwicklung. Neben der Bindungssicherheit wirkt sich das Interaktionsverhalten von Bezugspersonen auch auf andere Bereiche aus.

Neben der Wichtigkeit der Qualität der Eltern-Kind-Interaktion auf die geistige und motorische Entwicklung von Säuglingen (Evans & Porter, 2009) konnten auch der Einfluss auf die sozial-emotionale (Kochanska, Aksan & Carlson, 2008) sowie kognitive Entwicklung (Evans & Porter, 2009) nachgewiesen werden. Ebenfalls konnten Stright, Gallager und Kelley (2008) zeigen, dass eine gute Eltern-Kind-Interaktion die akademische Kompetenzen im mittleren Kindesalter so wie die sozialen Fähigkeiten steigert.

Im Folgenden wird für die beiden Skalen *Feinfühligkeit* und *Rückmeldung* theoretisches Hintergrundwissen aufgearbeitet. Für die genauen Beobachtungskriterien innerhalb des Beobachtungssystems siehe Kapitel 7.1.2.

2.1 Feinfühligkeit

Das grundlegende Konzept der mütterlichen Feinfühligkeit geht auf Mary D. S. Ainsworth zurück (Ainsworth, Bell & Stayton, 1974). Ainsworth definierte die

mütterliche Feinfühligkeit anhand von 4 Merkmalen (Grossmann & Grossmann, 2004):

- die Wahrnehmung der kindlichen Signale
- die richtige Interpretation der Signale
- sowie Angemessenheit und
- Promptheit der mütterlichen Reaktion

Die Annahme besteht darin, dass alle gesetzten Aktionen des Säuglings Informationen für die Mutter bieten. Wenn die Mutter die Signale des Kindes wahrnimmt, lernt sie das Kind kennen. Durch die Rückmeldung des Kindes (weinen, lachen, etc.) erhält die Mutter Bestätigung ob das gesetzte Verhalten die erwünschte Wirkung hervorruft.

Der Einfluss der Feinfühligkeit auf die Entwicklung eines Kindes ist unbestreitbar. Es konnten Auswirkungen auf die kognitive Entwicklung (Feldman, Eidelman & Rotenberg, 2004), die sozio-emotionale Entwicklung (Feldman et al., 2004; Kok et al., 2013; Bouvette-Turcot, Bernier & Leblanc, 2013) sowie einen positive Korrelation mit der kindlichen Entwicklung im Allgemeinen (Belsky & Fearon, 2002) festgestellt werden. Der exklusive, hohe Zusammenhang von Feinfühligkeit und Bindungssicherheit wie in Ainsworth et al. (1974) postuliert nahm über die Forschungsjahre hinweg ab (Goldsmith & Alansky, 1987; De Wolff & van Ijzendoorn, 1997; Atkinson et al., 2000) und wurde durch weitere Interaktionsmuster, wie weiter oben bereits erwähnt, erweitert. Die Kategorie *Feinfühligkeit* im Beobachtungssystem INTAKT basiert auf den Annahmen von Ainsworth et al. (1974), und wird in Kapitel 7.1.2.1 genauer behandelt.

2.2 Rückmeldung

Im allgemeinen Sinn kann elterliche Rückmeldung als eine Art von Informationsquelle für das Kind über das eigene Verhalten verstanden werden. Eine Rückmeldung bietet einen Ankerpunkt, an dem sich das Kind orientieren und seine

eigene Leistung in Bezug zu sich selbst aber auch der Umwelt setzen kann. Kinder können hierdurch die Angemessenheit ihrer eigenen Handlung einschätzen und ihre eigenen Gütekriterien denen der Eltern anpassen oder eben abwandeln (Trudewind, Unzer & Schneider, 1989).

Es ist aber nicht nur von Bedeutung ob eine Rückmeldung statt findet sondern auch, zu welchem Zeitpunkt und welche Art von Rückmeldung. Laut Harter (1978) und Kelley, Brownell und Campbell (2000) sind positive Rückmeldungen vor allem in der frühen Kindheit wichtig für die Entwicklung des Selbstwerts, der Motivation und des Kontrollgefühls über Erfolg und Misserfolg der eigenen Handlungen.

Positive Rückmeldungen in Form von Lob können von den Eltern auch dazu eingesetzt werden, ein warmes Umfeld zu schaffen. Eine liebevolle, unterstützende und fürsorgliche elterliche Umgebung regt nach Ainsworth et al. (1978), Bowlby (1988) eine sichere Eltern-Kind-Bindung an. Im Gegensatz dazu können negative Rückmeldungen Grenzen setzen welche ebenfalls wichtig für die Eltern-Kind-Bindung sind (Bowlby, 1988).

Neben der positiven und negativen Rückmeldung kann eine weitere Art der elterlichen Bewertung identifiziert werden. Wenn der Inhalt der Rückmeldung eher negativ oder kritisch ist, aber mit der Intention auf der Korrektur des kindlichen Verhaltens liegt und mit einem neutralen oder sogar positiven Tonfall gegeben wird, kann man von einer korrigierenden Rückmeldung sprechen (Kelley et al., 2000).

Die Implikationen der elterlichen Rückmeldung wurde überwiegend in Bezug zu den Emotionen Scham und Stolz untersucht. Hier konnte die Annahme bestätigt werden, dass negative Rückmeldungen einen Einfluss auf das gezeigte kindliche Schamverhalten haben (Alessandri & Lewis, 1993, 1997; Kelley et al., 2000). Hinsichtlich der positiven Rückmeldung konnte kein Zusammenhang zum kindlichen Stolz nachgewiesen werden, jedoch gab es eine positive Korrelation zur Ausdauer bei leistungsbezogenen Aufgaben. Negative Rückmeldung konnte mit weniger Ausdauer in Bezug gebracht werden. Diese Ergebnisse stützen die

Annahmen zum Einfluss der positiven Rückmeldung auf der motivationalen Ebene (Harter, 1978), jedoch ist dieses Verhältnis relativ komplex.

Auch bei den sozio-emotionalen Verhaltensweisen konnten Zusammenhänge zur elterlichen Rückmeldung gemacht werden. In der Mehrzahl dieser Studien werden positive Rückmeldungen als Lob und negative Rückmeldungen als Kritik definiert (Owen, Slep & Heyman, 2012; Jack, Mikami & Calhoun, 2011; Cuelar & Johnson, 2009). Owen et al. (2012) konnten in ihrer Metaanalyse zeigen, dass Maßregelungen und negative, non-verbale Rückmeldungen eine Erhöhung der Folgsamkeit bei Kindern hervorruft. Dieses doch eher ungewöhnliche Ergebnis wird dadurch erklärt, dass wenn die Maßregelung stärker als die belohnende Qualität des Nicht-Befolgens ist, ein Anstieg in der Folgsamkeit beobachtet werden kann. Die Erkenntnisse bei Lob und Folgsamkeit waren weniger geradlinig. Lob fungierte nur als Verstärker von Folgsamkeit, wenn es gepaart mit einem anderen Verstärker auftrat, wie positiver Aufmerksamkeit. Den Einfluss von elterlicher Rückmeldung auf den sozialen Status von Kindern unter ihren Peers untersuchten Jack et al. (2011). Die Kinder sollten nach einer Spielrunde mit einem Elternpaar ihre Peers beurteilen. Elterliches Lob senkte die Anzahl an negativen Bewertungen von anderen Kindern für Kinder mit wenig verbaler Aggression, steigerte aber die Nominierungen bei den Kindern mit hohem Aggressionspotential. Weiteres wurden Kinder, die Kritik von ihren Eltern erhielten, öfter negativ bewertet, aber nur unter der Bedingung hoher verbaler Aggression.

Einen Zusammenhang zur kognitiven Entwicklung bei Kindern und der verbalen Stimulation durch die Eltern konnten Page, Wilhelm, Wendy Chrsi Gamble und Card (2010) nachweisen. Unter verbaler Stimulation wurden animierende und aufklärende Äußerungen und Rückmeldungen, welche dem Kind Hilfestellungen zum Erfolg der gestellten Aufgaben boten, verstanden. Diese Verhaltensweisen waren ein signifikanter Prädiktor für die erzielten Testwerte zur kognitiven Entwicklung.

Die Interaktionsverhalten wie Rückmeldung oder mütterliche Feinfühligkeit sind komplex und es bedarf einer genauen Planung um sich überhaupt erst im

Klaren zu sein, welche Aspekte für die Beobachtung des Verhaltens herangezogen werden. Was genau eine Beobachtung wissenschaftlich macht, welche Formen es gibt und auf welche Fehlerquellen zu achten sind wird im nächsten Kapitel erarbeitet.

3 Wissenschaftliche Verhaltensbeobachtung

Nachdem in den vorherigen Kapiteln ein kurzer, theoretischer Überblick der für diese Arbeit und im INTAKT enthaltenen Kategorien (*Feinfühligkeit, Rückmeldung*) zur Beobachtung der Eltern-Kind-Interaktion gegeben wurde folgt nun ein kurzer Abstecher in die theoretischen Gefilde der Verhaltensbeobachtung.

Die Realität ist ein Abbild dessen, was wir mit all unseren Sinnen wahrnehmen. Wenn wir unsere Aufmerksamkeit auf Beispielsweise eine Situation lenken, in der eine Mutter mit ihrem Kind spielt, dann sehen wir bereits nur mehr einen Teil der Realität. Unsere Sinne sind auf diese spezielle Szenerie fokussiert. Wir haben also unseren Anzeigevorgang alleine durch die Beobachtung des Mutter-Kind-Verhaltens per se definiert. Faßnacht (1995) beschreibt Beobachtung als eben diesen „Anzeigevorgang“ von „Verhaltens-Realität“ und definiert weitere Anzeigevorgänge welche er primäre, sekundäre und tertiäre Beiordnung nennt. Bei der primären Beiordnung findet eine gänzlich organisch bedingte Selektion der „Realität“ statt, das bedeutet, dass der/die BeobachterIn etwas sieht und hört, aber ebenso andere Signale wie einen Geruch oder die Temperatur wahrnimmt. Bei der sekundären Beiordnung werden die Wahrnehmungen miteinander verbunden und zu einem Eindruck organisiert. Natürlich kann und wird dieser Eindruck von unseren Vorerfahrungen beeinflusst. In der tertiären Beiordnung werden die Eindrücke des/der BeobachterIn dann in Zeichensysteme wie Zahlen, Codes oder Sprache transformiert und somit die Übermittlung an andere Personen möglich gemacht. Um es stark vereinfacht auszudrücken: eine Beobachtung besteht aus Selektion, Organisation und Kodierung.

Spricht man aber von einer wissenschaftlichen Beobachtung, werden diese drei Aspekte erweitert. Laut Greve und Wentura (1997) ist die Absicht die erhobenen Daten auszuwerten sowie die Kriterien der Replizierbarkeit und Objektivität der Beobachtung kennzeichnend für die wissenschaftliche Methodik. Bodenmann (2006) definiert hierzu für den Fall des diagnostischen Zugangs weitere Punkte. Er schreibt von einer systematischen und kontrollierten Herangehensweise, aber auch die Definition von Beobachtungskategorien, Beobachtungs-

zeiträumen und Beobachtungsintervallen sowie offenem oder verdecktem Verhalten aus der fremd- oder selbstperspektive sind Definitionskategorien.

3.1 Formen der Verhaltensbeobachtung

In der Literatur gibt es viele verschiedene Klassifikationen von psychologischer Beobachtung. Die Unterschiede können jedoch stets anhand von Systematik, Setting, Status des/der BeobachterIn, dem Zeitpunkt sowie der Schärfe der Analyseinheiten getroffen werden. Beispiele hierfür sind systematische vs. unsystematische, frei vs. gebundene, verdeckte vs. offene, direkte vs. indirekte Beobachtung oder Selbst- vs. Fremdbeobachtung (Greve & Wentura, 1997; Bodenmann, 2006; Schölmerich, 2011). Diese Unterscheidungen sind eindeutig, da die Bezeichnungen der Klassifikationsdimensionen evident sind. Der Unterschied zwischen technisch vermittelnder und unvermittelnder Beobachtung (Greve & Wentura, 1997) ist für diese Arbeit insofern interessant, da die Frage aufgeworfen wird, ab wann ein technisches Hilfsmittel als solches überhaupt definiert wird. Worin besteht der Unterschied im Bezug zur technischen Unterstützung zwischen einer Videokamera, einem Nachtsichtgerät oder einer einfachen Brille? Ein angemessenes Kriterium wäre laut Greve und Wentura (1997) die Möglichkeit des Abspeicherns der Beobachtung. In Kapitel 3.4 wird noch genauer auf die Videotechnik eingegangen.

3.2 Quantifizierung in der Verhaltensbeobachtung

Im Folgenden sollen die drei häufigsten Verfahren zur Quantifizierung von Verhalten laut Faßnacht (1995) näher beleuchtet werden: das Event-Sampling-Verfahren, das Time-Sampling-Verfahren und das Rating-Verfahren. Alle drei Verfahren kommen bei den Kategorien *Feinfühligkeit* und *Rückmeldung* des Beobachtungssystems INTAKT zu tragen.

Im Event-Sampling-Verfahren werden ein oder mehrere Verhaltensweisen möglichst genau im Sinne von Auftreten und Nicht-Auftreten kodiert, d.h. im na-

türlichen Verhaltensstrom wird dann kodiert, wenn das gewählte Verhalten auftritt (Bodenmann, 2006). Stellt man sich die zu beobachtenden Sequenz als eine gerade Linie vor, dann wird jedes Mal, wenn ein vorher definiertes Verhalten auftritt, auf dieser Linie ein Strich vermerkt. Natürlich ist es auch möglich, die Länge des aufgetretenen Verhaltens zu kodieren womit man ein gutes Abbild über die Häufigkeit und Dauer des Verhaltensmerkmals erhält. Beim Time-Sampling-Verfahren wiederum wird der Verhaltensstrom in kurze, kontinuierlich aufeinanderfolgende Zeiteinheiten (Einheitsintervalle) unterteilt und analysiert (Faßnacht, 1995). Es wird also die zu beobachtende Sequenz in gleich lange Intervalle unterteilt und jedes einzelne Intervall wird von dem/der BeobachterIn mittels Ratingskala bewertet. Bei beiden Methoden besteht natürlich die Möglichkeit der Bewertung des Verhaltens durch eine Ratingskala. Hier wird davon gesprochen, dass der/die BeobachterIn den Ausprägungsgrad der Verhaltensweise an vorher definierten Ankerpunkten abschätzt.

3.3 Beobachtungssysteme: Kodierung

Neben der Form und der Art der Verhaltensbeobachtung muss auch definiert werden, mit welchem System das Verhalten beschrieben wird bzw. wie das zu beobachtende Verhalten kodiert wird. Bei diesen Beschreibungssystemen handelt es sich um tertiäre Anzeigesysteme (Faßnacht, 1995), welche erst durch den Bezug zum beobachteten Verhalten wie *Feinfühligkeit* und *Rückmeldung* an Bedeutung gewinnen.

Verbal-Systeme stellen die einfachste Form der Verhaltenskodierung da. Es wird in vollem Umfang mit sprachlichen Mitteln das zu beobachtende Verhalten beschrieben. Faßnacht (1995) unterscheidet zwischen Tagebuchaufzeichnungen, Selbst- und Fremdaufzeichnungen, Verlaufsprotokollen und der Methode der kritischen Vorfälle.

Nominal-Systeme sind einfache, qualitative Klassifikationssysteme welche im Gegensatz zu Verbal-Systemen auf ganze Sätze oder das Einhalten von Grammatik verzichten. Es müssen nicht nur Wörter verwendet werden, es ist auch möglich

auf Graphiken zurückzugreifen. Nominal-Systeme können in Zeichen- und Indexsysteme oder Kategorien-Systeme (Faßnacht, 1995) unterteilt werden. Bei Ersterem werden vorher definierte Zeichen für ein bestimmtes Verhalten festgelegt, welche dann als Indikator für die übergeordnete, zu beobachtende Eigenschaft oder das Konstrukt gelten. Diese Zeichen für eine Verhaltensweise sind beliebig erweiterbar und schließen sich nicht gegenseitig aus, ebenso kann es zu kodierfreien Phasen kommen. Weiters kann durch die Häufigkeit der Kodierung eines oder mehrerer Indikatoren auch auf die Intensität des Verhaltensaspekts geschlossen werden (Schölmerich, 2011).

Bei Kategorien-Systemen schließen sich die Prädikatoren gegenseitig aus und decken vollständig die zu erfassende Kategorie ab. Die Kategorien sind unabhängig voneinander und sollen den zu untersuchenden Verhaltensaspekt umfassend beschreiben, daher handelt es sich um ein logisch geschlossenes und inhaltlich vollständiges System.

Bei Dimensional-Systemen verhält sich die Relation der Einheiten wie bei einem Kategorien-System, mit dem Zusatz der geradlinigen Geordnetheit (Faßnacht, 1995). Man stelle sich eine Gerade mit formal gleichen Einheiten vor, die lückenlos angeordnet sind. Trotz der Gleichheit der Einheiten sind sie aufgrund ihrer Position auf der Geraden in Bezug auf einen Referenzpunkt doch nicht gleich. Es lassen sich drei verschiedene Dimensional-Systeme definieren: Ordinalskala, Intervallskala und Ratioskala.

Das letzte Beschreibungssystem laut Faßnacht (1995) nennt sich Struktural-System. In diesem System befinden sich die Einheiten in einer komplexeren, räumlichen Anordnung als es bei den Dimensional-Systemen der Fall ist, sie können gleich als auch verschieden sein.

Beim INTAKT handelt es sich also um ein systematisches, gebundenes, offenes und indirektes Beobachtungssystem welches stimulierte Verhalten durch Fremdbeobachtung anhand der Kategorien *Feinfühligkeit*, *Rückmeldung* und *Joint-Attention* quantifiziert. Systematisch und gebunden da Beginn, Dauer und Auswertung sowie klare inhaltliche Konzepte zur Beobachtung vorliegen. Offen und

indirekt da Wissen über die Erhebungssituation bei den zu beobachteten Personen gegeben ist und es sich um eine Videoanalyse mit Datengewinnung durch Kodierung im nach hinein handelt. Zum genauen Durchführungsprozedere wird auf Kapitel 7 verwiesen.

Die angeführten Begriffsdefinitionen bei der Verhaltensbeobachtung und Kodierung wurden gegeben, da sich die Frage stellte, inwiefern die Kodiersysteme für die Skalen der *Feinfühligkeit* und *Rückmeldung* des INTAKT einer Neuausrichtung in Bezug auf die Kodierart bei computergestützt vs. händischer Kodierung erfahren muss. Die Überlegungen von Zouzoula (2015) wurden hierfür adaptiert und können im Kapitel 8.3 nachgelesen werden.

3.4 Verhaltensbeobachtung mittels Videotechnik

Die Videotechnik hat in der Verhaltensbeobachtung, aber auch in der Verhaltens- und Interaktionsanalyse, einen hohen Stellenwert. Thiel (2011) schreibt der Video- und Filmtechnik sogar eine essentielle Bedeutung zu wenn er postuliert dass sich die Verhaltensbeobachtung ohne diese technische Hilfe nicht zur relevanten Methode entwickelt hätte.

Natürlich lassen sich die Vorteile der Videotechnik nicht von der Hand weisen. Das Geschehen kann beliebig oft betrachtet werden, es kann je nach Programm beliebig genau betrachtet werden (Minuten, Sekunden, Bilder pro Sekunde) und es können beliebig viele BeobachterInnen die Situation bewerten.

Durch die Beobachtung mittels Videotechnik wird eine gewisse „Außenperspektive“ eingenommen, ob diese jedoch rein objektiver Natur ist, darf angezweifelt werden. Denn schon alleine durch die Position der Kamera oder des Aufnahmegerätes, die aus praktikablen Gründen meist statisch ist, ist das zu beobachtende Szenario eingeschränkt. Doch diese Beschränkungen hängen nicht mit dem System der Videoaufzeichnung per se zusammen, sondern haben den Ursprung bereits im vorab definierten Beobachtungsablauf. Eine Kamera bildet nur den von ForscherInnen und BeobachterInnen vorher definierten Ausschnitts der „Realität“ ab. Jedoch kann eine Videoaufzeichnung das Ausgewählte in einer anderen Art

wiedergeben als ein/e BeobachterIn, da diese/r bestimmten Einflussfaktoren unterstellt ist (Thiel, 2011) (siehe Kapitel 3.5). Da auch nur bei kurzen Sequenzen einer dyadischen Interaktion viele Eindrücke auf den/die BeobachterIn einwirken können, ist es bei der Auswertung immens wichtig vorab das Beobachtungssystem genau zu definieren und die Beschreibungssysteme bzw. die Quantifizierung des Verhaltens einer genauen Ausarbeitung zu unterziehen.

In welchen Formen das Kodieren eines Verhaltens erfolgen kann wurde bereits in Kapitel 3.3 abgehandelt, jedoch definiert Faßnacht (1995) explizit für die Video-Beobachtung und im speziellen beim Event-Sampling Verfahren vier unterschiedliche Methoden.

Bei der Zeit-Einblendung handelt es sich um eine Methode, bei der eine Uhr im Videobild eingeblendet wird durch welche die Anfangs- und Endzeit der zu beobachtenden Verhaltenseigenschaft abgelesen und tabellarisch notiert wird. Eine der Live-Beobachtung ähnliche Methode ist die Merkmal-kontingente-Koppelung bei welcher das Verhalten mit einem Ereignisschreiber ohne zeitliche Koppelung kodiert wird. Der Nachteil dieser Methode ist jedoch, dass die Kodierung der gesamten Videosequenz ohne Unterbrechung erfolgen muss. Bei der CTL-Koppelung ist ein Computer direkt mit der Videoaufnahme verbunden, was wiederum eine sehr genau Kodierung zu erwarten lässt. Am genauesten kodiert werden kann mit der Zeitkode-Koppelung. Hier wird der Ereignisschreiber mittels Zeitkode der Videoaufnahme synchronisiert.

Natürlich ist die Verhaltensbeobachtung mit Hilfe der Videoaufzeichnung auch durch eine gründliche Planung nicht vor Fehlern gefeit. Welche Beobachtungsfehler auftreten können und wie diese auch in Bezug zur Videotechnik stehen wird im nächsten Kapitel erläutert.

3.5 Beobachtungsfehler in der wissenschaftlichen Verhaltensbeobachtung

Wie zuvor schon erwähnt ist die Realität das, was wir mit unseren Sinnen erfassen. Jedoch ist es nicht so dass unsere Umwelt einfach ungefiltert wahrgenommen wird. Unsere Gehirn „hilft“ uns, zum Beispiel bei optischen Täuschungen, oder verbindet bereits abgespeicherte Konzepte mit den aufgenommen Reizen unserer Sinnesorgane.

Ein Beispiel hierfür wäre die Relativität von Größen auf ein Bild oder Kippfiguren. Dies bedeutet, dass unser Gehirn bis zu einem gewissen Grad automatisch entscheiden kann welche Informationen für uns wichtig sind und welche nicht. Es existiert also ein komplexes System welches Einfluss auf unsere Wahrnehmung aber auch auf unser Denken nimmt. Natürlich können in diesem System auch Fehler entstehen wie vor allem die kognitive Psychologie zeigt („bottom up“ vs. „top down processing“-Prozesse). Abgesehen von diesen elementaren Fehlern in der Wahrnehmung definieren Greve und Wentura (1997) eine Reihe von Problemen die im „Prozess“ von Selektion, Organisation und Kodierung auftreten können, nämlich: Fehler zu Lasten des/der BeobachterIn, Fehler zu Lasten der Beobachtung und Fehler zu Lasten äußerer Bedingungen.

3.5.1 Fehler zu Lasten des/der BeobachterIn

Persönlichkeitsfaktoren wie die Stimmung, die zum Zeitpunkt der Beobachtung vorhandene Konzentrationsfähigkeit aber auch persönliche Dispositionen gegenüber des zu beurteilenden Individuums beeinflussen die Bewertung des/der BeobachterIn negativ im Sinne der Qualität. Greve und Wentura (1997) unterscheiden drei Fehlerquellen: Wahrnehmungsfehler, Deutungsfehler und Interpretationsfehler sowie Erinnerungsfehler (Kapazitätsprobleme und Selektions- bzw. Modifikationsfehler).

Faßnacht (1995) definiert bei der Beurteilung von Personen bzw. bei der Beobachtung einer Situation mit mehreren Personen 5 häufige Fehler: zentrale Ten-

denzen, Milde-Härte-Fehler, logische Fehler, Halo-Effekt und Primacy-Recency-Effekt.

Der zentrale Tendenzen Effekt ist ein für Ratingskalen relevantes Phänomen. Hier wurde beobachtet dass bei wiederholten Erhebungen Extremwerte zur Skalenmitte tendieren. Gerade bei der Kategorie *Feinfühligkeit* des INTAKT, bei der eine Ratingskala zum Einsatz kommt, sollte darauf geachtet werden die zu bewertenden Sequenzen los gelöst vom gesamten Video zu betrachten. Ein weiterer Fehler bei der Personenbeurteilung ist das systematisch zu positiv oder negativ Bewerten einer Person, der Milde-Härte-Fehler. Bei Greve und Wentura (1997) würden diese beiden Fehler unter die Klassifikation Interpretationsfehler fallen, also Fehler welche das Wahrgenommene beeinflussen.

Der logische Fehler lässt sich damit erklären, dass der/die BeobachterIn eine persönliche Theorie in die Beobachtungssituation mit einbringt und diese dann die Beurteilung der Daten beeinflusst. Ähnliches passiert beim Halo-Effekt. Der/Die BeobachterIn lässt sich bei der Beurteilung eines Persönlichkeitsmerkmals oder einer Eigenschaft vom Gesamteindruck der zu beobachtenden Person beeinflussen. Dieser Effekt tritt gehäuft auf, wenn das zu beobachtende Merkmal im Allgemeinen schwer zu beobachten sowie ungewöhnlich ist, aber auch wenn es zuvor schlecht definiert wurde (Bortz & Döring, 2006). Für den/die BeobachterIn und KodiererIn der INTAKT-Videosequenzen gilt es darauf zu achten Sympathien oder Antipathien nicht in die Wertung der Feinfühligkeitsskala einfließen zu lassen. Diese Beobachtungsfehler bestimmen bereits das bewusst Wahrgenommene durch den/die BeobachterIn und werden daher nach Greve und Wentura (1997) Wahrnehmungsfehler bezeichnet.

Als Erinnerungsfehler wird zum Beispiel der Primacy-Recency-Effekt bezeichnet. Hier hat die Anfangs- und Endposition des zu bewertenden Objekts maßgeblichen Einfluss auf die nachfolgende Beurteilung. Der erste sowie der neuste Eindruck haben starken Einfluss auf die Bewertung. Auch hier sollte man sich bei der Bewertung der 2-minütigen Videosequenzen zur *Feinfühligkeit* des Effekts bewusst sein.

Diesen Fehlern kann jedoch mit einer gründlichen Einschulung des/der BeobachterIn vorbeuge geleistet werden. Je besser das BeobachterInnentraining und die Auswahl der BeobachterInnen selbst, desto weniger Fehler werden begangen (Bodenmann, 2006).

3.5.2 Fehler zu Lasten der Beobachtung

Das BeobachterInnentraining kann von noch so einer hohen Qualität sein, ein unangemessens Beobachtungssystem beeinflusst das Ergebnis. Sind die vorab definierten Kategorien nicht gut ausdifferenziert kann es möglicherweise die Reliabilität der BeobachterInnen oder KodiererInnen beeinflussen (Faßnacht, 1995; Greve & Wentura, 1997).

3.5.3 Fehler zu Lasten äußerer Bedingungen

Gerade bei der Videobeobachtung kann es dazu kommen, dass die „Realität“ durch ungünstigen Lichteinfall oder Probleme mit der Tonspur nicht richtig bei dem/der BeobachterIn ankommt. Diese Fehlerquellen können durch Sorgfalt in der Planung eingedämmt werden.

Nach dem Aufzeigen der wichtigsten Beobachtungsfehler in der Verhaltensbeobachtung sollen im nächsten Kapitel die grundlegenden Konzepte der Objektivität, Reliabilität und Validität im Bezug auf das händische Kodiervorgang des INTAKT beleuchtet werden. Denn wie schon in der Einleitung erwähnt sind die Gütekriterien ein Schwachpunkt bei einer großen Anzahl an Verhaltensbeobachtungsverfahren.

4 Gütekriterien der Verhaltensbeobachtung

Um die Vergleichbarkeit von diagnostischen Verfahren wie Tests, Fragebögen oder wie im Fall dieser Arbeit Beobachtungen sicher zu stellen, werden im Sinne der Qualitätssicherung bestimmte Gütekriterien herangezogen. Lienert und Raatz (1998) geben drei Hauptgütekriterien an, welche in diesem Abschnitt erläutert werden sollen. Auf eines der vier Nebengütekriterien, die Ökonomie, wird ebenfalls eingegangen.

4.1 Objektivität

Die Objektivität gibt an, in welchem Ausmaß die Ergebnisse der Beobachtung unabhängig von dem/der BeobachterIn sind (Bortz & Döring, 2006). Die BeobachterInnenunabhängigkeit kann in drei Unterpunkte aufgeteilt werden: Durchführungsobjektivität, Auswertungsobjektivität und Interpretationsobjektivität. Gerade die Objektivität ist ein Gütekriterium das nach Bortz und Döring (2006) „recht unproblematisch“ ist, da die Standardisierung der Durchführung, Auswertung und Interpretation des gesamten Verfahrens die Objektivität prinzipiell garantiert. Hier kann auf die ausführliche Einschulung zum INTAKT in Kapitel 8.1 und auf das Manual verwiesen werden.

4.2 Validität

Von Validität spricht man wenn das Verfahren jenes Merkmal misst, welches gemessen werden soll. Die Validität ist das wichtigste Gütekriterium überhaupt denn sie erlaubt eine Generalisierung des in der Test oder Beobachtungssituation erfassten psychologischen Merkmals auf das Zielmerkmal außerhalb der Testsituation. Es ist aber auch das am schwierigsten zu erfassende Kriterium (Bortz & Döring, 2006; Kubinger, 2009), denn: Hohe Reliabilität ist eine notwendige Voraussetzung für hohe Validität, aber kein Nachweis (Greve & Wentura, 1997).

Man unterscheidet zwischen Inhaltlicher Validität, Kriteriumsvalidität und Kon-

struktvalidität. Für diese Arbeit ist vor allem die Kriteriumsvalidität von Bedeutung da sie sich statistisch erfassen lässt. Gemessen wird sie durch Korrelation eines „Außenkriteriums“ mit dem interessierenden Test. Außenkriterien können in zwei verschiedene Kategorien unterteilt werden: Kriterien welche in ferner Zukunft liegen („prognostische Validität“) oder Kriterien welche auf eine „Übereinstimmungsvalidität“ abzielen, d.h. auf die Korrelation mit einem anderen Test, der dasselbe Konstrukt erfasst. Im Fall dieser Arbeit ist das Außenkriterium die händische Kodierung, welche mit der Kodierung über das Computerprogramm INTERACT verglichen wird.

4.3 Reliabilität

Unter Reliabilität versteht man die Akkuratez der Messung des Verfahrens hinsichtlich des zu messenden Merkmals. Losgelöst davon, was gemessen wird, sollten die unter den gleichen Bedingungen gewonnenen Ergebnisse einer Testperson übereinstimmen (Kubinger, 2009). In der klassischen Testtheorie wurden mehrere Verfahren zur Bestimmung der Reliabilität entwickelt: Retest-Reliabilität, Paralleltest-Reliabilität, Split-Half-Reliabilität und Innere Konsistenz.

Bei der Retest-Reliabilität wird ein Test einer Testperson an zwei verschiedenen Zeitpunkten vorgelegt und die Korrelation zwischen den beiden Testzeitpunkten errechnet. Durch diese Methode wird aber hauptsächlich die zeitliche Stabilität des Testwerts ermittelt, umgemünzt auf das Beobachtungssystem INTAKT kann dadurch die Stabilität der Benutzung der Skala des/der BeobachterIn über die Zeit hinweg erfasst werden. Hier wird auf das Kapitel 8.4 zur Begutachtung der Konsistenz des Kodierers dieser Arbeit verwiesen. Wirtz und Caspar (2002) nennen zum Vergleich der Beurteilung von BeobachterInnen noch die Interraterreliabilität. Wie genau die Messung dieser erfolgt wird noch näher in Kapitel 8.4 im Zuge der Intraklassenkorrelation erläutert.

4.4 Ökonomie

Lienert und Raatz (1998) bezeichnen einen Test oder ein Verfahren als ökonomisch wenn geringe Material- und Zeitkosten (Durchführung und Auswertung) sowie eine leichte Handhabung gegeben sind. Bortz und Döring (2006) und Kubinger (2009) differenzieren ein wenig genauer und ermitteln zwei Komponenten: den zeitlichen und den finanziellen Faktor.

Der zeitliche Aufwand setzt sich nicht nur aus der Nettotestzeit zusammen sondern auch aus der Vorbereitung, der Auswertung und der Rückmeldung. Doch der für diese Arbeit entscheidendere Aspekt ist die Wirtschaftlichkeit. Hierunter fallen die eigentlichen Anschaffungskosten für das Verfahren, Personalkosten aber auch Lizenzgebühren.

Da die händische Kodierung des Beobachtungssystems INTAKT auch aus der Notwendigkeit eines ökonomischeren Kodier- und Auswertungsschemas entstanden ist wird im nächsten Kapitel auf die Vor- und Nachteile von Computer- und Papier-Bleistift-Verfahren eingegangen.

5 Papier-Bleistift oder computergestützte Kodierung

Die Transformation eines Beobachtungssystem, welches computerbasiert kodiert und ausgewertet wird hin zu einem händischen Kodier- und Auswertungssystem ist in der vom Autor gesichteten Literatur nicht gängig. Es kann aber auf Studien von Gottman (1983), Richard Bakeman und Adamson (1984) und Richard Bakeman und Brownlee (1982) verwiesen werden, wo mit dem Bleistiftverfahren kodiert wurde und gleichzeitig technische Hilfsmittel in Gebrauch waren.

Es gilt auch zwischen der eigentlichen Computerdiagnostik und computergestützten Test zu unterscheiden. Nur der bloße Einsatz zur Auswertung eines Verfahrens kann nach Kubinger (2009) nicht als Computerdiagnostik bezeichnet werden. Jedoch stellt sich die Frage, ob die Annehmlichkeiten bei der Kodierung, welche durch ein Abspielprogramm wie dem INTERACT (siehe Kapitel 7.1.3) entstehen, auch Auswirkung im Vergleich zu einem schlichteren Programm und einer Kodierung und Auswertung per Hand zeigen.

Im Lauf der Zeit wurden viele Verfahren „modernisiert“ und so von Papier-Bleistift-Verfahren hin zu einer computerbasiert Unterstützung transformiert (Mead & Drasgow, 1993; Schulenberg & Yutrzenka, 1999; Kindsvater & Sturm, 2003; Wijndaele et al., 2007). Hier handelt es sich oft um Tests und weniger um Beobachtungssysteme, da computergestützte Automatisierung beim tatsächlichen Akt der sekundären Beiordnung (noch) nicht möglich ist. Gerade hier ist die Äquivalenzforderung von großer Bedeutung und unumgänglich, da die gewonnen Normen, Reliabilitäts- und Validitätsnachweise nur durch eine Überprüfung übertragbar sind. Nach Meinung des Autors dieser Arbeit sind die Forderungen nach psychometrischer Äquivalenz, subjektiv perzipierte bzw. erfahrungsbezogener Äquivalenz und populationsspezifischer und individuumsspezifischer Äquivalenz (Klinck, 2006) auch in die andere Richtung, d.h. von einer vorher computergestützten Kodierung zu einer Papier-Bleistift-Kodierung, unerlässlich.

Die Vorteile einer computergestützten Testung sind einleuchtend und schnell aufgezählt: neben hoher Durchführungs- und Auswertungsobjektivität sind auch eine höhere Ökonomie sowie ausgeweitete Testbedingungen (Einsatz von inno-

vativen Itemtypen, besserer Testschutz sowie spezifischeres Erfassen des Testverhaltens) Punkte die für eine Computerisierung sprechen.

Roger Bakeman und Gottman (1997) geben zu bedenken, dass computergestütztes Kodieren einen hohen „Einkaufspreis“ haben kann, nicht nur durch die Anschaffungskosten, sondern auch durch die Zeit, welche benötigt wird sich mit dem Programm vertraut zu machen. Wobei sich hier der positive Aspekt der Ökonomie für computerbasierte Kodiersysteme ins Gegenteil verkehrt und auch einen Grund für diese Arbeit liefert. Weiters können der leichtere Transport, die geringeren Ersatzkosten (und Kosten im Allgemeinen) sowie der universelle Einsatz unabhängig von den Gegebenheiten (Live-Beobachtung oder Videoanalyse) als Vorteile für eine Papier-Bleistift-Kodierung genannt werden (Roger Bakeman & Gottman, 1997).

Empirischer Teil

6 Hintergrund der Untersuchung

Im bisherigen Teil dieser Arbeit wurde ein Einblick in die Eltern-Kind-Interaktion und im Speziellen auf die Konzepte *Feinfühligkeit* und *Rückmeldung* gegeben. Ebenfalls wurde auf die Verhaltensbeobachtung sowie die verschiedenen Kodiermöglichkeiten und Systeme aber auch die möglichen Gefahren bei der wissenschaftlichen Beobachtung eingegangen und ein kurzer Einblick in die Videoanalyse bzw. computer- und händische Kodierung sowie deren Gütekriterien gegeben.

Das Beobachtungssystem INTAKT (Hirschmann et al., 2013) bietet die Möglichkeit der qualitativen Erfassung der Interaktion von Bezugsperson und Kind anhand der Kategorien *Feinfühligkeit*, *Rückmeldung* und *Joint-Attention*. Die bisherigen Untersuchungen zur Validierung sowie Reliabilitäts- und Objektivitätsanalyse wurden mittels INTERACT und somit mit einem computerbasiertem Auswertungsprogramm durchgeführt. Im Sinne des Nebengütekriteriums der Ökonomie wurde eine Auswertungsverfahren, welches auch händisch und ohne zusätzliche Software auskommt, angestrebt. Bei Zouzoula (2015) und Thoden (2014) wurde genau dieser Gedanke umgesetzt und für die Kategorien *Feinfühligkeit* und *Rückmeldung* respektive *Joint-Attention* ein Papier-Bleistift-Verfahren zur Kodierung entwickelt welches hinsichtlich der Hauptgütekriterien überprüft wurde.

In dieser Arbeit soll das händische Kodierungsschema von Zouzoula (2015) mit der herkömmlichen Computerkodierung verglichen werden, um weitere Rückschlüsse auf die Validität zu ziehen. Gemäß der Kriteriumsvalidität, im genaueren der Übereinstimmungsvalidität, soll computergestützte Kodierung als externes Kriterium für die händische Kodierung dienen. Die Validierung erfolgt nur für die Kategorien *Feinfühligkeit* und *Rückmeldung*.

6.1 Forschungsfrage

Ziel dieser Arbeit ist es, das bereits konstruierte händische Auswertungssystem für das Beobachtungssystem INTAKT mit den Ergebnissen der computergestützten Kodierung zu vergleichen.

Die zu beantwortende Frage lautet also: Inwieweit stimmen die Ergebnisse der händischen Kodierung des Beobachtungssystems INTAKT bezüglich der Kategorien *Feinfühligkeit* und *Rückmeldung* mit denen der computergestützten Kodierung überein?

Es wird eine Überprüfung der BeobachterInnenübereinstimmung mittels Intraklassenkorrelation (ICC) und dem gewichteten Kappa-Koeffizient nach Cohen (κ) erfolgen. Zur Überprüfung der Übereinstimmungsvalidität wird eine Produkt-Moment-Korrelation (r_{PM}) errechnet. Alle Berechnungen werden für die Kategorien *Feinfühligkeit* und *Rückmeldung* erfolgen.

7 Untersuchungsinstrument

In diesem Kapitel wird das in der Untersuchung verwendete Beobachtungssystem INTAKT (Hirschmann et al., 2013) sowie dessen Durchführung und Beobachtungskategorien näher beschrieben.

7.1 Beobachtungssystem INTAKT

Das Beobachtungssystem INTAKT ist ein Video-Beobachtungsverfahren zur Einschätzung der Qualität der Mutter-Kind-Interaktion. Hirschmann et al. (2011) konnten zeigen dass die definierten Beobachtungskategorien reliabel und valide durch den INTAKT erfasst werden können. Im weiteren Verlauf wird kurz auf das Setting bei der Videoaufnahme und die Beobachtungskategorien des INTAKT eingegangen. Grundlage für diese Informationen ist das 2. unveröffentlichte Manual des INTAKT (Hirschmann et al., 2013).

7.1.1 Beobachtungssituation

Zur Durchführung der Videoaufzeichnung wurden zwei verschiedene Situationen kreiert. Bei der Bastelaufgabe wird den Mutter-Kind-Dyaden Materialien zur Verfügung gestellt mit der standardisierten Instruktion die Schablone eines Hauses auf Papier zu verschönern. Die freie Spielsituation wird durch eine Kiste mit diversen Spielmaterialien ergänzt. Die genaue Sitzposition des Kindes und der Mutter sowie die Anordnung der Bastel- und Spielmaterialien sind im Manual angeführt. Zeitbeschränkungen für die Bastelsituation sind nicht vorgesehen, bei der Spielsituation kann jedoch nach 20-30 Minuten eine Beendigung eingeleitet werden.

7.1.2 Beobachtungskategorien

Im folgenden werden die Verhaltensdimensionen sowie deren systematische Erfassung innerhalb des INTAKT-Beobachtungssystems erläutert.

7.1.2.1 Feinfühligkeit

Die mütterliche *Feinfühligkeit* ist laut Ainsworth et al. (1978) ein wesentlicher Baustein für eine sichere Bindungsbeziehung und ebenso ein wichtiger Teil der Mutter-Kind-Interaktion. Im genaueren betrifft *Feinfühligkeit* die Fähigkeit der Mutter unverzüglich und angemessen auf die Signale und Bedürfnisse des Kindes zu reagieren. Dies setzt natürlich voraus dass sie im Vorhinein wahrgenommen und richtig interpretiert wurden. Ainsworth et al. (1974) entwickelten hierzu eine 9-stufige Skala mit fünf definierten Kategorien welche auch als Basis der 7-stufigen Ratingskala des INTAKT dient. Für den theoretischen Hintergrund siehe Kapitel 2.1.

Im INTAKT wird auf der Skala von (1) *sehr geringe Feinfühligkeit*, (2) *geringe Feinfühligkeit*, (3) *eher geringe Feinfühligkeit*, (4) *mittlere Feinfühligkeit*, (5) *eher hohe Feinfühligkeit*, (6) *hohe Feinfühligkeit* bis (7) *sehr hohe Feinfühligkeit* kodiert.

Im folgenden wird die Verhaltensdefinition für den höchsten und den niedrigsten Ankerpunkt der Skala angeführt. Im Manual sind ebenfalls Definitionen zu den Ratings (3) und (5) zu finden, falls keiner der Ankerpunkte auf das Verhalten zutrifft wird eine der Zwischenstufen (2), (4) oder (6) herangezogen.

Ankerpunkt (1) sehr geringe Feinfühligkeit:

Die Bezugsperson verfolgt ihre eigenen Ideen und Handlungsabsichten, ohne das Kind miteinzubeziehen und stimmt ihr Handeln nicht auf die kindlichen Wünsche und Ideen ab. Ebenso wenig fördert sie die Initiativen des Kindes. Allgemein wirkt die Bezugsperson wenig interessiert an der Interaktion mit dem Kind und kann den kindlichen Blickwinkel schwer einnehmen. Die Bezugsperson reagiert erst bei mehrmaligen verbalen Äußerungen des Kindes, da nonverbale Signale nicht wahrgenommen werden. Der Kommunikationsstil der Bezugsperson ist nicht kindgemäß und die Bezugsperson gibt dem Kind keinerlei motivierende Rückmeldungen (Hirschmann et al., 2013).

Ankerpunkt (7) sehr hohe Feinfühligkeit:

Der Bezugsperson gelingt es, im vollen Maße auf die kindlichen Wünsche und Ideen einzugehen und ihre Handlungen an diese anzupassen. Sie reagiert angemessen und prompt auf die kindlichen Signale und motiviert das Kind fortlaufend durch ihre Rückmeldungen. Die Interaktion wirkt harmonisch und reibungslos. Die Bezugsperson nimmt den kindlichen Blickwinkel sehr gut ein und nutzt stets einen kindgerechten Kommunikationsstil (Hirschmann et al., 2013).

Feinfühligkeit - unkodierbar:

Wenn mindestens die Hälfte des Intervall die Mutter oder das Kind nicht sichtbar sind wird die Kategorie unkodierbar gewählt.

Mit Hilfe des Time-Sampling-Verfahrens erfolgt eine Rating über jeweils eine 2-Minuten-Einheit. Dieser Rhythmus wird über die gesamte Dauer des Videos beibehalten.

7.1.2.2 Rückmeldung

Die theoretische Grundlage für die Kategorie *Rückmeldung* findet sich in Kapitel 2.2 wieder, hier soll ein Überblick zur Kodierung und Definition im INTAKT-Beobachtungssystem gegeben werden.

Positive Rückmeldung:

Es erfolgt in einer emotional positiv konnotierten Art und Weise bzw. in einem freundlichen Tonfall eine Zustimmung oder Verstärkung hinsichtlich des kindlichen Verhaltens. Inbegriffen sind hier auch lobende Äußerungen hinsichtlich der Ergebnisse oder Handlungen des Kindes. Im Manual sind Beispiele für *positive Rückmeldungen* wie „Super!“ und „Gut gemacht!“ angeführt und es wird darauf hingewiesen dass die Äußerungen „Mhm“ nur in Kombination mit einer weiteren

positiven Aussage kodiert wird, ebenso „genau“ bzw. „ja“, welche klar positiv moduliert ausgesprochen werden müssen (Hirschmann et al., 2013).

Korrigierende Rückmeldung:

Unter der *korrigierenden Rückmeldung* werden informative und hilfreiche Äußerungen und Handlung der Bezugsperson dem Kind gegenüber verstanden, wodurch es seine Handlungen bzw. Ergebnisse verbessern kann. Diese Art von Rückmeldung wird in einem neutralen oder freundlichen Tonfall übermittelt, wenngleich sie inhaltlich negativ oder kritisch ist. Die korrigierende Rückmeldung kann als Entscheidungshilfe für das Kind dienen und so entwicklungsunterstützend sein. Ein wichtiger Punkt um eine Kodierung dieser Kategorie vorzunehmen ist das tatsächliche Handeln des Kindes. Das bedeutet es wird erst kodiert wenn das Kind tatsächlich etwas getan hat, was korrigiert werden kann. Weiteres ist es nicht ausreichend wenn dem Kind nur gesagt wird, was etwas nicht ist bzw. werden falsche Erklärungen ebenfalls nicht kodiert (Hirschmann et al., 2013).

Negative Rückmeldung:

Es wird Ablehnung und Unzufriedenheit über die Handlungen oder über die Ergebnisse des Kindes in einer emotional negativen Art ausgedrückt (Hirschmann et al., 2013).

Keine Rückmeldung:

Es erfolgt keine Rückmeldung der Bezugsperson.

Rückmeldung – unkodierbar:

Diese Kategorie wird kodiert, wenn die Situation durch das Verlassen des Raums seitens der Mutter oder durch das Betreten einer dritten Person nicht einschätzbar ist. Auch wenn die Mutter mit einer dritten Person spricht oder das Gesagte unverständlich ist wird diese Kategorie gewählt (Hirschmann et al., 2013).

Um die Klassifikation der Kategorie zu treffen müssen auch das nonverbale Verhalten der Bezugsperson sowie die Konnotation der Rückmeldung mit einbezogen werden. Die Quantifizierung erfolgt bei der computergestützten Kodierung mittels Event-Sampling Verfahren. Die einzelnen Kategorien schließen sich gegenseitig aus und decken das zu untersuchende Verhalten vollständig ab. Auf die von Zouzoula (2015) entwickelte händische Kodierung für die Kategorien *Feinfühligkeit* und *Rückmeldung* wird in Kapitel 8.3 näher eingegangen.

7.1.2.3 Joint-Attention

Die Kategorie *Joint-Attention* wird in dieser Untersuchung nicht bearbeitet und findet sich hier nur der Vollständigkeit wegen wieder. Man versteht darunter die Fähigkeit des Kindes, seine Aufmerksamkeit sowohl auf den sozialen Partner als auch auf ein Objekt oder Ereignis zu fokussieren. Durch das Verhalten der Bezugsperson in der Interaktion mit dem Kind kann eine fördernde oder hemmende Wirkung entstehen (Saxon, Colombo, Robinson & Frick, 2000; Claussen, Mundy, Mallik & Willoughby, 2002). Im INTAKT werden neben den Kategorien *Keine Joint-Attention* und *Unkodierbar – Joint Attention* weitere 5 Unterteilungen vorgenommen:

- Aktive Aufrechterhaltung der Joint Attention-Episode auf Handlungsebene
- Aktive Aufrechterhaltung der Joint Attention-Episode auf verbaler Ebene
- Passive Aufrechterhaltung der Joint Attention-Episode
- Aufmerksamkeitslenkung im laufenden Spiel
- Aufmerksamkeitswechsel zu einem neuen/anderen Spiel

Zum exakten Kodiervorgang und einer Beschreibung der einzelnen Kategorien wird auf das Manual von (Hirschmann et al., 2013) verwiesen.

7.1.3 INTERACT

Die Auswertung der am Computer kodierten Videos erfolgt über die INTERACT Software der Firma Mangold. Mit diesem Programm ist es sowohl möglich Live- als auf Videobeobachtungen auszuwerten. Das Programm erlaubt individuelle Beobachtungs- und Kodiermethoden zu definieren und so Videosequenzen am Computer zu kodieren und auszuwerten. Es lassen sich eigene Codes erstellen und diese können wiederum einzelnen Tasten zugeordnet werden.

Im INTERACT kann mittels Event-Sampling-Verfahren aber auch mit Time-Sampling-Verfahren oder einer Kombination aus beidem gearbeitet werden. Beim Time-Sampling-Verfahren besteht die Möglichkeit das gesamte Video in gleich lange Sequenzen zu teilen und diese einzeln abzuspielen bzw. zu kodieren. Die exakte Zeitdauer von vorher definierten Verhaltensweisen kann per Knopfdruck durch einen Anfangs- sowie Endzeitpunkt bestimmt werden.

Der Timecode wird in einem sehr genauen Format ausgegeben und gespeichert: Stunden:Minuten:Sekunden:Bilder pro Sekunde.

Natürlich besitzt die INTERACT Software auch jede beliebige Funktion einer üblichen Videoabspielsoftware wie Anpassung der Abspielgeschwindigkeit, pausieren und vor und zurückspielen des Videos.

8 Untersuchungsdesign

Im Folgenden wird der Aufbau der Untersuchung erläutert und Erklärungen zur händischen Adaption von Zouzoula (2015) gegeben.

8.1 INTAKT: Einschulung und Probekodierung

Vor der Durchführung der Untersuchung erhielt der Beobachter eine ausführliche Einschulung zum INTAKT (Hirschmann et al., 2013). Unter Anleitung von Mag. Dr. Hirschmann wurde das Auswertungssystem INTERACT der Firma Mangold in der 9 Version sowie die Dimensionen *Feinfühligkeit*, *Rückmeldung* und *Joint-Attention* erklärt. Die Probekodierung fand anhand von 15-Minuten Videos am Computer statt. Es wurde eine BeobachterInnenübereinstimmung von $\kappa = .70$ mit Mag. Dr. Hirschmann als unterster Schwellenwert festgelegt. Nach Rücksprache und Beseitigung von etwaigen Unklarheiten wurde so lange kodiert bis der Mindestübereinstimmungswert erzielt wurde.

8.2 Stichprobe

Die für diese Arbeit herangezogenen Videos bestehen aus 13 Mutter-Kind-Dyade welche bei den bereits im Kapitel 7.1.1 näher erläuterten Spiel- und Baselsituationen gefilmt worden waren. Die Kameraperspektive ist statisch und wird während der gesamten Beobachtungssituation nicht verändert.

Die Videos wurden im Zeitraum von Jänner 2012 bis Mai 2013 von Huber (2013), Reischer (2013) und Vogler (2015) aufgenommen sowie mittels INTERACT kodiert. Das längste Video hat eine Dauer von 63,16 Minuten, das kürzeste Video beträgt 27,05 Minuten ($M = 44,62$; $SD = 11,43$). Alle Mütter gaben ihre schriftliche Einverständnis zur Videoaufzeichnung und es wurden zu jeder Mutter und zu jedem Kind soziodemographische Variablen erhoben.

Das jüngste Kind zum Beobachtungszeitpunkt war 41 Monate alt und das Älteste 69 Monate. Das Durchschnittsalter der beobachteten Kinder lag bei 58

Monaten ($SD = 7,48$). Eine relativ ausgeglichene Aufteilung ergibt sich beim Geschlecht: 46,2 Prozent der Kinder waren weiblich, 53,8 Prozent männlich.

Das Alter der Mütter bewegte sich zwischen 27 und 50 Jahren mit einem Mittelwert von 34,69 Jahren ($SD = 6,96$). Der am häufigsten abgeschlossene Bildungsweg war die Fachschule/Lehre mit 46,2 Prozent gefolgt von Matura und Hochschulabschluss (jeweils 23,1 Prozent) und dem Pflichtschulabschluss mit 7,7 Prozent. Weiters waren 69,2 Prozent der Mütter in den Videos der Stichprobe berufstätig (30,8 Prozent waren nicht berufstätig).

Der Großteil der beobachteten Mutter-Kind-Dyaden stammt aus Niederösterreich (69,2 Prozent), das zweithäufigste Bundesland ist Wien mit 23,1 Prozent gefolgt vom Burgenland mit 7,7 Prozent.

8.3 Händische Kodierung: Adaption und Ablauf

Zum besseren Verständnis des Prozesses der händischen Kodierung für die Beobachtungskategorien *Feinfühligkeit* und *Rückmeldung* wird nun eine Abriss des Vorganges dargelegt. Zu einer detaillierteren Erörterung wird auf die Arbeit von Zouzoula (2015) verwiesen, aus welcher das Prozedere adaptiert wurde.

8.3.1 Skala Feinfühligkeit

Die Erfassung der INTAKT-Beobachtungskategorie *Feinfühligkeit* wie in Kapitel 7.1.2.1 beschrieben kann bei der Durchführung der Kodierung mittels Papier-Bleistift-Verfahren genau so gehandhabt werden wie bei der computergestützten Variante. Die siebenstufige Ratingskala wurde somit beibehalten, ebenso das Time-Sampling Verfahren mit den 2-Minuten-Intervallen.

8.3.2 Skala Rückmeldung

Mit Hilfe des Computerprogramms INTERACT ist eine Erfassung der Häufigkeit aber auch der Dauer der einzelnen Rückmeldungskategorien – *positive*, *negative*, *korrigierende*, *keine* sowie *unkodierbare Rückmeldung* – möglich. Die Kategorien

schließen sich gegenseitig aus und somit wird das gesamte Verhalten erfasst. Wie Zouzoula (2015) schon in ihrer Arbeit schreibt, wäre das Festhalten an diesem Kodierschema bei der händischen Kodierung nicht praktikabel und nicht im Sinne des Gütekriteriums der Ökonomie.

Aufgrund dieser Umstände wird bei der händischen Kodierung auf ein Index-System zurückgegriffen, jedoch weiterhin mittels Time-Sampling-Verfahren kodiert. Durch das Index-System können Intervalle gebildet werden (15 Sekunden) in welchen die Häufigkeit des Auftretens einer der Rückmeldungskategorien erfasst wird. Dadurch fällt zwar die Erfassung der genauen Dauer der Rückmeldung weg, es ist jedoch möglich in einem 15-Sekunden-Intervall mehrere verschiedene Rückmeldungen zu kodieren. Durch die Anwendung des Index-Systems sind auch kodierfreie Phasen möglich und somit wird, auch auf Grund der nicht umsetzbaren Adaption der Kategorie *Keine Rückmeldung*, bei der händischen Kodierung eben diese Kategorie nicht berücksichtigt.

Zouzoula (2015, S. 53) hat basierend auf ihren Überlegungen folgende Kodierregeln aufgestellt:

- (a) Die Kodierung der einzelnen Rückmeldungskategorien ist nach deren Anfangssekunde und nicht nach deren Schlusssekunde auf dem Beobachtungsbogen zu protokollieren.
- (b) Die Rückmeldungskategorien *positiv*, *korrigierend* und *negativ* sind anhand ihrer Auftretshäufigkeit erneut zu kodieren, wenn die Rückmeldung einen Verhaltensaspekt, eine Handlung oder Leistung des Kindes betrifft, die sich vom zuvor kodierten Aspekt unterscheidet. Erstreckt sich eine Rückmeldung, die ein und dieselbe Verhaltensweise oder Leistung des Kindes betrifft, ohne zeitliche Unterbrechung über ein 15-Sekunden-Intervall hinaus, ist dies nicht erneut zu kodieren.
- (c) Bei einem über ein 15-Sekunden-Intervall hinausgehenden Auftreten der Kategorie *Rückmeldung* – *unkodierbar* ist dies pro 15-Sekunden-Intervall jeweils erneut zu kodieren. Treten innerhalb eines 15-Sekunden-Intervalls

mehrere *unkodierbare Rückmeldungen* – beispielsweise Bezugsperson nicht sichtbar sowie akustisch unverständliche Aussagen der Bezugsperson – auf, sind diese anhand ihrer Auftrittshäufigkeiten zu kodieren.

- (d) Aufeinanderfolgende Wortwiederholungen innerhalb einer Rückmeldungs-kategorie, beispielsweise „Super! Super!“ – sind als eine einzige Rückmel-dung zu kodieren.

Es sei noch erwähnt dass die 15-Sekunden-Intervalle als Orientierungshilfe auf dem Protokollbogen (siehe Anhang IV) gedacht sind und es sich trotz der zeitlichen Unterteilung um ein Event-Sampling-Verfahren handelt.

8.3.3 Ablauf

Nach erfolgreicher Absolvierung der Einschulung und Probekodierung wurden die vorhandenen Videos aus der Stichprobe mittels Papier-Bleistift-Verfahren und zur Hilfenahme der Protokollbögen von Zouzoula (2015) kodiert. Bei jedem Video wurde zuerst die Kategorie *Feinfühligkeit* und im Anschluss die Kategorie *Rückmeldung* eingeschätzt.

Das Programm VLC-Media-Player ® (Version 2.2.4) diene als Videoplayer für die Videoaufzeichnungen. Für die Kategorie *Feinfühligkeit* wurde nach Beendigung jedes 2-Minuten-Intervalls der Abspielmodus gestoppt und eine Bewertung durchgeführt. Dieser Ablauf wurde bis zum Ende des Videos beibehalten. Bei der Kategorie *Rückmeldung* wurde beim Auftreten einer zu kodierenden Rückmeldung ebenfalls gestoppt und mit Hilfe der Zeitleiste eine Kodierung in dem entsprechenden 15-Sekunden-Intervall des Protokollbogens getätigt.

8.4 Retest-Reliabilität

Das Retest-Reliabilitätsmaß gibt an, inwiefern derselbe/dieselbe BeobachterIn bei wiederholten Messungen zu den gleichen Ergebnissen kommt. Das Maß wurde in dieser Arbeit errechnet um Aufschluss darüber zu erhalten inwiefern der Beobachter bei der händischen Kodierung die zu bewertenden Skalen *Feinfühligkeit*

und *Rückmeldung* stabil oder konsistent gebraucht. Hierzu wurden 20 Prozent der Videos aus der Stichprobe nochmals kodiert und die Summenwerte zum ersten und zum zweiten Zeitpunkt der Kodierung miteinander korreliert. Die Ergebnisse für jede Kategorie der Rückmeldung pro Video sowie für die Skala Feinfühligkeit pro Video und auch als Gesamtwert werden in Tabelle 1 gezeigt.

Tabelle 1

Retest-Reliabilität ($Rel_{Retest-Methode}$) für die Skalen Feinfühligkeit (FF) und Rückmeldung (RM) händisch

Videocode	positive RM	korr. RM	negative RM	unko. RM	FF
	Rel_{Retest}	Rel_{Retest}	Rel_{Retest}	Rel_{Retest}	Rel_{Retest}
JV17	.893	.893	.952	.952	.931
NR17	.771	.865	.901	1.000	.821
PH13	.803	.575	1.000	.610	.861
Gesamt	.802	.802	.934	.830	.885

Unabhängig von der Rückmeldungskategorie konnte bei 80% der Videos eine $Rel_{Retest} > .800$ erzielt werden. Die zwei höchsten Wert betragen 1.000 bei der *negativen Rückmeldung* in Video PH13 sowie der *unkodierbaren Rückmeldung* bei Video NR17. Mit .575 bei der *korrigierenden Rückmeldung* in Video PH13 wurde der niedrigste Wert erzielt. Bei den Werten über alle Videos hinweg ist keine $Rel_{Retest} < .800$.

Wirtz und Caspar (2002) weisen jedoch darauf hin, dass es bei dieser Art von Reliabilität zu Problemen kommen kann. So können verschiedene BeobachterInnen in ihrem eigenen Urteil über eine Skala bei wiederholten Messungen konsistent sein, jedoch muss in diesem Fall keine Reliabilität zwischen verschiedenen BeobachterInnen bestehen. Das bedeutet man benötigt Koeffizienten, die sich auf die Reliabilität zwischen unterschiedlichen BeobachterInnen beziehen. Da in dieser Arbeit ein Vergleich von händischer und computergestützter Kodierung statt-

finden soll und die jeweilige Kodierart von unterschiedlichen BeobachterInnen durchführt wurde, soll in Kapitel 10 die in dieser Arbeit verwendeten Koeffizienten erläutert werden.

9 Koeffizienten zur Überprüfung der BeobachterInnenübereinstimmung

9.1 Intraklassenkorrelation

Die Intraklassenkorrelation (ICC) dient als die am häufigsten verwendete Methode zur Reliabilitätsbestimmung bei intervallskalierten Beurteilungsdaten und basiert auf der klassischen Testtheorie (Wirtz & Caspar, 2002).

Die Beobachtungskategorie *Feinfühligkeit* wird mittels siebenstufiger Ratingskala mit detailliert ausformulierten Ankerpunkten erfasst. Aufgrund dieser theoriegestützten Beurteilungsskala und der ausdifferenzierten Beurteilungshilfen der einzelnen Ratingskalen wird von einem Intervallskalenniveau ausgegangen.

Bei der Beobachtungskategorie *Rückmeldung* werden jeweils die Auftrittshäufigkeiten von *positiver*, *negativer*, *korrigierender Rückmeldung* sowie *Rückmeldung - unkodierbar* in einem 15-Sekunden-Intervall erfasst. Zum Vergleich mit der händischen Kodierung wurden die Daten aus der Computerauswertung auf die Protokollbögen zur händischen Kodierung übertragen. Dadurch war es möglich, jedes 15-Sekunden-Intervall der händischen mit der computergestützten Auswertung zu vergleichen. Wenn man davon ausgeht, dass in einem Intervall die Größe der Differenz der Auftrittshäufigkeiten einer Kategorie der *Rückmeldung* auch Aufschluss über die Qualität der Übereinstimmung gibt dann kann von einem intervallskaliertem Datenniveau ausgegangen werden.

Da mit Hilfe des INTERACT eine framegenaue Kodierung möglich ist konnte der Fall eintreten dass das Auftreten einer Rückmeldung genau an der Grenze eines 15-Sekunden-Intervalls bei der händischen Kodierung erfolgt. So war die Möglichkeit gegeben dass eine Kodierung nicht übereinstimmt obwohl es sich beim Beginn der Rückmeldung nur um einen Unterschied eines Bilds pro Sekunde handelt. Um diesen Umstand entgegen zu wirken wurden für die gesamte Dauer jedes Videos jeweils die 0 bis 15-Sekunden- und 15-30-Sekunden-Intervalle sowie die 30-45-Sekunden- und 45-60-Sekundenintervalle zusammengefasst. Dadurch

wurde eine Minute nicht mehr in 4 Einheiten unterteilt sondern in 2.

Eine Ratingskala kann nur dann als reliabel eingeschätzt werden, wenn verschiedene BeobachterInnen das selbe Objekt ähnlich einschätzen und jeder/jede BeobachterIn unterschiedliche Personen unterschiedlich beurteilen. Varianz kann durch Unterschiede in den Ratings der Ratingskala entstehen (Varianz des Merkmals) aber auch durch alle Störeinflüsse, die einen/eine BeobachterIn dieselbe Person unterschiedlich einschätzen lassen (Varianz unsystematischer Fehler). Sind diese Störfaktoren klein, wird das BeobachterInnenurteil zum großen Teil nur von den gleichen systematischen Unterschieden zwischen den Personen erzeugt was wiederum auf ein reliables Messinstrument schließen lässt (Wirtz & Caspar, 2002).

Die ICC ist ein Maß für die Varianzaufklärung. Dieser Umstand gewinnt bei der Auswahl zur ein- oder zweifaktoriellen Variante an Bedeutung, denn die Unterscheidung folgt daraus, welche Komponenten als Teil der Fehlervarianz definiert werden (Wirtz & Caspar, 2002).

Beim einfaktoriellen Modell ist eine Durchführung auch möglich wenn eine Person von mehreren BeobachterInnen beurteilt wurde, beim zweifaktoriellen Modell müssen alle Personen von dem/der selben BeobachterIn beurteilt werden. Die Voraussetzungen für die Berechnung der einfaktoriellen ICC sind denen der einfaktoriellen Varianzanalyse gleich: Normalverteilung der Daten für die BeobachterInnen und Varianzhomogenität überprüft durch den Levene-Test. Zusätzlich sollte die Anzahl der BeobachterInnen für jede Person zehn nicht unterschreiten.

Weitere Voraussetzung bedarf es bei der zweifaktoriellen ICC: Notwendig ist eine Überprüfung der Interaktionseffekte mittels des Additivitätstest von Tukey da keine systematischen Interaktionseffekte zwischen den Personen und BeobachterInnen auftreten dürfen. Zusätzlich sollten die Trennschärfen analysiert werden, denn besteht hier eine deutliche Inhomogenität zwischen den BeobachterInnen wird diese Voraussetzung verletzt (Wirtz & Caspar, 2002). Falls eine der Bedingungen nicht erfüllt wird sollte beachtet werden dass die Reliabilität systematisch

fehlerhaft geschätzt wird.

Eine weitere Differenzierung gibt durch die Unterscheidung zwischen justierten und unjustierten Modell der ICC. Bei der ICC_{unjust} werden die Mittelwertsunterschiede der einzelnen BeobachterInnen als Fehlerquellen angesehen und fließen somit zu Lasten der Reliabilitätsschätzung in die Berechnung ein. Wohingegen bei der ICC_{just} die Reliabilität vom Effekt unterschiedlicher Mittelwerte der BeobachterInnen bereinigt ist. Eine zusätzliche Unterscheidung ist, dass die ICC_{just} nur berechnet werden kann, wenn alle BeobachterInnen alle Objekte eingeschätzt haben (Wirtz & Caspar, 2002). Die ICC_{unjust} nimmt im Allgemeinen niedrigere Werte an als die ICC_{just} , denn die Mittelwertsunterschiede sind Indikator für mangelnde Reliabilität da sie als Fehlerquellen angesehen werden und somit den Wert der ICC_{unjust} drücken. Da in dieser Arbeit die Mittelwertsunterschiede in die Berechnung mit einfließen sollen wurde stets mit dem unjustierten Modell gerechnet.

Neben der Entscheidung nach justiert/unjustiert muss auch festgelegt werden, ob der/die BeobachterIn als „fixed“ oder als „random“ angesehen wird. Dies ist vor allem für die Interpretation der gewonnen Werte wichtig da beim Faktor „random“ der/die BeobachterIn eine repräsentative Stichprobe aller BeobachterInnen darstellt, wohingegen beim Faktor „fixed“ sich die Reliabilitätseinschätzung auf die untersuchte Stichprobe bezieht und keine Verallgemeinerung vorgenommen werden kann (Wirtz & Caspar, 2002). In dieser Untersuchung wurde der Faktor „random“ gewählt.

Der Wertebereich der ICC liegt definitionsgemäß bei 0 bis 1. Ein Wert von 0 deutet auf keinen Zusammenhang zwischen den BeobachterInnen hin, eine Zunahme des Werts im positiven Bereich lässt auf eine höher Reliabilität schließen. Werte größer .70 können im Allgemeinen als Indikator für eine gute BeurteilerInnenübereinstimmung gesehen werden (Wirtz, 2006). Werte, die sich im negativen Bereich bewegen werden mit 0 gleichgesetzt und würden als 0% Varianzerklärung interpretiert werden (Wirtz & Caspar, 2002). Der Ausprägungsgrad des Koeffizienten sollte jedoch immer in Relation zum messenden Merkmal und der verwen-

deten Stichprobe gesehen werden (Wirtz & Caspar, 2002).

9.2 Kappa-Koeffizient nach Cohen (κ)

Laut Wirtz und Caspar (2002) ist Cohens Kappa das am häufigsten verwendete Maß zur Bestimmung der BeobachterInnenübereinstimmung für mindestens nominalskalierte Daten und basiert auf der prozentualen Übereinstimmung (PÜ). Bei der PÜ wird angegeben, in wie viel Prozent der Fälle ein oder mehrere BeobachterInnen das selbe Urteil abgeben. Der Vorteil von κ gegenüber PÜ ist, dass das Verhältnis der beobachteten zu der bei Zufall erwarteten Übereinstimmung berücksichtigt wird (Wirtz & Caspar, 2002). Man kann also von einer standardisierten Maßzahl sprechen, bei der die Zufallserwartung bei beobachteten Übereinstimmungen korrigiert wird, jedoch werden die genauen Übereinstimmungsinformationen verrechnet.

Aus diesem Grund wurde für diese Untersuchung mit dem gewichteten Kappa-Koeffizienten (κ_w) gerechnet, bei dem die Wertedifferenzen miteinbezogen werden. Dadurch wird nicht nur Rechnung getragen, ob die BeobachterInnen übereinstimmen, sondern auch wie nahe beisammen oder wie weit entfernt voneinander sie sind. Es werden bei der Berechnung relative Gewichte eingeführt. Die Gewichtung kann im Prinzip frei gewählt werden (Faßnacht, 1995), oft wird aber entweder ein linearer oder quadratischer Ansatz gewählt. Bei der quadratischen Gewichtung werden Fehlübereinstimmung stärker gewertet, das bedeutet je mehr Differenz desto weniger Übereinstimmung. Greve und Wentura (1997) sprechen sich ebenfalls für die quadratische Gewichtung aus, da „harmlose“ Fälle von Nicht-Übereinstimmungen weniger stark eingehen, aber große Differenzen stärker.

Ein weiterer Vorteil des gewichteten Kappa-Koeffizient ist der direkte Zusammenhang mit der ICC. Laut Wirtz und Caspar (2002) ist κ bei der quadratischen Gewichtung identisch mit der unjustierten Intraklassenkorrelation weshalb im Falle der Nichterfüllung der Voraussetzungen der ICC der gewichtete Kappa-Koeffizient gewählt wurde.

Bei der Interpretation der Werte werden die postulierten Klassifikationen von

Wirtz und Caspar (2002), Fleiss, Levin und Paik (2003) und Bortz und Döring (2006) herangezogen bzw. adaptiert: ein Koeffizient $\geq .75$ steht für eine ausgezeichnete Übereinstimmung, wohingegen ein Koeffizient $\leq .40$ für eine schlechte Übereinstimmung steht. Eine gute Übereinstimmung wird bei Werten zwischen .60 bis .75 angenommen. So ergibt sich für diese Arbeit:

- $\kappa \leq .40$ schlechte Übereinstimmung
- $\kappa > .40$ bis $\leq .60$ akzeptable Übereinstimmung
- $\kappa \geq .60$ bis $< .75$ gute Übereinstimmung
- $\kappa \geq .75$ ausgezeichnete Übereinstimmung

10 Ergebnisse der Untersuchung

Die Ergebnisse der Untersuchung für die Beobungskategorie *Feinfühligkeit* und *Rückmeldung* werden in diesem Kapitel dargestellt. Ziel war es den Vergleich von händischer und computergestützter Kodierung herzustellen.

10.1 Feinfühligkeit: Vergleich händischer mit computerbasierter Kodierung

Den ersten gewählten Vergleich zwischen der händischen und computerbasierten Kodierung für die Beobungskategorie *Feinfühligkeit* soll die Abbildung 1 geben. Dargestellt ist der Verlauf der Mittelwerte für die händische und die computergestützte Kodierung pro Videoaufzeichnung.

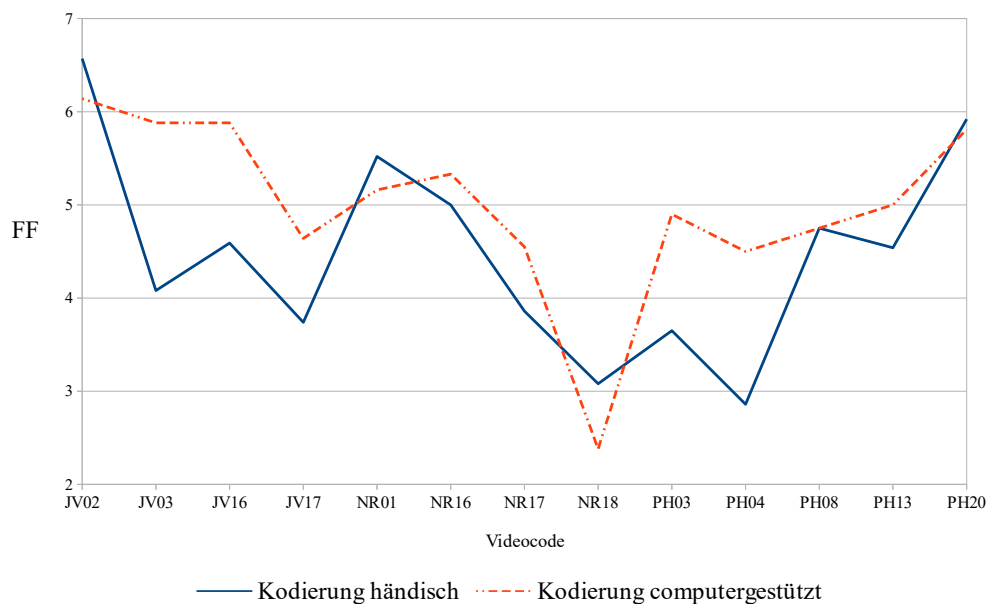


Abbildung 1. Verlauf der Mittelwerte Beobungskategorie *Feinfühligkeit* für händische und computergestützte Kodierung pro Video

Die Mittelwerte und Standardabweichungen für die händische Kodierung sowie die Computerkodierung der Beobachtungskategorie *Feinfühligkeit* für die gesamte Stichprobe (N=13) sowie für jedes Video finden sich in Tabelle 2.

Tabelle 2

Mittelwerte (M) und Standardabweichungen (SD) der Beobachtungskategorie Feinfühligkeit für händische und computergestützte Kodierung per Video und über alle Videoaufzeichnungen (N=13)

Videocode	Computerkodierung		händische Kodierung	
	M	SD	M	SD
(1) JV02	6.14	0.848	6.57	0.634
(2) JV03	5.88	0.612	4.08	0.974
(3) JV16	5.88	0.600	4.59	0.939
(4) JV17	4.63	1.422	3.74	1.147
(5) NR01	5.16	0.454	5.52	0.724
(6) NR16	5.33	0.485	5.00	0.485
(7) NR17	4.55	0.510	3.86	0.640
(8) NR18	2.38	0.506	3.08	0.494
(9) PH03	4.90	0.718	3.65	1.040
(10) PH04	4.50	0.694	2.86	0.705
(11) PH08	4,75	0.608	4.75	0.676
(12) PH13	5.00	0.816	4.54	0.967
(13) PH20	5.81	0.567	5.92	0.628
Gesamt	5.10	1.076	4.58	1.354

Anmerkung. (1) sehr geringe Feinfühligkeit bis (7) sehr hohe Feinfühligkeit.

Der höchste Mittelwert der Beobachtungskategorie *Feinfühligkeit* aller Videos findet sich für die Computerkodierung als auch für die händische Kodierung in Video (1) JV02 mit 6.14 (SD = 0.848) und 6.57 (SD = 0.634). Video (8) NR18 weist den geringsten Mittelwert mit 2.38 (SD = 0.506) für die computergestützte Kodierung aus, für die händische Kodierung ist es das Video mit dem Code (10) PH04 mit 2.86 (SD = 0.705). Der Mittelwert für alle 13 Videos beträgt 5.10 (SD = 1.076) für die Computerkodierung und 4.58 (SD = 1.354) für die händische Kodierung.

In Tabelle 3 finden sich die relativen und absoluten Häufigkeiten für die gesamte Stichprobe mit der Unterteilung in computergestützte und händische Kodierung.

Tabelle 3

Häufigkeiten der Beobachtungskategorie Feinfühligkeit für händische und computergestützte Kodierung über alle 13 Videoaufzeichnungen

Feinfühligkeit	absolute Häufigkeit		relative Häufigkeit	
	händische Kodierung	Computerkodierung	händische Kodierung	Computerkodierung
(1) sehr geringe	– ^a	– ^a	– ^a	– ^a
(2) geringe	17	10	6.0%	3.5%
(3) eher geringe	50	12	17.7%	4.2%
(4) mittlere	67	37	23.7%	13.1%
(5) eher hohe	72	124	24.4%	43.8
(6) hohe	54	82	19.1%	29.0%
(7) sehr hohe	23	18	8.1%	6.4%
(0) unkodierbar	– ^a	– ^a	– ^a	– ^a
Gesamt	283	283	100%	100%

Anmerkung. ^a Kategorien wurden nicht kodiert.

Insgesamt wurden 283 2-Minuten-Einheiten kodiert. Am häufigsten tritt die Kategorie (5) *eher hohe Feinfühligkeit* bei der händischen sowie der Computerkodierung auf, Kategorie (2) *geringe Feinfühligkeit* wurde jeweils am wenigsten häufig kodiert. Die Kategorien *Feinfühligkeit - unkodierbar* und (1) *sehr geringe Feinfühligkeit* wurden von den BeobachterInnen nicht kodiert.

10.1.1 Korrelationskoeffizienten (r_{PM} , r_{IC} , κ_w) Feinfühligkeit

Zur Überprüfung des Ausmaßes der Übereinstimmung von händischer und computerbasierter Kodierung wurden drei verschiedene Korrelationskoeffizienten und deren jeweilige Verfahren herangezogen, der Kappa-Koeffizient nach Cohen (κ), die Produkt-Moment-Korrelation (r_{PM}) und die Intraklassenkorrelation (r_{IC}).

Um eine Aussage über die Übereinstimmungsvalidität der händischen Kodierung zu treffen, wurde die Produkt-Moment-Korrelation (r_{PM}) mit den Summenwerten der händischen und computergestützten Kodierart für die Kategorie *Feinfühligkeit* pro Video errechnet. Die Voraussetzungen zur Berechnung sind gegeben und in diesem Fall wurde als Außenkriterium die computergestützte Kodierung gewählt.

Die Berechnung der Intraklassenkorrelation fand mittels einfaktoriellen Modells statt, unter den in Kapitel 9.1 erläuterten Kriterien. Die Berechnungen für die Voraussetzungen des einfaktoriellen Modells finden sich in Anhang I wieder. Zur Überprüfung der Normalverteilung der Daten wurden die standardisierten Schiefe- und Kurtosiskoeffizienten ($|Z| < 1.96$) pro Video als auch im Gesamtwert berechnet, die Varianzenhomogenität wurde mittels Levene-Test als Gesamtwert über alle Videos erfasst ($F_{1,564}=37.587$, $p=<.001$).

Tabelle 4

Produkt-Moment-Korrelation (r_{PM}) und Intraklassenkorrelation (r_{IC}) für die Kategorie Feinfühligkeit über alle Videos ($N=13$): Vergleich händische und computerkodierte Auswertung

	Korrelationskoeffizient	Signifikanzniveau
r_{PM}	.901	<.001*
r_{IC}	.509	<.001

Anmerkung. *Korrelation ist bei Niveau 0,05 signifikant (einseitig).

Der Korrelationskoeffizient über alle Videoaufzeichnung beträgt $r_{PM}=.901$ ($F = 11$; $p=<.001 < \alpha = .05$) und die BeobachterInnenübereinstimmung zwischen händisch und computergestützter Kodierung über alle 13 Videos beträgt $r_{IC} = .509$ ($F = 3.077$; $p=<.001 < \alpha = .05$) (siehe Tabelle 4). Der wahre Wert des ICC-Koeffizienten befindet sich mit einer 95%iger Sicherheit im Bereich zwischen .418 und .591. Der Grad der Übereinstimmung ist signifikant.

Aufgrund der Strukturierung der Daten konnte eine Aufsplittung der KodiererInnen unter Berücksichtigung der Kodierart durchgeführt werden. In diesem Fall wurden alle Videos, die per Computer von einer einzelnen Kodiererin bewertet wurden, mit denen des Kodierers, welcher die händische Methode durchgeführt hatte, verglichen. Dadurch war die Konstellation für eine zweifaktorielle ICC gegeben (alle BeobachterInnen beurteilten alle Videos).

Zusätzlich zu den Voraussetzung des einfaktoriellen Modells dürfen keine systematischen Interaktionseffekte zwischen BeobachterInnen und Personen vorliegen, dies wurde mittels Nicht-Additivitäts-Test nach Tukey überprüft. Die Ergebnisse finden sich, genau so wie die Überprüfung auf Varianzhomogenität für alle KodiererInnen, in Anhang II wieder.

Da die Voraussetzungen bei Kodiererin JV und PH (siehe Anhang II) nicht gegeben waren, wurde anstelle des zweifaktoriellen Modells der ICC der gewichtete

Kappa-Koeffizient nach Cohen (κ_w) errechnet. Die Ergebnisse sind in Tabelle 5 abgebildet.

Tabelle 5

Intraklassenkorrelation (r_{IC}) und Kappa-Koeffizient nach Cohen (κ) für die Kategorie Feinfühligkeit über alle Videos der Kodierer NR (N=4), JV (N=4) und PH (N=5): Vergleich händische und computerkodierte Auswertung

Kodiererin	Intraklassenkorrelation (r_{IC})	Kappa-Koeffizient (κ)
NR	.773	.730
JV	_ ^a	.416
PH	_ ^a	.464

Anmerkung. ^a Voraussetzungen für Berechnung nicht gegeben.

Die BeobachterInnenübereinstimmung von Kodiererin NR mit der händischen Kodierung beträgt $r_{IC} = .773$ ($F = 6.430$; $p < .001 < \alpha = .05$). Der signifikante Koeffizient befindet sich mit 95%iger Sicherheit im Bereich zwischen .616 und .819. Für die Kodiererin JV ergab sich ein gewichteter Kappa-Wert von .416. Der Vergleich der händischen Kodierung mit den Ergebnissen der Kodiererin PH ergab $\kappa_w = .464$.

10.2 Rückmeldung: Vergleich händischer mit computerbasierter Kodierung

In diesem Kapitel soll ein Vergleich zwischen der händisch und computerbasierten Kodierung der Kategorie Rückmeldung gezogen werden, beginnend mit den absoluten Auftrittshäufigkeiten. In Tabelle 6 sind die aufsummierten Auftrittshäufigkeiten pro 15 Sekunden-Intervall unabhängig von den erfassten Rückmeldungskategorien (*positiv*, *korrigierend*, *negativ* und *unkodierbar*) angegeben, aufgeschlüsselt für die händischen und die computergestützte Kodierung.

Tabelle 6

Auftrittshäufigkeiten und Anteilswerte der Beobachtungskategorie Rückmeldung ohne Ausprägung unterteilt in händische und computergestützte Kodierung über alle Videoaufzeichnungen (N=13)

Rückmeldung	absolute Häufigkeit		relative Häufigkeit	
	händische Kodierung	Computerkodierung	händische Kodierung	Computerkodierung
0	2029	1963	86.82%	84%
1	288	345	12.32%	14.76%
2	20	27	0.86%	1.15%
3	- ^a	2	- ^a	0.09%
Summe	2337	2337	100%	100%

Anmerkung. ^a Häufigkeit kam nicht vor.

2337 Beobachtungseinheiten zu je 15 Sekunden-Intervallen.

Am Häufigsten für die händischen und die Computerkodierung mit 86,82% bzw. 84% wurde keine *Rückmeldung* kodiert. Wenn eine *Rückmeldung* in einem 15-Sekunden-Intervall vermerkt wurde, dann am häufigsten nur eine Rückmeldungskategorie. Die relativen Häufigkeiten betragen hier bei der händischen Kodierung 12,32% und bei der Computerkodierung 14,76%. Drei Kodierung in ei-

nem Intervall wurden lediglich zwei mal während der Computerkodierung vermerkt.

Die Häufigkeiten für die einzelnen Kategorien der Skala *Rückmeldung* pro 15-Sekunden Intervall werden in Tabelle 7 dargestellt.

Tabelle 7

Häufigkeiten der Beobachtungskategorie Rückmeldung positiv, korrigierend, negativ und unkodierbar der händische und computergestützte Kodierung über alle Videoaufzeichnungen (N=13)

Rückmeldung		absolute Häufigkeit		relative Häufigkeit	
		händische Kodierung	Computer-kodierung	händische Kodierung	Computer-kodierung
positiv	0	2231	2170	95.5%	92.9%
	1	91	146	3.9%	6.2%
	2	15	19	0.6%	0.8%
	3	- ^a	2	- ^a	0.1%
	Summe	2337	2337	100%	100%
korrigierend	0	2278	2287	97.5%	97.9%
	1	59	49	2.5%	2.1%
	2	- ^a	1	- ^a	0.0%
	Summe	2337	2337	100%	100%
negativ	0	2290	2265	98%	96.9%
	1	46	67	2%	2.9%
	2	1	5	0.0%	0.2%
	Summe	2337	2337	100%	100%

unkodierbar	0	2241	2252	95.9%	96.4%
	1	92	83	3.9%	3.6%
	2	4	2	0.2%	0.1%
	Summe	2337	2337	100%	100%
insgesamt	keine RM	2029	1963	86.82%	84%
	RM	308	374	13.18%	16%
	Summe	2337	2337	100%	100%

Anmerkung. ^a Häufigkeit kam nicht vor.

In allen 4 Kategorien der händischen und Computerkodierung (*positiv, korrigierend, negativ* und *unkodierbar*) überwiegt die Anzahl an Intervallen in denen keine Kodierung stattgefunden hat (>92%). Ohne Berücksichtigung der Auftrittshäufigkeit innerhalb eines 15-Sekunden-Intervalls wurden 4,5% *positive Rückmeldungen* bei der händischen Kodierung und 7,1% bei der Computerkodierung vermerkt. Bei 2,5% der 15-Sekunden-Intervallen wurde eine *korrigierende Rückmeldung* mittels händischer Kodierung kodiert, bei der computergestützten sind es 2,1%. Auf die *negative Rückmeldung* entfallen bei der händischen Kodierung 2% und bei der Computerkodierung 3,1%, wohingegen 4,1% und 3,7% der *Rückmeldung - unkodierbar* zugeteilt werden (händische respektive computergestützte Kodierung).

Der relative Anteil aller kodierter Kategorien der Beobachtungskategorie Rückmeldung über alle 15-Sekunden-Intervalle hinweg, aufgeschlüsselt nach händischer und computergestützter Kodierung wird in Abbildung 2 dargestellt.

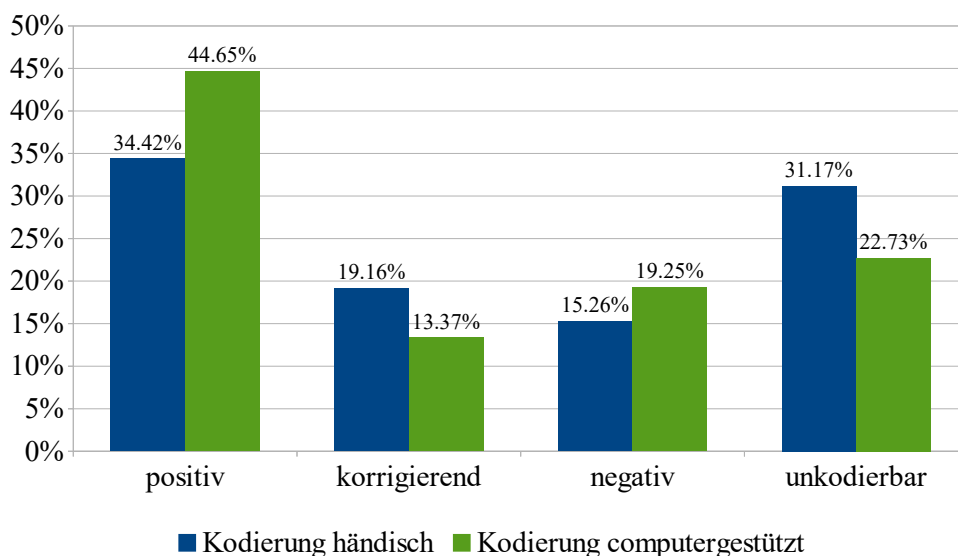


Abbildung 2. Prozentanzahl der tatsächlich kodierten Rückmeldungen nach Kategorie und Kodierart

Am Häufigsten wurde die Kategorie *positive Rückmeldung* kodiert, sowohl bei der händischen (34.42%) als auch bei der computergestützten Kodierung (44.65%). Bei dieser Kategorie herrscht ebenso die größte Differenz zwischen den Kodierarten (10.22%). Die zweithöchste Abweichung findet sich bei der Kategorie *Rückmeldung - unkodierbar* mit 8.44%. Die Differenz bei den Kategorien *korrigierend* und *negativ Rückmeldung* sind geringer (5.79% und 3.99%), jedoch wurde bei der händischen Kodierung häufiger *korrigierende Rückmeldungen* (19.16% zu 13.37%) und bei computergestützten häufiger *negative Rückmeldungen* (15.26% zu 19.25%) kodiert.

10.2.1 Korrelationskoeffizienten (r_{PM} , κ_w) Rückmeldung

Wie auch bereits bei der Beobachtungskategorie *Feinfühligkeit* wird bei der Beobachtungskategorie *Rückmeldung* zum Zweck der Überprüfung der Übereinstimmungsvalidität der händischen Kodierung die Produkt-Moment-Korrelation (r_{PM}) mit den Summenwerten der händischen und computergestützten Kodierart für alle Kategorien der Beobachtungskategorie *Rückmeldung* errechnet. Die Ergebnisse finden sich in Tabelle 8 genau so wieder wie die berechneten κ_w -Koeffizienten. Das einfaktorielle Modell der ICC konnte aufgrund der nicht erfüllten Voraussetzungen bei dem Vergleich der 30-Sekunden-Intervalle (siehe Anhang V und VI) nicht berechnet werden.

Wie bereits in Kapitel 9 beschrieben wurden zur geplanten Berechnung der ICC für die Kategorie *Rückmeldung* nicht die wie auf dem Protokollbogen (Anhang IV) angegeben Intervallsunterteilung verglichen, sondern immer jeweils zwei 15-Sekundenintervalle zusammengefasst. Durch die ungerade Anzahl der Intervalle wurde das letzte 15-Sekunden-Intervall bei der Aufsummierung ausgeschlossen und fließt somit auch nicht in die Berechnungen mit ein.

Tabelle 8
Korrelationskoeffizienten (κ_w , r_{PM}) für alle Rückmeldungs-kategorien

Rückmeldung	κ_w	r_{PM}
positiv	.699	.674
korrigierend	.394	.606
negativ	.540	.703
unkodierbar	.390	.726

Anmerkung. Korrelation ist bei Niveau 0,05 signifikant (einseitig).

Die höchste Korrelation wurde bei der Kategorie *Rückmeldung - unkodierbar* mit $r_{PM}=.726$ erzielt, gefolgt von der *negative Rückmeldung* mit $r_{PM}=.703$. In der Kategorie *positive Rückmeldung* beträgt der Korrelationskoeffizient $.606$ und der niedrigste Wert wurde bei der *korrigierenden Rückmeldung* mit $.606$ erzielt. Der Korrelationskoeffizient über alle Rückmeldungskategorien beträgt $.699$. Für alle Kategorien der Beobachtungskategorie *Rückmeldung* sind die Ergebnisse signifikant.

Der errechnete gewichtete κ_w -Koeffizienten über alle 13 Videoaufzeichnung für die 30-Sekunden-Intervalle betragen für die Kategorie *positive Rückmeldung* $.699$, für die *korrigierende Rückmeldung* $.394$, für die *negative Rückmeldung* $.540$ und für die Kategorie *Rückmeldung - unkodierbar* $.390$.

10.2.2 Kappa-Koeffizient nach Cohen (κ) pro Video

Zusätzlich zum Vergleich der händischen und computergestützten Kodierung pro Rückmeldungskategorie über alle 30-Sekunden-Einheiten aller Videos hinweg wurde für jedes Video einzeln ein Vergleich angestellt. In Abbildung 3 sind die κ_w -Koeffizienten je Videoaufzeichnung für die *positive* und *korrigierende Rückmeldung* für die 30-Sekunden-Intervalle dargestellt.

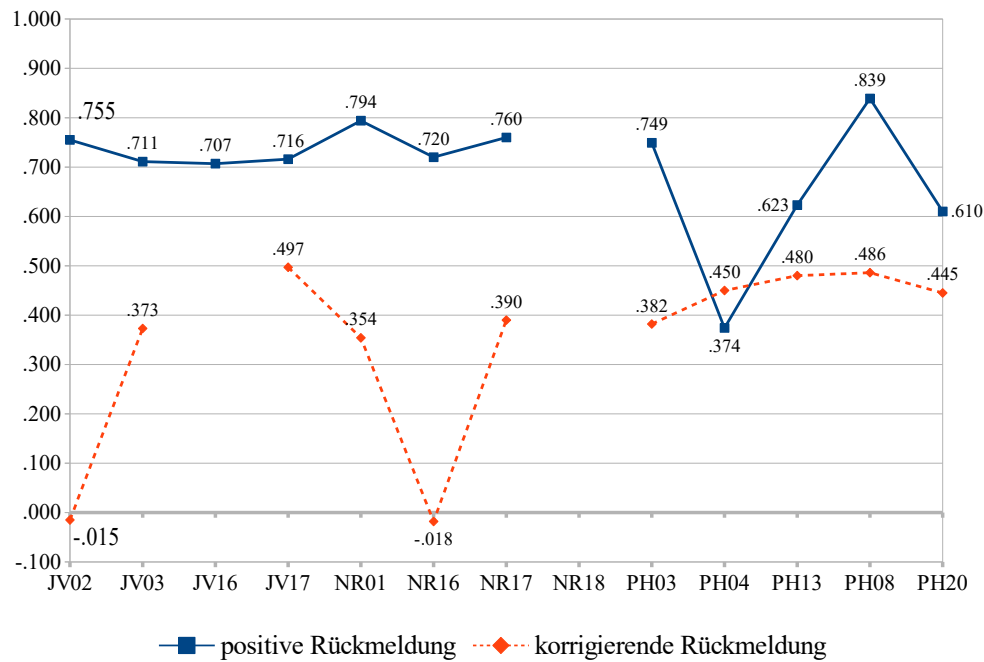


Abbildung 3. κ_w -Koeffizienten je Videoaufzeichnung für positive und korrigierende Rückmeldung für die 30-Sekunden-Intervalle

Bei 75% der κ_w -Koeffizienten der *positiven Rückmeldung* wurde ein Wert $>.710$ erreicht, der niedrigste Wert dieser Kategorie beträgt $.374$ und der höchste $.839$. Keine Berechnung fand bei der Videoaufzeichnung NR18 aufgrund von durchgehend konstant ausgefallener Beobachtungen auf Seiten der händischen Kodierung statt. Der niedrigste κ_w -Koeffizienten der Kategorie *korrigierende Rückmeldung* beträgt $-.018$ und der höchste $.497$. Es wurde bei 45,5% ein Wert von $>.445$ erzielt. Für die Videos NR16 und NR18 konnten keine Werte berechnet werden da bei der computergestützten Kodierung durchgehen konstant bewertet wurde.

In Abbildung 4 finden sich die κ_w -Koeffizienten für die *negative Rückmeldung* und die *Rückmeldung - unkodierbar* pro Videoaufzeichnung.

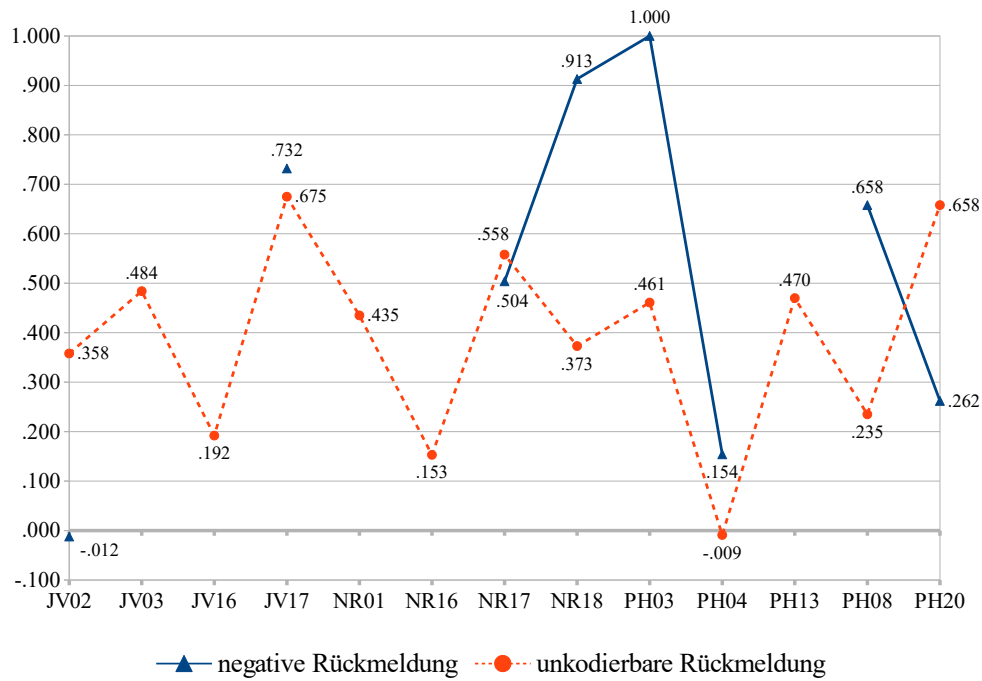


Abbildung 4. κ_w -Koeffizienten je Videoaufnahme für negative und unkodierbare Rückmeldung für die 30-Sekunden-Intervalle

Der höchste Wert wird mit 1.0 beziffert, der niedrigste mit -.012. Bei vier Videoaufzeichnungen (JV03, JV16, NR01, NR16, PH13) konnte keine Berechnung des Koeffizienten durchgeführt werden da aufgrund von konstanter Kodierung bei einer der beiden Kodierarten keine Varianz vorhanden war. Die κ_w -Koeffizienten der Kategorie *Rückmeldung - unkodierbar* befinden sich zwischen .675 und -.009.

Weiters wurde die Beobachtungskategorie *Rückmeldung* ohne die Differenzierung der einzelnen Kategorien *positiv*, *korrigierend*, *negativ* und *unkodierbar* über alle Videoaufzeichnungen unter der Bedingung der Kodierart (händisch oder computergestützt) verglichen. Der κ_w -Koeffizienten beträgt .647.

11 Diskussion

Diese Arbeit soll eine Überprüfung des händischen Kodierschemas (Zouzoula, 2015) für das Beobachtungssystem INTAKT (Hirschmann et al., 2013) in Bezug auf die Vergleichbarkeit der Ergebnisse aus der computergestützten Kodierung darstellen. Mittels 13 Videoaufzeichnung von Mutter-Kind-Interaktionen, computergestützt kodiert durch 3 Beobachterinnen und händisch kodiert durch den Verfasser dieser Arbeit, wurden die INTAKT-Beobachtungskategorien *Feinfühligkeit* und *Rückmeldung* auf etwaige Unterschiede überprüft.

Die INTAKT-Kategorie *Feinfühligkeit* wurde mittels siebenstufiger Rating-skala im Time-Sampling-Verfahren mit 2-Minuten-Intervallen jeweils für die händische und computergestützte Kodierung ermittelt.

Die INTAKT-Beobachtungskategorie *Rückmeldung* wurde bei der händischen Kodierung als auch bei der Computerkodierung mittels Event-Sampling-Verfahren erfasst. Der Unterschied bestand darin dass bei der händischen Kodierung ein Index-System verwendet wurde welches das mehrfache Auftreten einer Rückmeldungskategorie in einem 15-Sekunden-Intervall erlaubt. Zusätzlich zu dieser Unterscheidung wurde bei der händischen Kodierung nur zwischen *positiver*, *korrigierender*, *negativer* und *unkodierbarer Rückmeldung* differenziert, die Rückmeldungskategorie *keine Rückmeldung* aus der computergestützten Kodierung wurde nicht berücksichtigt.

Die Frage nach der Vergleichbarkeit der unterschiedlichen Kodiermethoden wurde durch die Berechnung der Korrelation zwischen der händischen und computerbasierten Auswertung versucht zu beantworten. Dadurch sollten Aufschlüsse über die Kriteriumsvalidität bzw. genauer die Übereinstimmungsvalidität gegeben werden. Die Intraklassenkorrelation (ICC) bzw. der gewichtete Kappa-Koeffizienten nach Cohen (κ_w) sollte Erkenntnisse zur BeobachterInnenübereinstimmung hinsichtlich der händischen und computergestützten Kodierart geben. Beim Kappa-Koeffizienten wurde eine quadratische Gewichtung gewählt, da diese nach Wirtz und Caspar (2002) identisch mit der unjustierten Intraklassenkorrelation ist.

Folgend werden die Ergebnisse nach diesen beiden Kriterien interpretiert.

11.1 Feinfühligkeit

11.1.1 Übereinstimmungsvalidität

Die Antwort auf die Frage nach der Übereinstimmungsvalidität der INTAKT-Beobachtungskategorie *Feinfühligkeit* lässt sich in Ansätzen bereits durch einen Blick auf die deskriptive Statistik beantworten. Wenn man den Verlauf der Mittelwerte in Abbildung 1 und auch die Gesamtwerte für die Mittelwerte in Tabelle 2 mit 4.58 (SD=1.354) und 5.10 (SD=1.076) für die händische und computergestützte Kodierung respektive betrachtet, so sind die Werte der jeweiligen Kodierarten recht nahe beieinander.

Betrachtet man die absoluten Häufigkeiten der einzelnen Kategorien der Beobachtungskategorie *Feinfühligkeit* so fällt auf dass insgesamt nahezu gleich oft *mittlere* und *hohe Feinfühligkeit* bei der händischen Kodierung vergeben wurde (23,7% und 24,4%) wohingegen bei der computergestützten Variante klar die Kategorie *eher hohe Feinfühligkeit* mit 43,8% vorne liegt. Diese mittige Verteilung der vergebenen Feinfühligkeitswerte legt einen zentrale Tendenzen Effekt, wie in Kapitel 3.5.1 beschrieben, bei der händischen Kodierung nahe.

Trotz dieses Fehlers zu Lasten des/der BeobachterIn kann durch den signifikanten Korrelationskoeffizienten der Produkt-Moment-Korrelation von $r_{PM}=.901$ auf eine ausgezeichnete Übereinstimmungsvalidität in der INTAKT-Beobachtungskategorie *Feinfühligkeit* für die händische Kodierung geschlossen werden.

11.1.2 BeobachterInnenübereinstimmung

Nach Betrachtung der Übereinstimmung der Gesamtwerte der Beobachtungskategorie *Feinfühligkeit* bzw. der Übereinstimmungsvalidität zwischen händischer und computergestützter Kodierung, wurde durch den Einsatz der Intraklassenkorrelation die Quantifizierung der Übereinstimmung zwischen den händischen und computergestützten KodiererInnen angestrebt. Mit einem signifikanten ICC-Koeffizienten von $r_{IC}=.509$ für alle 13 Videos wurde ein akzeptabler Wert erzielt, auch in Anbetracht der GröÙer der vorhandenen Stichprobe. Bei genauerer Be-

trachtung und dem Vergleich der einzelnen Kodiererinnen mit dem händischen Kodierer konnte ein ausgezeichneter Wert bei Kodiererin NR mit $r_{IC}=.773$ erzielt werden. Die gewichteten Kappa-Werte bei Kodiererin JV und PH waren mit .416 und .464 deutlich darunter, können aber trotzdem noch als akzeptabel eingestuft werden (zur Wahl der Koeffizienten siehe Kapitel 12.1.3 und 12.1.4). In Anbetracht der Vergleichbarkeit von κ_w mit r_{IC} (Wirtz & Caspar, 2002) und der hohen Strenge der ICC gegenüber anderen Verfahren wie der Produkt-Moment-Korrelation kann bei den oben genannten Werten für die Beobachtungskategorie *Feinfühligkeit* mit denen für die Arbeit definierten Klassifikationen zur Interpretation der Ergebnisse von einer akzeptablen BeobachterInnenübereinstimmung gesprochen werden.

11.2 Rückmeldung

11.2.1 Übereinstimmungsvalidität

Ein interessantes Ergebnis aus Tabelle 6 findet sich bei den Auftrittshäufigkeit der Computerkodierung: im Gegensatz zur händischen Kodierung gab es 2.82% mehr Kodierung ungeachtet der Rückmeldungskategorie. Liegt die Computerkodierung bei der Anzahl der Kodierungen pro 15-Sekunden-Intervall im Bereich der einmaligen Vermerkungen noch mit 2.44% vorne, so vermindert sich der Wert bei den zweifach Nennungen pro Intervall auf 0.29%.

Dieser Umstand kann durch einen Blick auf die Tabelle 7 näher beleuchtet werden. Bei der Aufschlüsselung der Auftrittshäufigkeiten nach Beobachtungskategorie wird ersichtlich dass die *positive* und *negative Rückmeldung* häufiger bei der computergestützten als bei der händischen Kodierung vermerkt wurden. Auch hier gilt wieder dass der Unterschied bei der einmaligen Vermerkung pro 15-Sekunden-Intervall größer ist als bei Doppel- oder Dreifachnennungen. Für die Beobachtungskategorien *korrigierende Rückmeldung* und *Rückmeldung - unkodierbar* trifft das Gegenteil zu: in beiden Fällen kam es bei der händischen Kodierung zu insgesamt mehr Kodierungen, wenn auch nur in einem sehr geringen

Bereich. Eine graphische Darstellung dieser Ergebnisse kann man in Abbildung 2 begutachten.

Ein möglicher Erklärungsansatz für diese Verteilungen könnte eine genauere Betrachtung der Korrelationswerte der Beobachtungskategorie *Rückmeldung* geben. Der Korrelationskoeffizient der *korrigierenden Rückmeldung* ist mit einem Wert von .606 der niedrigster aller 4 Kategorien. Zusammen mit den deutlich geringeren, tatsächlich kodierten händischen *korrigierenden Rückmeldungen* liegt die Vermutung nahe dass bei dieser Kategorie eventuell ein Validierungsproblem vorliegt und genauere Instruktionen und Beschreibungen zum Kodiervorgang von Nöten sind. Die Tatsache dass auch bei der Kategorie *Rückmeldung - unkodierbar* öfter bei der händischen Kodierung vermerkt wurde kann durch die Unterschiede in der Wahrnehmung der KodiererInnen bei der Schwere von Störungen durch äußere Bedingungen herrühren. Nichtsdestotrotz bleibt die erhöhte Häufigkeit bei den Kategorien *positive* und *negative Rückmeldung* für die Computerkodierung bzw. auch eine insgesamt häufigeren Kodierung von Rückmeldungen von 2.82% gegenüber der händischen Kodierung. Dies kann auch in einer höheren Genauigkeit der computergestützten Kodierung begründet sein.

Wie bereits angesprochen ist abgesehen von der *korrigierenden Rückmeldung* der lineare Zusammenhang zwischen der händischen und computergestützten Kodierung zur Übereinstimmungsvalidität auf einem guten (*positive Rückmeldung*: $r_{PM}=.674$) bis ausgezeichneten (*negative bzw. Rückmeldung - unkodierbar*: $r_{PM}=.703$ und $r_{PM}=.726$) Niveau.

11.2.2 BeobachterInnenübereinstimmung

Beim Vergleich der händischen und computerkodierten Ergebnisse in Bezug auf die BeobachterInnenübereinstimmung ist der erzielte gute Kappa-Wert bei der Beobachtungskategorie *positive Rückmeldung* mit .699 der höchste aller Kategorien. Auch bei der Aufschlüsselung der Kappa-Werte nach den jeweiligen Videos in Abbildung 3 kann die gute Übereinstimmung bestätigt werden. Jeder Kappa-Wert befindet sich über .610, mit Ausnahme von Video PH04.

Der zweithöchste Wert zur BeobachterInneneinstimmung liegt mit .540 in einem akzeptablen Bereich und konnte bei der *negativen Rückmeldung* erzielt werden. Bei der Betrachtung der κ_w -Koeffizienten je Videoaufzeichnung in Abbildung 4 fällt natürlich auf, dass mehrere Videos aus der Berechnung ausgeschlossen wurden. Grund hierfür sind die fehlenden Kodierungen auf Seiten der computergestützten (JV03, JV16, NR01 und NR16) sowie der händischen Variante (NR01, NR16 und PH13) womit ein Vergleich nicht mehr möglich ist. Ebenfalls zu beachten ist dass die Werte der verglichenen Videos sehr streuen, von $\kappa_w = -.012$ und .154 bis hin zu den höchsten Übereinstimmungswerten von $\kappa_w = 1.000$ und .913. Natürlich ist der Kappa-Wert über alle Videos insgesamt akzeptabel, wenn man aber die fehlenden Berechnungen der einzelnen Videos und die Streuung mit einbezieht sollte das Ergebnis trotzdem kritisch gesehen werden.

Die zwei niedrigsten Übereinstimmungswerte beim Vergleich der händischen und computergestützten Kodierung pro Rückmeldungskategorie wurden bei der *korrigierenden Rückmeldung* ($\kappa_w = .394$) sowie der *Rückmeldung - unkodierbar* ($\kappa_w = .390$) erzielt. Wie in Abbildung 3 erkennbar ist kann bei keinem Video zur Kategorie *korrigierende Rückmeldung* der κ_w -Wert von .500 überschritten werden, aber abgesehen von den Videos JV02 ($\kappa_w = -.015$) und NR16 ($\kappa_w = -.018$) mit negativen Werten (welche besagen dass die Übereinstimmung noch kleiner als eine zufällige Übereinstimmung ist) ist der Verlauf der Werte relativ kontinuierlich. Aus diesem Grund kann auch beim Gesamtwert über alle Videos hinweg bei der korrigierenden Rückmeldung von einer schlechten Übereinstimmung ausgegangen werden.

Der Verlauf der κ_w -Koeffizienten für die Beobachtungskategorie *Rückmeldung - unkodierbar* je Videoaufzeichnung zeigt eine klare Streuung der Werte, dies spiegelt sich auch im κ_w -Wert für die gesamte Beobachtungskategorie wieder. Mit .390 wird hier eine schlechte Übereinstimmung erzielt.

Zuletzt soll in diesem Kapitel noch der Übereinstimmungskoeffizient für die Beobachtungskategorie *Rückmeldung* ohne die Differenzierung der einzelnen Kategorien näher betrachtet werden. Der κ_w -Wert deutet mit .647 auf eine gute Über-

einstimmung zwischen den händischen und computergestützten Kodierungen hin. Aber wie gerade erläutert sind trotz dieses guten Wertes in den einzelnen Kategorien durchaus Verbesserungen möglich. Vorschläge für praxisbezogene Adjustierungen folgen im nächsten Kapitel.

11.3 Anmerkungen und Verbesserungsvorschläge

Zuerst muss festgehalten werden dass gerade die computerisierte, automatische Datenübertragung des Programms INTERACT ein großer Vorteil gegenüber des händischen Kodierungsschemas des INTAKT ist. Denn genau hier, bei der manuellen Übertragung von Daten, schleichen sich oft Fehler ein. Auch die framegenaue Erfassung des Zeitpunktes (Beginn und Ende) der Beobachtungskategorien kann durch den derzeitigen Protokollbogen nicht äquivalent repliziert werden, wahrscheinlich auch in Zukunft nicht. Ebenfalls in Verbindung mit dem Protokollbogen steht der offensichtliche Datenverlust beim Wechsel des Kodiersystems der Beobachtungskategorie *Rückmeldung* von einem Kategorie-System hin zu einem Index-System bei der händischen Kodierung. Die genaue Dauer der Rückmeldung ist nicht mehr erfassbar und es wurden in dieser Arbeit bei der händischen Kodierung insgesamt 2.82% weniger Kodierungen bei der Beobachtungskategorie *Rückmeldungen* in der händischen Variante getätigt.

Trotzdem ist die Übereinstimmung der händischen mit der computergestützten Kodierung der Beobachtungskategorie *Rückmeldung* anhand der Ergebnisse insgesamt als gut einzustufen. Jedoch fällt bei Betrachtung der einzelnen Kategorien die *korrigierende Rückmeldung* mit schlechten κ_w - und r_{PM} -Werten auf. Diese Kategorie hat eine direkte Verbindung zu den Beobachtungskategorien *positive* und *negative Rückmeldung*, wie durch die Beschreibung im Manual des INTAKT auch deutlich wird: „Diese Rückmeldungen sind für das Kind informativ bzw. hilfreich, wenn auch möglicherweise inhaltlich negativ oder kritisch, aber werden in neutralem oder freundlichem Tonfall vorgebracht.“ (Hirschmann et al., 2013, S. 13). Es wäre denkbar dass einige *korrigierende Rückmeldungen* als *positive* oder *negative Rückmeldungen* kodiert wurden. Ein Blick auf Abbildung 2 könnte diese

Theorie bestätigen: insgesamt wurden bei der Computerkodierung öfter *positive* und *negative Rückmeldungen* kodiert als bei der händischen Kodierung dafür aber insgesamt mehr *korrigierende Rückmeldungen* bei der händischen Kodierart vermerkt.

Da das Konstrukt Feinfühligkeit ein durchaus schwer zu erfassendes Merkmal darstellt und es dadurch auch zu Problemen bei der Interraterreliabilität kommen kann (Wirtz & Caspar, 2002), ist der erzielte Intraklassenkoeffizient von .509 bei der Beobachtungskategorie *Feinfühligkeit* als mehr als akzeptabel einzuordnen.

Zusammenfassend soll erwähnt werden dass der Wechseln zu einem händischen Kodierschema des Beobachtungssystems INTAKT womöglich mit einem Datenverlust gegenüber der computergestützten Kodierung einhergeht. Dennoch sind beim Vergleich der Beobachtungskategorien *Feinfühligkeit* und *Rückmeldung* akzeptable Werte bei der Übereinstimmungsvalidität und BeobachterInnenübereinstimmung erzielt worden.

Abstract (Deutsch)

Das standardisierte Video-Beobachtungssystem INTAKT erfasst die Qualität der Interaktion zwischen Kind und der primären Bezugsperson anhand von drei unterschiedlichen Beobachtungskategorien. Die vorliegende Arbeit soll einen Vergleich des bereits entwickelten händischen Kodierschemas mit dem der computergestützten Auswertung in Bezug auf die Übereinstimmungsvalidität sowie der BeobachterInnenübereinstimmung der Beobachtungskategorien *Feinfühligkeit* und *Rückmeldung* liefern. Zu diesem Zweck wurden 13 Videosequenzen von Mutter-Kind-Interaktion händisch durch den Verfasser der Arbeit und computergestützt durch drei unabhängige Beobachterinnen kodiert.

Die BeobachterInnenübereinstimmung der Beobachtungskategorien *Feinfühligkeit* und *Rückmeldung* wurde mittels Intraklassenkorrelation (r_{IC}) und des gewichteten Kappa-Koeffizient (κ_w) ermittelt. Aufschlüsse zur Übereinstimmungsvalidität sollte die Produkt-Moment-Korrelation (r_{PM}) liefern.

Insgesamt konnten hohe Werte zur Übereinstimmungsvalidität bei der Beobachtungskategorie *Feinfühligkeit* mit $r_{PM}=.901$ sowie *Rückmeldung* mit $r_{PM}=.699$ erzielt werden. Die Ergebnisse zur Interraterreliabilität für die Beobachtungskategorie *Feinfühligkeit* waren $r_{IC}=.509$, dies kann aufgrund der Komplexität des zu beobachtenden Konstrukts als akzeptabel eingestuft werden. Der gewichteten Kappa-Wert zur Beobachtungskategorie *Rückmeldung* zeugt mit $\kappa_w=.699$ ebenfalls von einer guten BeobachterInnenübereinstimmung.

Abstract (Englisch)

The standardised video-observational system INTAKT assesses the quality of the interaction between child and primary caregiver based on three different observation-categories. This current thesis aims to check the already developed manual coding scheme against the conventional computerised coding scheme regarding concurrent validity as well as inter-observer reliability and for the categories *sensitivity* and *feedback*. For this purpose 13 Videos of mother-child-interactions were coded with the manual coding scheme by the author of this thesis as well as by three different independent observers relying on the computerised coding scheme.

The inter-observer reliability for the INTAKT observation-category *sensitivity* and *feedback* were assessed using intraclass correlation (r_{IC}) and weighted kappa (κ_w). The product-moment correlation (r_{PM}) should shed light on the concurrent validity.

Overall, the results regarding the concurrent validity for the observational-categories *sensitivity* with $r_{PM}=.901$ and *feedback* with $r_{PM}=.699$ were high. The results concerning inter-rater reliability for the observational-categories *sensitivity* were $r_{IC}=.509$ and due to the complexity of the construct it can be considered acceptable. The weighted kappa value for the observational-categorie *feedback* with $\kappa_w=.699$ is also an indication for good inter-rater agreement.

Literatur

- Ainsworth, M. D. S., Bell, S. M. & Stayton, D. J. (1974). Infant-mother attachment and social development: "Socialization" as a product responsiveness to signals. In M. Richards (Hrsg.), *The integration of a child into a social world* (S. 99–135). Cambridge: Cambridge University Press.
- Ainsworth, M. D. S., Blehar, M. C., Waters, E. & Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Hillsdale: Erlbaum.
- Alessandri, S. M. & Lewis, M. (1993). Parental Evaluation and Its Relation to Shame and Pride in Young Children. *Sex Roles*, 29(5/6), 335–343.
- Alessandri, S. M. & Lewis, M. (1997). Differences in Pride and Shame in Maltreated and Nonmaltreated Preschoolers. *Child Development*, 67, 1857–1869.
- Atkinson, L., Niccols, A., Paglia, A., Coolbear, J., Parker, K. C. H., Poulton, L., ... Sitarenios, G. (2000). A meta-analysis of time between maternal sensitivity and attachment assessments: Implications for internal working models in infancy/toddlerhood. *Journal of Social and Personal Relationships*, 17(6), 791–810.
- Bakeman, R. [Richard] & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, 55, 1278–1289.
- Bakeman, R. [Richard] & Brownlee, J. R. (1982). Social rules governing object conflicts in toddlers and preschoolers. In K. H. Rubin & H. S. Ross (Hrsg.), *Peer relationships and social skills in childhood* (S. 99–111). New York: Springer.
- Bakeman, R. [Roger] & Gottman, J. M. [John Michi]. (1997). *Observing interaction: An introduction to sequential analysis* (Second Edition). New York: Cambridge University Press.
- Belsky, J. & Fearon, R. M. P. (2002). Early attachment security, subsequent maternal sensitivity, and later child development: Does continuity in development

- depend upon continuity of caregiving? *Attachment and Human Development*, 4(3), 361–387.
- Bodenmann, G. (2006). Beobachtungsmethoden. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (S. 151–159). Göttingen: Hogrefe.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaften* (4., überarbeitete Auflage). Berlin, Heidelberg: Springer.
- Bouvette-Turcot, A.-A., Bernier, A. & Leblanc, É. (2013). Maternal Psychosocial Maladjustment and Child Internalizing Symptoms: Investigating the Modulating Role of Maternal Sensitivity. *Journal of Abnormal Child Psychology*, 45, 157–170.
- Bowlby, J. (1988). *A secure base: Clinical applications of attachment theory*. London: Routledge.
- Claussen, A. H., Mundy, P. C., Mallik, S. A. & Willoughby, J. C. (2002). Joint attention and disorganized attachment status in infants at risk. *Development and Psychopathology*, 14(2), 279–291.
- Cuellar, A. K. & Johnson, S. L. (2009). Depressive Symptoms and Affective Reactivity to Maternal Praise and Criticism. *Journal of Social and Clinical Psychology*, 28(9), 1173–1194.
- De Wolff, M. S. & van Ijzendoorn, M. H. (1997). Sensitivity and Attachment: A Meta-Analysis on Parental Antecedents of Infant Attachment. *Child Development*, 68(4), 571–591.
- Deimann, P. & Kastner-Koller, U. (2007). Entwicklungspsychologie. In M. Haselhorn & W. Schneider (Hrsg.), *Handbuch der Entwicklungspsychologie* (S. 558–569). Göttingen: Hogrefe.
- Dunitz-Scheer, M., Scheer, P., Stadler, B. & Kaimbacher, P. S. (2011). Interaktionsdiagnostik oder die Entmystifizierung der empathisch beobachtenden Interpretation zwischen-menschlichen Verhaltens. In H. Keller (Hrsg.), *Hand-*

- buch der Kleinkindforschung* (4., überarbeitete Auflage, S. 962–983). Bern: Hans Huber.
- Evans, C. A. & Porter, C. L. (2009). The emergence of mother-infant co-regulation during the first year: Links to infants' developmental status and attachment. *Infant Behavior and Development*, 32, 147–158.
- Faßnacht, G. (1995). *Systematische Verhaltensbeobachtung: Eine Einführung in die Methodologie und Praxis* (Zweite, völlig neubearbeitete Auflage). München: Ernst Reinhardt Verlag.
- Feldman, R., Eidelman, A. I. & Rotenberg, N. (2004). Parenting Stress, Infant Emotion Regulation, Maternal Sensitivity, and the Cognitive Development of Triplets: A Model for Parent and Child Influences in a Unique Ecology. *Child Development*, 75(6), 1774–1791.
- Fleiss, J. L., Levin, B. & Paik, M. C. (2003). *Statistical Methods for Rates and Proportions* (Third Edition). New Jersey: Wiley.
- Goldsmith, H. H. & Alansky, J. A. (1987). Maternal and Infant Temperamental Predictors of Attachment: A Meta-Analytic Review. *Journal of Consulting and Clinical Psychology*, 55(6), 805–816.
- Gottman, J. M. [Jerry Monit]. (1983). How children become friends. *Monographs of the Society for Research in Child Development*, 48(3).
- Greve, W. & Wentura, D. (1997). *Wissenschaftliche Beobachtung: Eine Einführung*. Weinheim: Psychologie Verlags Union.
- Grossmann, K. & Grossmann, K. E. (2004). *Bindungen: Das Gefüge psychischer Sicherheit* (5., vollständig überarbeitete Auflage). Stuttgart: Klett-Cotta.
- Harter, S. (1978). Effectance Motivation Reconsidered: Toward a Developmental Model. *Human Development*, 21, 34–64.
- Hirschmann, N., Aigner, N., Deimann, P. & Kastner-Koller, U. (2013). INTAKT: Ein Video-Beobachtungsinstrument zur Erfassung der Mutter-Kind-Interaktion - Manual. *unveröffentlichtes Manuskript, 2. Auflage*.

- Hirschmann, N., Kastner-Koller, U., Deimann, P., Aigner, N. & Svecz, T. (2011). INTAKT: A new instrument for assessing the quality of mother-child interactions. *Psychological Test and Assessment Modeling*, 53(3), 295–311.
- Huber, P. (2013). *Zusammenhang zwischen mütterlicher Feinfühligkeit und kindlicher Entwicklung: Ein Beitrag zur Validierung des Video-Beobachtungssystems INTAKT* (Diplomarbeit, Universität Wien, Wien). Zugriff unter <http://othes.univie.ac.at/27552/>
- Jack, A., Mikami, A. Y. & Calhoun, C. D. (2011). The Moderating Role of Verbal Aggression on the Relationship Between Parental Feedback and Peer Status Among Children with ADHD. *Journal of Abnormal Child Psychology*, 39, 1059–1071.
- Jakob, A. (2016). *Interaktionsbeobachtung von Eltern und Kind: Methoden - Indikation - Anwendung: Ein Praxisbuch* (2., überarbeitete und erweiterte Auflage 2016). Stuttgart: Kohlhammer Verlag.
- Kelley, S. A., Brownell, C. A. & Campbell, S. B. (2000). Mastery Motivation and Self-Evaluative Affect in Toddlers: Longitudinal Relations with Maternal Behavior. *Child Development*, 71(4), 1061–1071.
- Kim, S., Kochanska, G., Boldt, L. J., Nordling, J. K. & O'Brien, J. J. (2014). Developmental trajectory from early responses to transgression to future antisocial behavior: Evidence for the role of the parent-child relationship from two longitudinal studies. *Development and Psychopathology*, 26, 93–109.
- Kindsvater, S. & Sturm, W. (2003). Computer- vs. Papier-Bleistiftvorgabe: Äquivalenzstudie zum nonverbalen Lerntest (NVLT). *Zeitschrift für Neuropsychologie*, 14(1), 13–21.
- Klinck, D. (2006). Computerisierte Methoden. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (S. 226–232). Göttingen: Hogrefe.

- Kochanska, G., Aksan, N. & Carlson, J. J. (2008). Temperament, Relationships, and Young Children's Receptive Cooperation With Their Parents. *Developmental Psychology*, 41(4), 648–660.
- Kok, R., Linting, M., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Jaddoe, V. W. V., Hofman, A., ... Tiemeier, H. (2013). Maternal Sensitivity and Internalizing Problems: Evidence from Two Longitudinal Studies in Early Childhood. *Child Psychiatry and Human Development*, 44, 751–765.
- Kubinger, K. D. (2009). *Psychologische Diagnostik: Theorie und Praxis psychologischen Diagnostizierens* (2., überarbeitete und erweiterte Auflage). Göttingen: Hogrefe.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Weinheim: Beltz.
- Main, M. & Solomon, J. (1986). Discovery of an insecure-disorganised/disoriented attachment pattern: Procedures, findings and implications for the classification of behaviour. In T. B. Barzelton & M. W. Yogman (Hrsg.), *Affective development in infancy* (S. 95–124). Norwood: Ablex.
- Mead, A. D. & Drasgow, F. (1993). Equivalence of Computerized and Paper-and-Pencil Cognitive Ability Tests: A Meta-Analysis. *Behavior Research Methods, Instruments and Computers*, 114(3), 449–458.
- Owen, D. J., Slep, A. M. S. & Heyman, R. E. (2012). The Effect of Praise, Positive Nonverbal Response, Reprimand, and Negative Nonverbal Response on Child Compliance: A Systematic Review. *Clinical Child and Family Psychology Review*, 15, 364–385.
- Page, M., Wilhelm, M. S., Wendy Chrisi Gamble & Card, N. A. (2010). A Comparison of maternal sensitivity and verbal stimulation as unique predictors of infant social-emotional and cognitive development. *Infant Behavior and Development*, 33, 101–110.
- Reischer, N. (2013). *Zusammenhang zwischen mütterlicher Rückmeldung und kindlicher Entwicklung: Ein Beitrag zur Validierung des Video-Beobachtungs-*

- systems INTAKT* (Diplomarbeit, Universität Wien, Wien). Zugriff unter <http://othes.univie.ac.at/27599/>
- Sarimski, K. (2006). Frühdiagnostik. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (S. 603–610). Göttingen: Hogrefe.
- 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(1), 9–25.
- Schneewind, K. A. (2006). Familiendiagnostik. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (S. 587–592). Göttingen: Hogrefe.
- Schölmerich, A. (2011). Verhaltensbeobachtung. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (4., überarbeitete Auflage, S. 768–790). Bern: Hans Huber.
- Schulenberg, S. E. & Yutrzenka, B. (1999). The equivalence of computerized and paper-and-pencil psychological instruments: Implications for measures of negative affect. *Behavior Research Methods, Instruments and Computers*, 31(2), 315–321.
- Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment and Human Development*, 7(4), 349–367.
- Stright, A. D., Gallager, K. C. & Kelley, K. (2008). Infant Temperament Moderates Relations Between Maternal Parenting in Early Childhood and Children's Adjustment in First Grade. *Child Development*, 79(1), 186–200.
- Thiel, T. (2011). Film- und Videotechnik in der Psychologie. Eine Entwicklungsgeschichte aus erkenntnistheoretisch-methodischer Perspektive. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (4., überarbeitete Auflage, S. 792–818). Bern: Hans Huber.
- Thoden, D. (2014). *Beobachtungssystem INTAKT: Übereinstimmung von Computercodierung und händischer Codierung bezüglich der Kategorie Joint At-*

- tion (Diplomarbeit, Universität Wien, Wien). Zugriff unter <http://othes.univie.ac.at/32252/>
- Trudewind, C., Unzer, L. & Schneider, K. (1989). Die Entwicklung der Leistungsmotivation. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (S. 962–983). Bern: Hans Huber.
- Vogler, V. (2015). *Zusammenhang zwischen der Joint Attention und kindlicher Entwicklung* (Diplomarbeit, Universität Wien, Wien). Zugriff unter <http://othes.univie.ac.at/38618/>
- Wijndaele, K., Matton, L., Duvigneaud, N., Duquet, J. L. W., Thomis, M., De Bourdeaudhuij, I. & Philippaerts, R. (2007). Reliability, equivalence and respondent preference of computerized versus paper-and-pencil mental health questionnaires. *Computers in Human Behavior*, 23, 1958–1970.
- Wirtz, M. (2006). Methoden zur Bestimmung der Beurteilerübereinstimmung. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (S. 369–380). Göttingen: Hogrefe.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Zimmermann, P. (2007). Bindungsentwicklung im Lebenslauf. In M. Hasselhorn & W. Schneider (Hrsg.), *Handbuch der Entwicklungspsychologie* (S. 326–346). Göttingen: Hogrefe.
- Zouzoula, M.-A. (2015). *Beobachtungssystem INTAKT: Übereinstimmung von Computerkodierung und händischer Kodierung bezüglich der Kategorien Feinfühligkeit und Rückmeldung* (Diplomarbeit, Universität Wien, Wien). Zugriff unter <http://othes.univie.ac.at/39572/>

Abbildungsverzeichnis

1	Verlauf der Mittelwerte Beobachtungskategorie <i>Feinfühligkeit</i> für händische und computergestützte Kodierung pro Video	45
2	Prozentanzahl der tatsächlich kodierten Rückmeldungen nach Kategorie und Kodierart	54
3	κ_w -Koeffizienten je Videoaufzeichnung für positive und korrigierende Rückmeldung für die 30-Sekunden-Intervalle	57
4	κ_w -Koeffizienten je Videoaufzeichnung für negative und unkodierbare Rückmeldung für die 30-Sekunden-Intervalle	58

Tabellenverzeichnis

1	Retest-Reliabilität ($Rel_{\text{Retest-Methode}}$) für die Skalen Feinfühligkeit (FF) und Rückmeldung (RM) händisch	38
2	Mittelwerte (M) und Standardabweichungen (SD) der Beobachtungskategorie <i>Feinfühligkeit</i> für händische und computergestützte Kodierung per Video und über alle Videoaufzeichnungen (N=13) .	46
3	Häufigkeiten der Beobachtungskategorie <i>Feinfühligkeit</i> für händische und computergestützte Kodierung über alle 13 Videoaufzeichnungen	47
4	Produkt-Moment-Korrelation (r_{PM}) und Intraklassenkorrelation (r_{IC}) für die Kategorie Feinfühligkeit über alle Videos (N=13): Vergleich händische und computerkodierte Auswertung	49
5	Intraklassenkorrelation (r_{IC}) und Kappa-Koeffizient nach Cohen (κ) für die Kategorie Feinfühligkeit über alle Videos der Kodierer NR (N=4), JV (N=4) und PH (N=5): Vergleich händische und computerkodierte Auswertung	50
6	Auftrittshäufigkeiten und Anteilswerte der Beobachtungskategorie Rückmeldung ohne Ausprägung unterteilt in händische und computergestützte Kodierung über alle Videoaufzeichnungen (N=13)	51
7	Häufigkeiten der Beobachtungskategorie Rückmeldung positiv, korrigierend, negativ und unkodierbar der händische und computergestützte Kodierung über alle Videoaufzeichnungen (N=13)	52
8	Korrelationskoeffizienten (κ_w, r_{PM}) für alle Rückmeldungskategorien	55

Anhang I - Feinfühligkeit: standardisierte Schiefe und Kurtosiskoeffizienten

Videocode	Computerkodierung		händische Kodierung	
	Standardisierte Schiefe (z)	Standardisierte Kurtosis (z)	Standardisierte Schiefe (z)	Standardisierte Kurtosis (z)
(1) JV02	-1.542	-0.218	-2.278	0.633
(2) JV03	0.123	-0.100	1.578	-0.175
(3) JV16	0.036	0.221	0.874	-0.863
(4) JV17	-1.309	-0.867	-0.782	-1.199
(5) NR01	1.689	1.541	-1.477	-0.256
(6) NR16	1.422	-1.536	0	2.355
(7) NR17	-0.399	-2.274	0.232	-0.336
(8) NR18	0.875	-1.726	0.425	2.160
(9) PH03	1.555	1.741	-0.260	-1.078
(10) PH04	-0.812	0.026	0.474	-1.001
(11) PH08	0.355	-0.378	0.733	-0.725
(12) PH13	-1.645	1.912	-1.266	-0.414
(13) PH20	-0.092	0.078	0.110	-0.230
Gesamt	-5.352	3.488	-0.207	-2.716

Anmerkung. * Zur Überprüfung der Normalverteilung der Kodierung: Werte innerhalb $\pm 1.96 = NV$

Anhang II - Feinfühligkeit: Überprüfung der Voraussetzung für zweifaktorielles Model der ICC

Rater ¹	Levene-Test ²	Nicht-Additivitäts-Test nach Tukey ³
	$F_{(df1,df2)} (p)$	$x^2_{(df1,df2)} (p)$
JV	$F_{1,174}=18.977 (< .001)^*$	$x^2_{1,86}=13.637 (< .001)^*$
NR	$F_{1,166}= 1.157 (.284)$	$x^2_{1,82}= 0.000 (.990)$
PH	$F_{1,220}=50.885 (< .001)^*$	$x^2_{1,109}=52.461 (< .001)^*$

Anmerkung. * : Annahme der H_1 ; $\alpha = .25$ (Wirtz & Caspar, 2002)

¹ : JV: 02, 03, 16, 17; NR: 01, 16, 17, 18; PH: 03, 04, 08, 13, 20

² : zur Überprüfung der Varianzhomogenität

³ : zur Überprüfung der Interaktion zwischen Beobachterin und beurteilter Person

Anhang III - Protokollbogen Feinfühligkeit

Feinfühligkeit

BeobachterIn:

Videocode:

00:02:00	00:04:00	00:06:00	00:08:00	00:10:00	00:12:00	00:14:00
00:16:00	00:18:00	00:20:00	00:22:00	00:24:00	00:26:00	00:28:00
00:30:00	00:32:00	00:34:00	00:36:00	00:38:00	00:40:00	00:42:00
00:44:00	00:46:00	00:48:00	00:50:00	00:52:00	00:54:00	00:56:00
00:58:00	01:00:00	01:02:00	01:04:00	01:06:00	01:08:00	01:10:00
01:12:00	01:14:00	01:16:00	01:18:00	01:20:00	01:22:00	01:24:00
01:26:00	01:28:00	01:30:00	01:32:00	01:34:00	01:36:00	01:38:00
01:40:00	01:42:00	01:44:00	01:46:00	01:48:00	01:50:00	01:52:00

- 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

Videolänge:

Page 1

Anhang IV - Protokollbogen Rückmeldung

Rückmeldung		BeobachterIn:												Video-code:								
		00:00:00			00:01:00			00:02:00			00:03:00			00:04:00			00:05:00			00:06:00		
		0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	
Rückmeldung																						
positiv																						
korrigierend																						
negativ																						
unkodierbar																						

		00:07:00			00:08:00			00:09:00			00:10:00			00:11:00			00:12:00			00:13:00		
		0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	
Rückmeldung																						
positiv																						
korrigierend																						
negativ																						
unkodierbar																						

		00:14:00			00:15:00			00:16:00			00:17:00			00:18:00			00:19:00			00:20:00		
		0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	
Rückmeldung																						
positiv																						
korrigierend																						
negativ																						
unkodierbar																						

		00:21:00			00:22:00			00:23:00			00:24:00			00:25:00			00:26:00			00:27:00		
		0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	
Rückmeldung																						
positiv																						
korrigierend																						
negativ																						
unkodierbar																						

		00:28:00			00:29:00			00:30:00			00:31:00			00:32:00			00:33:00			00:34:00		
		0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	0	15	30	45	
Rückmeldung																						
positiv																						
korrigierend																						
negativ																						
unkodierbar																						

Video-länge:

Anhang V - Rückmeldung: standardisierte Schiefe und Kurtosiskoeffizienten

Rückmeldung	händische Kodierung		Computerkodierung	
	Standardisierte Schiefe (z)	Standardisierte Kurtosis (z)	Standardisierte Schiefe (z)	Standardisierte Kurtosis (z)
positiv	57.282	133.049	45.155	79.832
korrigierend	66.690	163.923	76.155	220.482
negativ	90.408	354.958	79.718	284.720
unkodierbar	58.591	124.238	54	104.832

Anmerkung. * Zur Überprüfung der Normalverteilung der Kodierung: Werte innerhalb $\pm 1.96 = NV$

Anhang VI - Rückmeldung: Überprüfung der Voraussetzung: Varianzhomogenität

Rückmeldung	Levene-Test ¹
	F (df1,df2) (p)
positiv	F _{1,2324} =45.071 (<.001)*
korrigierend	F _{1,2324} = 2.090 (.148)*
negativ	F _{1,2324} =20.931 (<.001)*
unkodierbar	F _{1,2324} = 1.466 (.226)*

Anmerkung. *: Annahme der H₁; $\alpha = .25$ (Wirtz & Caspar, 2002)

¹: zur Überprüfung der Varianzhomogenität.

Gisela-Legath-Gasse 3/D/5
1220 Wien
0699/101 347 22

LEBENS LAUF

Name: Georg Pavlata

Geburtsdatum: 06. 01. 1988

Geburtsort: Wien

Staatsbürgerschaft: Österreich

Adresse: 1220 Wien, Gisela-Legath-Gasse 3/D/5

E-mail: Georg.Pavlata@gmail.com

Schulische Laufbahn:

1994 - 1998: Volksschule Kittsee

1998 - 2002: Gymnasium Neusiedl am See

2002 – 2007: Höhere Lehranstalt für Wirtschaft am Pannoneum, Schwerpunkt Medieninformatik, Neusiedl am See

Abschluss mit Reife- und Diplomprüfung

seit 2008: Diplomstudium Psychologie und Bachelorstudium Bildungswissenschaften an der Hauptuniversität Wien

seit 2015: Masterstudiengang der Psychologie an der Hauptuniversität Wien

Qualifikationen:

2013/14: Gewaltpräventionskurs IFP – 6 Module

2014: ZARA Workshop „Zivilcourage“

2014: Weiterbildungsreihe „Gesunde Parks“ WIG - Gesunde Ernährung

2015: Weiterbildungsreihe „Gesunde Parks“ WIG -
Seelische Gesundheit bei Kindern und Jugendlichen

2015: Workshop „Kollektiv abwertende Einstellung und
Identitäten bei Jugendlichen“ - think.difference

2015: Bachelor of Science Psychologie

Berufspraxis:

August 2004: Ferialpraktikant bei der Firma
Römerquelle Ges.mbH, 2413 Edelstal

Juni-August 2005: Ferialpraktikant bei der SV Group
Catering & Services, 1040 Wien

August 2006: Ferialpraktikant in der Gemeinde
Edelstal

August 2007: Ferialpraktikant in der Raiffeisen Zentral
Bank Wien (Abt.: Cards & Systems), 1030 Wien

September 2007: Ferialpraktikant in der Raiffeisen
Zentralbank Wien (Abt.: Treasury), 1030 Wien

Oktober-Juni 2008/2009: Zivildienstler beim
Samariterbund Burgenland (Rettungssanitäter), 7331
Weppersdorf

Mai 2010 bis November 2016: Freizeitpädagoge bei
den Kinderfreunden Wien Aktiv

November 2016 bis Jänner 2017: Pflichtpraktikum am
Aktivspielplatz Rennbahnweg

seit August 2017: mobiler Jugendarbeiter beim Verein
Wiener Jugendzentren – Back on Stage 5

**Kenntnisse in
Fremdsprachen**

Deutsch (Muttersprache)

English (fließend in Wort und Schrift)

Französisch (Grundwissen)

Hobbys

Fußball spielen, laufen, kochen

Sonstiges:

A + B Führerschein