



universität  
wien

# DIPLOMARBEIT

Titel der Diplomarbeit

„Anwendbarkeit des Beobachtungsverfahrens INTAKT auf  
den Altersbereich der 1- bis 3-Jährigen im klinischen  
Setting: Eine Reliabilitätsanalyse“

Verfasserin

Alexandra Gogl

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2015

Studienkennzahl: A 298

Studienrichtung: Psychologie

Betreuerin: Ass.-Prof. Dr. Pia Deimann



Verhaltensbeobachtung ist ein sehr altes Verfahren, um Informationen über Individuen zu sammeln. Wenn wir sie nicht wissenschaftlich und nicht systematisch betreiben, verwenden wir die wohl älteste Methode des Datensammelns: das, was unsere Erfahrung ausmacht, was wir im täglichen Umgang mit Menschen bewußt oder unbewußt aufnehmen und was wir zu Ansichten, Meinungen, Voreingenommenheiten und Urteilen über »den Menschen« verarbeiten, ist zu einem großen Teil durch unsystematische Verhaltensbeobachtung bestimmt. (Faßnacht, 1979, S. 192)



## **Vorwort**

Bei Frau Ass.-Prof. Dr. Pia Deimann und Frau Ass.-Prof. Dr. Ursula Kastner-Koller möchte ich mich ganz herzlich für die Bereitstellung meines Diplomarbeitsthemas bedanken. Sie haben mein Interesse für die wissenschaftliche Verhaltensbeobachtung und Eltern-Kind-Interaktionen im Allgemeinen geweckt und mir dadurch eine Vorstellung meines Berufswunschs gegeben. Außerdem sei hier erwähnt, dass die äußerst zuvorkommende und persönliche Betreuung, die ich durch sie erhalten habe, nicht selbstverständlich ist.

Auch Frau Dr. Nicole Hirschmann gilt mein ganz besonderer Dank. Stand sie mir trotz Karenz unermüdlich für Fragen inhaltlicher, technischer und methodischer Natur jederzeit zur Verfügung. Durch ihre fachlich kompetente und ausdauernde Art, erleichterte sie mir das Verfassen dieser Arbeit.

Mein Dank geht auch an Frau Dr. Christine Sonn-Rankl und Frau Dr. Josephine Schwarz-Gerö, den Begründerinnen der Säuglingspsychosomatik des Wilhelminenspitals in Wien. Durch ihre Kooperation und die Bereitstellung der Interaktionsvideos der Station war diese Untersuchung überhaupt erst möglich.

Meinem Kollegen Hossein Valai danke ich für die entspannte Zusammenarbeit und sein sonniges Wesen. Auch meiner Studienkollegin Anna-Maria Zouzoula danke ich von ganzem Herzen für ihre große Hilfe und die gute kollegiale Zusammenarbeit.

Bei meinen Eltern möchte ich mich nach all‘ den Jahren für ihre bedingungslose Liebe, aufopfernde Art sowie großherzige und geduldige Unterstützung bedanken. Ihr lebt mir tagtäglich vor, was gute Eltern ausmacht. Papa, du hast dafür gesorgt, dass ich mich frei von Geschlechterstereotypen entwickeln konnte und dafür danke ich dir sehr.

Zu guter Letzt‘ gebührt mein Dank meinem Freund Johannes. Du bist mir mein ganzes Studium eine große Stütze gewesen. Ich danke dir für deine sensible Art, deine Wissbegierde, deinen Humor und deine aufrichtige Liebe.



## Inhaltsverzeichnis

|                                                                                                                  |           |
|------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Einleitung.....</b>                                                                                           | <b>1</b>  |
| <b>I. THEORETISCHER TEIL.....</b>                                                                                | <b>3</b>  |
| <b>1 Bindung und Bindungssicherheit .....</b>                                                                    | <b>3</b>  |
| 1.1 Abgrenzung Bindung von Bindungsverhalten.....                                                                | 3         |
| 1.2 Bindung als sichere Basis .....                                                                              | 4         |
| 1.3 Arten der Bindung .....                                                                                      | 4         |
| 1.4 Auswirkungen einer sicheren Bindung.....                                                                     | 5         |
| <b>2 Mütterliche Feinfühligkeit .....</b>                                                                        | <b>6</b>  |
| 2.1 Definition der Feinfühligkeit .....                                                                          | 6         |
| 2.2 Studien zur mütterlichen Feinfühligkeit .....                                                                | 6         |
| 2.2.1 Bindungssicherheit, Feinfühligkeit und Trennungsangst .....                                                | 6         |
| 2.2.2 Feinfühligkeit, kindlichen Problemlösefähigkeiten, sozialen Fähigkeiten und<br>berichtete Einsamkeit ..... | 8         |
| 2.2.3 Feinfühligkeit, sozioemotionale und kognitive Entwicklung .....                                            | 9         |
| 2.2.4 Schutzfaktoren der Elterlichen Sensitivität .....                                                          | 11        |
| 2.2.5 Selbstständigkeit, Emotionsregulation und soziale Kompetenzen .....                                        | 12        |
| <b>3 Rückmeldung von Bezugspersonen.....</b>                                                                     | <b>14</b> |
| 3.1 Definition von Rückmeldung .....                                                                             | 14        |
| 3.2 Studien zur mütterlichen Rückmeldung .....                                                                   | 14        |
| 3.2.1 Schamgefühl und Verhalten in herausfordernden Situationen .....                                            | 14        |
| <b>4 Joint Attention .....</b>                                                                                   | <b>17</b> |
| 4.1 Definition der Joint Attention.....                                                                          | 17        |
| 4.2 Studien zur Joint Attention.....                                                                             | 17        |
| 4.2.1 Sprache und intellektuelle Fähigkeiten .....                                                               | 17        |
| 4.2.2 Spielstil.....                                                                                             | 19        |
| 4.2.3 Soziale Kompetenz.....                                                                                     | 19        |
| 4.2.4 Emotionsregulation .....                                                                                   | 20        |

|          |                                                                    |           |
|----------|--------------------------------------------------------------------|-----------|
| <b>5</b> | <b>Systematische Verhaltensbeobachtung.....</b>                    | <b>21</b> |
| 5.1      | Arten der Verhaltensbeobachtung .....                              | 22        |
| 5.1.1    | Unstrukturierte versus strukturierte Beobachtung .....             | 22        |
| 5.1.2    | Offene versus verdeckte Beobachtung .....                          | 22        |
| 5.2      | Beobachtungsfehler in der systemischen Verhaltensbeobachtung ..... | 23        |
| 5.2.1    | Fehler zu Lasten der BeobachterIn .....                            | 23        |
| 5.2.1.1  | Wahrnehmungsfehler .....                                           | 23        |
| 5.2.1.2  | Interpretationsfehler .....                                        | 24        |
| 5.2.2    | Fehler zu Lasten der Beobachtung .....                             | 25        |
| 5.2.2.1  | Reaktivitäts- und Erwartungseffekte .....                          | 25        |
| 5.2.2.2  | Beobachtungsbedingungen .....                                      | 25        |
| 5.2.2.3  | Probleme des Beobachtungssystems .....                             | 25        |
| 5.2.3    | Vermeidung von Beobachtungsfehlern .....                           | 26        |
| 5.3      | Quantifizierungsmethoden.....                                      | 26        |
| 5.3.1    | Time-sampling-Verfahren .....                                      | 27        |
| 5.3.2    | Event-sampling-Verfahren .....                                     | 28        |
| 5.3.3    | Rating-Verfahren.....                                              | 28        |
| 5.4      | Beobachtungssysteme.....                                           | 29        |
| 5.4.1    | Verbal-Systeme .....                                               | 29        |
| 5.4.2    | Nominal-Systeme .....                                              | 29        |
| 5.4.2.1  | Indexsysteme.....                                                  | 30        |
| 5.4.2.2  | Kategoriensysteme .....                                            | 30        |
| 5.4.3    | Dimensional-Systeme .....                                          | 30        |
| 5.4.4    | Struktural-Systeme .....                                           | 31        |
| <b>6</b> | <b>Reliabilitätsanalysen .....</b>                                 | <b>32</b> |
| 6.1      | Gütekriterien.....                                                 | 32        |
| 6.1.1    | Reliabilität .....                                                 | 32        |
| 6.1.2    | Objektivität .....                                                 | 33        |
| 6.1.3    | Zusammenhang von Objektivität und Reliabilität .....               | 33        |
| 6.2      | Intraklassenkorrelation .....                                      | 34        |
| 6.2.1    | Allgemeines zur Intraklassenkorrelation .....                      | 34        |
| 6.2.2    | Modellauswahl in der Intraklassenkorrelation.....                  | 35        |
| 6.3      | Cohens Kappa.....                                                  | 36        |
| 6.4      | Prozentuelle Übereinstimmung .....                                 | 38        |

## Inhaltsverzeichnis

---

|            |                                                                         |           |
|------------|-------------------------------------------------------------------------|-----------|
| <b>II.</b> | <b>EMPIRISCHER TEIL .....</b>                                           | <b>39</b> |
| <b>7</b>   | <b>Hintergrund der Untersuchung.....</b>                                | <b>39</b> |
| 7.1        | Zentraler Ausgangspunkt und Forschungsinteresse .....                   | 39        |
| 7.2        | Kooperation mit der Säuglingspsychosomatik des Wilhelminenspitals ..... | 39        |
| 7.3        | Forschungsfragestellung.....                                            | 40        |
| <b>8</b>   | <b>Untersuchungsinstrument INTAKT .....</b>                             | <b>42</b> |
| 8.1        | Feinfühligkeit .....                                                    | 42        |
| 8.2        | Rückmeldung .....                                                       | 44        |
| 8.3        | Joint Attention .....                                                   | 45        |
| <b>9</b>   | <b>Untersuchungsdesign und Durchführung .....</b>                       | <b>47</b> |
| 9.1        | Design der Untersuchung .....                                           | 47        |
| 9.2        | Durchführung der Datenerhebung .....                                    | 47        |
| 9.2.1      | Einschulung der BeobachterInnen.....                                    | 47        |
| 9.2.2      | Sichtung der Videos .....                                               | 48        |
| 9.2.3      | Aufteilung der Videos .....                                             | 49        |
| 9.2.4      | Auswertungssoftware ‚INTERACT 9‘ von Mangold .....                      | 49        |
| 9.2.5      | Vorbereitung der Daten für die statistische Analyse .....               | 50        |
| <b>10</b>  | <b>Beschreibung der Stichprobe.....</b>                                 | <b>51</b> |
| 10.1       | Inhalte der Interaktionsvideos .....                                    | 51        |
| 10.2       | Beschreibung der Gesamtstichprobe .....                                 | 51        |
| 10.3       | Beschreibung der Gegenkodierungsstichprobe .....                        | 53        |
| 10.4       | Vergleich der Stichproben.....                                          | 54        |
| <b>11</b>  | <b>Ergebnisse der Untersuchung.....</b>                                 | <b>56</b> |
| 11.1       | Skala Feinfühligkeit .....                                              | 56        |
| 11.1.1     | Betrachtung über alle Situationen .....                                 | 56        |
| 11.1.1.1   | Darstellung der Häufigkeiten.....                                       | 56        |
| 11.1.1.2   | Darstellung der Intraklassenkorrelation .....                           | 58        |
| 11.1.2     | Betrachtung der Essens-Situationen .....                                | 60        |
| 11.1.2.1   | Ergebnisse zum Essens-Kontext .....                                     | 60        |

## Inhaltsverzeichnis

---

|           |                                                             |            |
|-----------|-------------------------------------------------------------|------------|
| 11.1.2.2  | Kodierprobleme im Essens-Kontext .....                      | 63         |
| 11.1.3    | Betrachtung der Pflegesituationen .....                     | 63         |
| 11.1.3.1  | Ergebnisse zum Pflege-Kontext .....                         | 63         |
| 11.1.3.2  | Kodierprobleme im Pflege-Kontext .....                      | 66         |
| 11.1.4    | Betrachtung der Spielsituationen .....                      | 68         |
| 11.1.4.1  | Ergebnisse zum Spiel-Kontext .....                          | 68         |
| 11.1.4.2  | Kodierprobleme im Spiel-Kontext .....                       | 70         |
| 11.2      | Skala Rückmeldung .....                                     | 71         |
| 11.2.1    | Interrater-Reliabilität der Kategorie Rückmeldung .....     | 71         |
| 11.2.2    | Kodierprobleme der Rückmeldung .....                        | 75         |
| 11.3      | Skala Joint Attention .....                                 | 76         |
| 11.3.1    | Interrater-Reliabilität der Kategorie Joint Attention ..... | 77         |
| 11.3.2    | Kodierprobleme der Joint Attention .....                    | 80         |
| <b>12</b> | <b>Diskussion und Interpretation .....</b>                  | <b>85</b>  |
| 12.1      | Skala Feinfühligkeit .....                                  | 86         |
| 12.2      | Skala Rückmeldung .....                                     | 89         |
| 12.3      | Skala Joint Attention .....                                 | 90         |
| 12.4      | Conclusio .....                                             | 92         |
| <b>13</b> | <b>Zusammenfassung .....</b>                                | <b>94</b>  |
| <b>14</b> | <b>Abstract .....</b>                                       | <b>96</b>  |
| <b>15</b> | <b>Literatur .....</b>                                      | <b>98</b>  |
| <b>16</b> | <b>Abbildungsverzeichnis .....</b>                          | <b>102</b> |
| <b>17</b> | <b>Tabellenverzeichnis .....</b>                            | <b>103</b> |
| <b>18</b> | <b>Anhang .....</b>                                         | <b>106</b> |

## Einleitung

Die hier vorliegende Arbeit stellt einen ersten Versuch dar, das Beobachtungsinstrument INTAKT (Hirschmann, Kastner-Koller, Deimann, Aigner & Svecz, 2011) auf einen anderen Altersbereich, nämlich auf die 1- bis 3-Jährigen, und in einem anderen Setting anzuwenden. Ursprünglich konzipiert für eine Bastelsituation und eine anschließende freie Spielsituation unter standardisierten Bedingungen, soll INTAKT nun auch Anwendung im klinischen Alltag finden. Die Studie dient der Erprobung des Manuals und des Instruments auf verschiedenste Situationen des klinischen Alltags. Durch die Kooperation mit der Säuglingspsychosomatik des Wilhelminenspitals in Wien standen Interaktionsvideos aus drei verschiedenen Kontexten (Essen, Pflege, Spiel), zu zwei Zeitpunkten (vor und nach erfolgter Intervention) zur Verfügung. Anhand der Beurteilungen zu diesen Interaktionsvideos, getätigt von zwei unabhängigen Ratern, wurde eine Reliabilitätsanalyse durchgeführt, die BeobachterInnenübereinstimmung eruiert und mögliche Differenzen zwischen den BeurteilerInnen erörtert.

Hirschmann, Kastner-Koller, Deimann, Aigner und Svecz (2011) erläutern, dass es einen Mangel an standardisierten Verfahren zur Überprüfung der Qualität von Eltern-Kind-Interaktionen gibt. Das von ihnen entworfene Beobachtungsinventar INTAKT soll diese Lücken schließen. INTAKT ist die Agglomeration der Wörter ‚Interaktion‘ und ‚intakt‘, was auf eine intakte Mutter-Kind-Beziehung anspielt. Es bewertet mütterliche Feinfühligkeit, mütterliche Rückmeldung und mütterliche Vorgehensweisen in Joint Attention-Episoden. Die INTAKT-Beobachtungsdimensionen haben gezeigt, dass sie eine objektive, reliable und valide Erhebung der Mutter-Kind-Interaktionsqualität zulassen (Hirschmann et al., 2011).

Bei den Probanden der vorliegenden Untersuchung bestand zum Zeitpunkt ihrer Aufnahme akuter Handlungsbedarf. Die häufigsten Aufnahmegründe der Station stellen exzessives Schreiverhalten, Fütterungsprobleme, Verhaltensauffälligkeiten sowie aktuelle Beziehungsprobleme zwischen Mutter und Kind dar. Die Station der Säuglingspsychosomatik des Wilhelminenspitals sieht sich mit einem enormen Andrang konfrontiert, der sich in einer Warteliste (monatelanges Warten auf einen Stationsplatz) verdeutlicht. Der dringende Handlungsbedarf im Säuglings- und Kleinkind-Alter, die

Unterstützung von Müttern und somit auch das Verlangen nach einem diagnostischen Instrument für diese Altersgruppe ist mehr als gegeben. Es sei weiter erwähnt, dass die zur Verfügung gestellten Videos mit größtem Respekt, vertraulich und anonymisiert behandelt wurden.

Der Stellenwert mütterlicher Feinfühligkeit sowie einer sicheren Bindung und deren Auswirkungen auf die kindliche Entwicklung sind durch das Vorliegen von wissenschaftlichen Erkenntnissen seit Jahrzehnten klar untermauert. Raikes und Thompson (2008) zeigten, dass Bindungssicherheit und mütterliche Sensitivität zu einer gesteigerten sozialen Problemlösefähigkeit und einem anpassungsfähigen Sozialverhalten beitragen. Mütterliche Depressionssymptome stellen einen Risikofaktor für die kindliche Problemlösefähigkeit dar. Es besteht ein signifikant positiver Zusammenhang zwischen einer sicheren Bindung und einer gelungenen kindlichen Entwicklung, dazu gehören eine gesteigerte Selbstständigkeit, eine erhöhte Fähigkeit zur emotionalen Regulation und die Ausbildung besserer sozialer Kompetenzen (Sroufe, 2005).

Auch die Art der Rückmeldung, die Kinder von ihren Eltern erhalten, kann verschiedene Bereiche ihrer Entwicklung und das kindliche Gefühlsleben beeinflussen. So können negative Rückmeldungen bei Kleinkindern Scham erzeugen. Positive sowie korrigierende Rückmeldungen wirken hingegen fördernd auf die Ausdauer und Beharrlichkeit des Kindes in schwierigen Aufgaben (Kelley, Brownell & Campbell, 2000).

## **I. THEORETISCHER TEIL**

### **1 Bindung und Bindungssicherheit**

Mütterliche Feinfühligkeit und Bindungssicherheit sind stark miteinander verknüpfte Komponenten, die das Fundament einer Mutter-Kind-Beziehung und in weiterer Folge einer gelungenen kindlichen Entwicklung bilden. Die Bedeutung einer sicheren Bindung für das kindliche Verhalten und deren Auswirkungen auf die kindliche Entwicklung werden in diesem Kapitel näher erläutert.

#### **1.1 Abgrenzung Bindung von Bindungsverhalten**

Eine Unterscheidung zwischen Bindung und Bindungsverhalten nahm Bowlby (2008) vor. Unter Bindung definiert er ein natürliches Überlebensmuster, das sich klar vom Nahrungs- und Sexualtrieb abgrenzt. Bindung setzt weiters ein starkes Kontaktbedürfnis gegenüber bestimmten Personen voraus und ist als dauerhaftes, stabiles und situationsabhängiges Merkmal des Bindungssuchenden zu verstehen. Auch Grossmann und Grossmann (2008) beschreiben Bindung als beständig sowie über Raum und Zeit kontinuierlich vorhanden. Ein Kind kann eine dauerhafte Bindung nur zu wenigen erlesenen Personen entwickeln, Bindungsverhalten hingegen kann auch bei anderen als diesen ausgesuchten Personen stattfinden (Bowlby, 2008). Unter Bindungsverhalten versteht Bowlby (2008) jenes „Verhalten, das darauf ausgerichtet ist, die Nähe eines vermeintlich kompetenteren Menschen zu suchen oder zu bewahren, ein Verhalten, das bei Angst, Müdigkeit, Erkrankung und entsprechendem Zuwendungs- oder Versorgungsbedürfnis am deutlichsten wird“ (Bowlby, 2008, S. 21). Nur bei drohender Gefahr der Trennung, des Verlustes oder in anderen Belastungssituationen wird Bindungsverhalten aktiviert (Bowlby, 2008; Grossmann & Grossmann, 2008). Wird ein Kind beispielsweise von seiner Mutter getrennt, so wird Bindungsverhalten hervorgerufen. Allerdings hat ein Kind, das in Gegenwart seiner Mutter spielt und kein Bindungsverhalten aufweist, trotzdem eine Bindung zu seiner Mutter (Grossmann & Grossmann, 2008).

## **1.2 Bindung als sichere Basis**

Die Anwesenheit der Mutter fördert das Explorationsverhalten des Kindes und hat somit die Funktion einer sicheren Basis (Ainsworth & Bell, 1970). Nach Bowlby (2008) bilden Eltern nur dann eine verlässliche Basis, wenn sie das Bindungsverhalten ihrer Kinder intuitiv erfassen, respektieren und als angeborenes Merkmal akzeptieren. Kinder sind nur dann in der Lage psychische Stabilität und Selbstsicherheit zu erlangen, wenn die Eltern ihre Autonomiebestrebungen fördern. Eine sensible Mutter hinterfragt ihr eigenes Verhalten stets, indem sie auf die Signale des Kindes achtet und gegebenenfalls ihr Verhalten korrigiert. Hierfür dürfte die eigene Elternbindung einen starken Einfluss auf die Gefühle und Verhaltensweisen der Mutter gegenüber ihrem Baby ausüben. Kinder sensibler, beruhigend einwirkender Mütter schreien seltener (Ainsworth, Blehar, Waters & Wall, 1978) und waren bereits ab ihrem zweiten Lebensjahr in der Lage Tröstungs- und Hilfbereitschaft zu leisten (Bowlby, 2008). Umgekehrt kümmern sich Mütter, die in weniger optimalen Verhältnissen aufgewachsen sind, signifikant weniger um ihre eigenen Kinder und verfügen über eine ärmere Kommunikation (Bowlby, 2008).

## **1.3 Arten der Bindung**

Nach Ainsworth, Blehar, Waters und Wall (1978) gibt es die Einteilung in drei Arten der Bindung: unsicher-vermeidend, unsicher-ambivalent und sicher gebunden. Dieses Konzept wurde 1986 von Main und Solomon um den desorganisierten Bindungsstil erweitert. Kinder mit einer sicheren Bindung zu beiden Eltern, verhalten sich gegenüber Fremden und neuen Aufgaben optimistisch und selbstbewusst. Sie haben gelernt, dass ihre Eltern in Stress- oder Angstsituationen emotional zur Verfügung stehen. Außerdem bieten Eltern sicher gebundener Kinder einen förderlichen Rückhalt in Explorationssituationen, was für eine gewisse Gelassenheit und Zuversicht in unbekannten Situationen sorgt. Kinder mit einer unsicher-ambivalenten Bindung zu ihren Eltern leben mit einer Ungewissheit, ob sie sich auf diese verlassen können. Sie entwickeln Trennungsängste und weisen nur selten einen Explorationsdrang vor. Der unsicher-vermeidende Bindungsstil ist durch kontinuierliche Ablehnung seitens der Eltern gekennzeichnet, weshalb diese Kinder fortwährend auf fremde Hilfe und Zuneigung verzichten (Bowlby, 2008).

Mütter, die ihren Kindern eine sichere Bindung gewähren, sind aufmerksam, sensibel, hilfsbereit und ermuntern ihr Kind selbst bei Misserfolgen. Hingegen Mütter, die einen unsicher-vermeidenden oder unsicher-ambivalenten Bindungsstil zu ihren Kindern pflegen, agieren weit weniger aufmerksam sowie sensibel und greifen zum falschen Zeitpunkt oder inadäquat ein. Die Gefühle des Kindes werden missachtet, Ermutigungen entfallen und Hilfestellungen bleiben aus (Bowlby, 2008).

Sicher gebundene Kinder werden als kooperativ, beliebt, anpassungsfähig und kreativ beschrieben. Kinder, die einer unsicher-vermeidenden Bindung entstammen, werden als emotional isoliert, feindselig, unsozial und geltungsbedürftig wahrgenommen. Für die Gruppe der unsicher-ambivalenten Kinder konnten Adjektive wie angespannt, impulsiv, leicht frustriert, passiv oder hilflos assoziiert werden (Bowlby, 2008).

### **1.4 Auswirkungen einer sicheren Bindung**

Eine sichere Bindung zwischen Mutter und Kind steht in einem engen Zusammenhang mit einer positiven kindlichen Entwicklung und bedingt unter anderem eine gesteigerte Selbstständigkeit, eine größere Kapazität an Emotionsregulation, bessere Sozialkompetenzen sowie bessere Coping-Strategien (Sroufe, 2005). Des Weiteren zeigen Ergebnisse, dass eine früh erworbene Bindungssicherheit mit einer gesteigerten sozialen Problemlösefähigkeit und einer geringer empfundenen kindlichen Einsamkeit verknüpft ist. Außerdem bedingt eine sichere Bindung, dass aggressive Lösungswege weniger oft vom Kind gewählt werden (Raikes & Thompson, 2008).

Die mütterliche Feinfühligkeit im ersten Lebensjahr bedingt und sagt Bindungssicherheit im zweiten Lebensjahr voraus. Es konnte gezeigt werden, dass sechs bis neun Monate alte Säuglinge feinfühligere Mütter weniger oft weinen und weniger oft Aggression oder Ängstlichkeit zeigen. Das mütterliche Verhalten hat somit eine steuernde Wirkung auf die Entwicklung des Kindes (Grossmann & Grossmann, 2008).

## **2 Mütterliche Feinfühligkeit**

### **2.1 Definition der Feinfühligkeit**

Das Konzept der Feinfühligkeit geht davon aus, dass alle Verhaltensweisen und Äußerungen des Säuglings, der Mutter wichtige Mittelungen liefern und als Informationsträger fungieren. Durch Mimik und Gestik des Neugeborenen erhält die Mutter Rückmeldung darüber, ob ihr Verhalten beim Kind die erwünschte Wirkung erzielt hat. Mütterliche Feinfühligkeit kann sich in vielen Facetten von Verhalten äußern. Besonders der Redestil der Mutter erwies sich als markantes Merkmal mütterlicher Feinfühligkeit. Der Säugling versteht zwar die Inhalte des Gesprochenen noch nicht, dennoch scheint die Art und Weise wie die Mutter mit dem Säugling spricht, prägnant zu sein (Grossmann & Grossmann, 2008).

Feinfühligkeit wird von Ainsworth, Bell und Stayton (1974) als Fähigkeit, die Signale des Kindes zu erkennen, richtig zu interpretieren und darauf prompt und angemessen zu reagieren, definiert. Zwischen mütterlicher Feinfühligkeit und der späteren kindlichen Bindungssicherheit konnte ein hoher Zusammenhang ausfindig gemacht werden (Ainsworth, Blehar, Waters & Wall, 1978). Mütterliche Feinfühligkeit prägt die Mutter-Kind-Beziehungsqualität maßgebend.

Es werden nun einschlägige Erkenntnisse aus der Forschung abgehandelt, die die Bedeutung und Wichtigkeit mütterlicher Feinfühligkeit für den Aufbau von Bindungssicherheit sowie für eine geglückte kindliche Entwicklung aufzeigen.

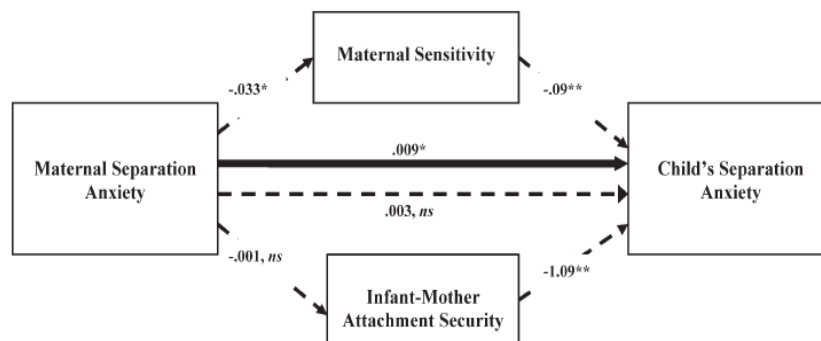
### **2.2 Studien zur mütterlichen Feinfühligkeit**

#### ***2.2.1 Bindungssicherheit, Feinfühligkeit und Trennungsangst***

Dallaire und Weinraub (2005) untersuchten in einer Längsschnittstudie von 99 Mutter-Kind-Dyaden, ob kindliche Trennungsangst in einem Alter von sechs Jahren in Zusammenhang mit Bindungssicherheit, der mütterlichen Angst vor Trennungssituationen

und der mütterlichen Feinfühligkeit steht. Die Partizipanten entstammten der NICHD Study of Early Childcare und wurden regelmäßig in einem Alter zwischen sechs Monaten und sechs Jahren besucht.

Die Autoren (Dallaire & Weinraub, 2005) konnten eine negative Korrelation zwischen der Trennungsangst des Kindes und der mütterlichen Feinfühligkeit finden. Je feinfühlicher die Mutter auf ihr Kind reagierte und einging, desto weniger Symptome der Trennungsangst konnten bei dem Kind gemessen werden. Positiv korreliert waren die von der Mutter berichtete Trennungsangst und die kindliche Angst vor Trennungssituationen in einem Alter von 6 Jahren. Des Weiteren berichteten unsicher gebundene Kinder über mehr Trennungsangst als sicher gebundene. Mittels Regressionsanalyse erhielten Dallaire und Weinraub (2005) die beiden Prädiktoren mütterliche Feinfühligkeit und Bindungssicherheit für die Vorhersage der kindlichen Trennungsangst. Die Variable mütterliche Trennungsangst enthielt keinen erklärenden Charakter zur Vorhersage der kindlichen Trennungsangst. Aufgrund der anschließenden Post-Hoc Analyse konnte darauf geschlossen werden, dass der Einfluss der mütterlichen Trennungsangst über die Mediatorvariable mütterliche Feinfühligkeit auf die kindliche Trennungsangst wirkt (siehe Abbildung 1).



**Abbildung 1:** Wirkungsmodell der kindlichen Trennungsangst. ns = not significant; \* $p < .05$ ; \*\* $p < .01$  (Dallaire & Weinraub, 2005, S. 404)

### **2.2.2 Feinfühligkeit, kindliche Problemlösefähigkeiten, soziale Fähigkeiten und berichtete Einsamkeit**

Raikes und Thompson (2008) untersuchten in ihrer Studie die möglichen Auswirkungen von Bindungssicherheit und mütterlicher Sensitivität auf die spätere kindliche soziale Entwicklung. Des Weiteren betrachteten die Autoren die Risikofaktoren mütterliche Depressionssymptome sowie fehlende Bindungssicherheit und deren Konsequenzen näher. Raikes und Thompson (2008) rekrutierten ihre Partizipanten im Zuge der NICHD Studie (Early Child Care and Youth Development) über die Geburtsstationen verschiedener Spitäler in mehreren Städten Amerikas. Insgesamt wurden 1364 Kinder vom ersten Lebensmonat bis zum sechsten Lebensjahr begleitet.

Aus den Analysen von Raikes und Thompson (2008) ergab sich ein signifikanter Zusammenhang zwischen früher Bindungssicherheit und dem elterlichen Erziehungsstil bezogen auf die Repräsentationen des Kindes, die es über Gleichaltrige hat. Sicher gebundene Kinder zeigten vermehrt positive soziale Repräsentationen und anpassungsfähigeres Sozialverhalten im Gegensatz zu unsicher gebundenen Kindern. Eine Interaktion zwischen Bindungssicherheit und Erziehungsqualität fand statt, die vorhersagend auf die kindlichen Repräsentationen wirkte.

Des Weiteren ergab die Studie einen negativen Zusammenhang zwischen den Depressionssymptomen der Mutter und den kindlichen Problemlösefähigkeiten. Der Einfluss mütterlicher Depressionssymptome auf die kindlichen Problemlösefähigkeiten war jedoch geringer, wenn das Kind sicher gebunden war. Bindungssicherheit kann somit als Ressource und protektiver Faktor fungieren. Fehlende Bindungssicherheit und mütterliche Depressionssymptome stellen hingegen kumulative Risikofaktoren für die kindlichen Problemlösefähigkeiten dar. Prinzipiell wiesen Kinder der unsicher-vermeidend gebundenen Gruppe weniger sozial kompetente Lösungen auf als sicher gebundene. Bindungssicherheit und Feinfühligkeit stellen somit wichtige Einflussgrößen für die kindlichen Problemlösefähigkeiten dar (Raikes & Thompson, 2008).

Sicher gebundene Kinder berichteten auch über weniger Einsamkeit und wiesen außerdem weniger negative soziale Attributionen als unsicher gebundene Kinder auf. Bindungssicherheit und frühe mütterliche Feinfühligkeit stellen negative Prädiktoren für die kindliche Einsamkeit dar (Raikes & Thompson, 2008).

Für die aggressiven Lösungswege seitens des Kindes konnte eine negative Korrelation zur mütterlichen Feinfühligkeit und eine positive Korrelation zur mütterlichen depressiven Symptomatik gefunden werden. Kinder, mit einer weniger feinfühligem Mutter und Kinder depressiver Mütter, generierten häufiger aggressive Lösungswege. Mütterliche Feinfühligkeit zeigte sich als negativer Prädiktor für die aggressiven Reaktionen des Kindes in sozialen Situationen (Raikes & Thompson, 2008).

Raikes und Thompson (2008) stellten außerdem fest, dass Bindungssicherheit und eine mögliche Depressionssymptomatik der Mutter auch für die Art der Attribuierung des Kindes eine wichtige Rolle spielen. Unsicher-vermeidend gebundene Kinder produzierten mehr negative Attributionen als sicher gebundene. Kinder mit Müttern, die schon sehr früh im Leben ihres Kindes eine schwere Depressionssymptomatik zeigten, tätigten in der ersten Schulstufe mehr negative Attributionen als Kinder mit einer Mutter ohne Depressionssymptome.

### ***2.2.3 Feinfühligkeit, sozioemotionale und kognitive Entwicklung***

Stams, Juffer und van Ijzendoorn (2002) untersuchten in einer Adoptionsstudie im Längsschnitt-Design 146 international adoptierte Kinder, die vor ihrem sechsten Lebensmonat in der nicht-biologischen Familie platziert und von den Forschern bis zu ihrem siebten Lebensjahr begleitet wurden. Die Auswahl, welches Kind zu welchen Eltern zugeteilt wurde, erfolgte über die Wartelisten der Adoptionsorganisationen. Die Familien wurden sieben Jahre lang begleitet. Alle Adoptiionseltern waren kaukasisch-weiß und in allen Adoptionsfamilien war primär die Mutter für die Erziehung zuständig. Die Familien gehörten überwiegend der Mittelschicht oder der oberen Mittelschicht an. Von den Adoptionskindern stammten 79 ursprünglich aus Sri Lanka, 48 aus Südkorea und 19 aus Kolumbien (Stams et al., 2002, S. 809).

Als Prädiktoren für die soziale kindliche Entwicklung konnten Stams et al. (2002) das Geschlecht des Kindes, mütterliche Sensitivität, Bindungssicherheit und frühes kindliches Temperament ausfindig machen. Dabei wiesen Mädchen im Generellen höhere Werte in der sozialen Entwicklung auf als Buben. Mütterliche Feinfühligkeit bewirkte höhere Werte in der sozialen Entwicklung sowie eine sichere Bindung, die für eine optimale soziale

Entwicklung verantwortlich war. Ein schwieriges Temperament in der frühen Kindheit bedingte hingegen eine weniger gut ausgeprägte soziale Entwicklung.

Vorhersagend für Ego-Resilienz (die Fähigkeit seine Gefühle und Impulse flexibel an die situativen Gegebenheiten und Anforderungen anzupassen) waren die mütterliche Feinfühligkeit und das kindliche Temperament in der frühen Kindheit. Je höher die Feinfühligkeit der Mutter ausgeprägt war, desto besser war auch die kindliche Fähigkeit zur Ego-Resilienz entwickelt. Ein schwieriges kindliches Temperament hingegen war mit einer verminderten Fähigkeit zur situativen Anpassung verknüpft (Stams et al., 2002).

Erklärend für die Ego-Kontrolle des Kindes zeigten sich das Geschlecht des Kindes, die Bindungssicherheit zur Mutter und das kindliche Temperament. Mädchen waren optimaler Ego-kontrolliert als Jungen. Es fand eine Interaktion zwischen Bindungssicherheit und kindlichem Temperament statt. Kinder mit einem als schwierig wahrgenommenen Temperament und einer desorganisierten Bindung zur Mutter wiesen die ungünstigste Ego-Kontrolle auf (Stams et al., 2002).

Für die kognitive Entwicklung von Kindern konnten die Autoren (Stams et al., 2002) die Prädiktoren mütterliche Feinfühligkeit, Bindungssicherheit und kindliches Temperament finden. Kinder mit einem schwierigen Temperament in der frühen Kindheit zeigten ein niedrigeres Niveau auf der kognitiven Ebene. Eine hohe mütterliche Feinfühligkeit zu einem frühen Zeitpunkt im Leben des Kindes war mit einer besseren kognitiven Entwicklung verknüpft. Außerdem wiesen Kinder mit einem als schwierig eingestuften Temperament und einer desorganisierten Bindung zur Mutter, die niedrigsten Werte auf der kognitiven Ebene auf.

Für das externalisierende Verhalten galt, dass Mädchen im Generellen geringere Werte aufwiesen als Burschen. Weiters zeigten Kinder, die in ihrer frühen Kindheit ein schwieriges Temperament aufwiesen, vermehrt externalisierendes Verhalten, wie auch internalisierte Verhaltensauffälligkeiten in der mittleren Kindheit (Stams et al., 2002).

Stams et al. (2002) betonten die Wichtigkeit der Prädiktoren mütterliche Feinfühligkeit, Bindungssicherheit und kindliches Temperament für die Vorhersage der Entwicklung in der mittleren Kindheit. Die Kombination aus zwei Risikofaktoren, nämlich Bindungsunsicherheit und schwieriges Temperament, bedingt ein niedrigeres Niveau in

den kognitiven Fertigkeiten sowie eine weniger optimale Ego-Kontrolle. Die frühe Mutter-Kind-Beziehung ist ausschlaggebend für die spätere kognitive und soziale Entwicklung des Kindes, was abermals den Einfluss früher Umweltfaktoren als vorhersagende Kraft kindlicher Entwicklung hervorhebt.

### **2.2.4 Schutzfaktoren der Elterlichen Sensitivität**

Hänggi, Benz-Fragnière, Haberkorn, Furler und Perrez (2013) sehen elterliche Sensitivität als wichtige Determinante für eine gesunde kindliche Entwicklung. Demgegenüber steht ein Mangel an elterlicher Feinfühligkeit, der die Entstehung kindlicher Störungen begünstigt.

Das Ziel ihrer Studie war es, Risiko- und Schutzfaktoren zu eruieren, die die Sensitivität von Müttern und Vätern beeinflussen. Zu diesem Zweck verwendeten sie ein Sensitivitätstraining, das nur einer Hälfte der Eltern zur Schulung ihrer erzieherischen Kompetenzen angeboten wurde. Die Eltern der Experimentalgruppe lernten im Zuge der Intervention, wie sie die Signale ihrer Säuglinge zu deuten haben, Perspektivenübernahme, das Teilen von positiven Emotionen und die Kontrolle von Ärger. Die Rekrutierung der Stichprobe erfolgte über das Fachpersonal von Spitälern (z.B. Hebammen) oder über öffentliche Ausschreibungen. Insgesamt nahmen 91 Mütter und 91 Väter derselben Familien an der Studie teil (Hänggi et al., 2013).

Im Zuge der Analysen konnte ein signifikant positiver Zusammenhang zwischen den protektiven Faktoren und mütterlicher Sensitivität eruiert werden. Schutzfaktoren, wie individuelles Coping, positive Einstellung zur Elternschaft, Selbstwirksamkeit bei der Erziehung, Partnerschaftsqualität oder dyadisches Coping, beeinflussen die Feinfühligkeit der Eltern positiv. Des Weiteren konnte ein negativer Zusammenhang zwischen den Risikofaktoren, wie psychische Belastungen oder ein schwieriges Temperament des Kindes, und der mütterlichen Sensitivität gefunden werden (Hänggi et al., 2013).

Entgegen der Erwartungen der Autoren berichteten Eltern, mit einem Kind, über eine höhere Feinfühligkeit als Eltern, die mehrere Kinder hatten. Dies begründeten Hänggi et al. (2013) damit, dass diese zwar mehr Erfahrung in der Kindererziehung haben, jedoch auch weniger Zeit für jedes einzelne Kind aufbringen können. Als Prädiktoren für die spätere

Sensitivität der Eltern erwiesen sich bei den Müttern eine funktionale Stressbewältigung und bei den Vätern eine positive Einstellung zur Elternschaft. Es zeigte sich, dass personennahe Faktoren gegenüber sozialen Variablen einen größeren Einfluss auf die Sensitivität ausüben. Hänggi et al. (2013) postulieren, dass ihre Erkenntnisse nur auf höher gebildete Eltern, die an einem Sensitivitätstraining teilgenommen haben, übertragbar sind.

### ***2.2.5 Selbstständigkeit, Emotionsregulation und soziale Kompetenzen***

In der Minnesota Studie der 1970er Jahre wurden 200 Mütter, die ein Risiko für elterliche Erziehungsprobleme aufwiesen und armutsgefährdet waren, rekrutiert. Sroufe (2005) präsentierte nach 30 Jahren Längsschnittstudie den positiven Zusammenhang zwischen einer sicheren Bindung und einer gelungenen kindlichen Entwicklung, dazu gehören eine gesteigerte Selbstständigkeit, eine erhöhte Fähigkeit zur emotionalen Regulation und die Ausbildung besserer sozialer Kompetenzen.

Sroufe (2005) konnte zeigen, dass Mütter mit unsicher gebundenen Kindern (dazu gehören unsicher-vermeidend und unsicher-ambivalent gebundene Kinder) niedrigere Werte in der Feinfühligkeit aufweisen, als Mütter mit sicher gebundenen Kindern. Außerdem weisen Mütter von unsicher-ambivalent gebundenen Kindern ein geringes psychologisches Bewusstsein auf und deren Kinder Entwicklungseinbußen. Mütter von Kindern mit einer unsicher-vermeidenden Bindung beschreiben negative Gefühle gegenüber der Elternschaft. Des Weiteren sind diese Mütter angespannt, leicht reizbar, kümmern sich nur sehr nachlässig um ihre Kinder und zeigen eine mangelhafte emotionale Bindung bzw. Beteiligung.

Körperliche Misshandlung und psychische Unerreichbarkeit seitens der Erziehungsperson sagen einen desorganisierten Bindungsstil vorher. Nach Sroufe (2005) spiegelt der desorganisierte Bindungsstil einer Mutter stark ihre vergangenen Beziehungsqualitäten, sozusagen ihre Bindungshistorie, wieder.

Sicher gebundene Kinder sind hingegen in der Lage ihre Bezugsperson als Ausgangspunkt für Erkundungen zu nutzen. Sie sind später unabhängiger und weisen ein höheres Selbstvertrauen auf. Ein weiterer Befund der Studie legt offen, dass sicher gebundene Kinder selbstbewusster sind, einen höheren Selbstwert aufweisen und mehr Ego-Resilienz

(die Fähigkeit seine Gefühle und Impulse flexibel an die situativen Gegebenheiten und Anforderungen anzupassen) besitzen. Außerdem tendieren sicher gebundene Kinder weniger oft dazu aufzugeben und reagieren seltener mit Frustration oder Aggression auf soziale Probleme. Ihre Coping-Strategien sind gekennzeichnet durch Ausdauer und Flexibilität. Was die soziale Kompetenz von sicher gebundenen Kindern betrifft, so haben sie positive Erwartungen gegenüber anderen und zeigen sich auch sozial kompetenter als unsicher gebundene Kinder. Sie sind aktiver in der Gruppe Gleichaltriger, weniger oft isoliert, werden von der Peer-Gruppe öfter als sympathisch eingestuft und haben vermehrt tiefgründige, wechselseitige Freundschaften. In der Pubertät zeigen Kinder des sicheren Bindungsstils öfter Führungsqualitäten.

Neue Situationen, kognitive Herausforderungen und hohe Stimulationen sind für unsicher-ambivalent gebundene Kinder besonders schwierig. Sie zeigen häufiger Unschlüssigkeit, eine weniger intensive Exploration, weniger Flexibilität und weniger Effektivität in der Lösung von Problemen. In sozialen Problemsituationen ist ihre Strategie oft das Verlassen der Situation. Für unsicher-vermeidend gebundene Kinder sind Situationen mit interpersoneller Nähe besonders herausfordernd. Freundschaften zu schließen stellt sich als sehr schwierig heraus.

Ein weiterer Befund der Studie hebt hervor, dass Eltern nicht die einzige soziale Einflussquelle darstellen. Auch Geschwister und Gleichaltrige der Peer-Group prägen das Kind und somit auch die kindliche Entwicklung. Die Kombination aus Bindung, elternbezogenen Variablen und peerbezogenen Variablen ist ein starker Prädiktor für die kindliche Entwicklung.

### **3 Rückmeldung von Bezugspersonen**

#### **3.1 Definition von Rückmeldung**

Die Beurteilung von kindlichen Handlungen und Ergebnissen durch die Eltern beginnt schon sehr früh und überdauert die gesamte Kindheit, was den Einfluss von Eltern und deren Bewertungen verdeutlicht (Kelley et al., 2000). Die mütterliche Rückmeldung ist einerseits als Informationsquelle zu betrachten und dient dem Kind als eine Art Bezugsnorm, durch die es die Angemessenheit seiner Handlungen einstufen kann. Auf der anderen Seite geben Rückmeldungen Aufschluss darüber, ob das Kind die elterlichen Erwartungen erfüllt oder nicht. Rückmeldungen beeinflussen die Motivation, das Selbstbewusstsein, den Selbstwert und das erlebte Kontrollgefühl des Kindes (Harter, 1978; Kelley et al., 2000). Positive Rückmeldungen zu kindlichen Handlungen steigern das Kompetenz- und Kontrollgefühl, was einen positiven Selbstwert zur Folge hat (Harter, 1978). Außerdem fördern positive Rückmeldungen die intrinsische Motivation für „mastery situations“, was Freude und Interesse an solchen Aufgaben bedingt (Kelley et al., 2000). Eine positive Übereinstimmung mit der Umwelt führt zu der Überzeugung und dem Gefühl von Kompetenz sowie von Selbstbestimmung, hier herrscht also eine interne Kontrollüberzeugung vor. Negative Rückmeldungen haben den gegenteiligen Effekt. Sie führen zu einem geringen Selbstbewusstsein, einem Mangel an Kontroll- und Kompetenzüberzeugung sowie zu einer externalisierenden Kontrollüberzeugung. Außerdem bedingen negative Rückmeldungen in „mastery situations“ Angst und mindern die Motivation zur Selbstwirksamkeit (Harter, 1978).

#### **3.2 Studien zur mütterlichen Rückmeldung**

##### ***3.2.1 Schamgefühl und Verhalten in herausfordernden Situationen***

In der Untersuchung von Kelley, Brownell und Campbell (2000) wurden die Auswirkungen der mütterlichen Rückmeldung (operationalisiert anhand des Kontrollstils und der Art der Rückmeldung der Mutter) auf das kindliche Verhalten erforscht. 75 Kleinkinder wurden im Zuge der NICHD Study of Early Child Care rekrutiert, davon

waren 35 weiblich und 40 männlich. Die Erhebung fand, zum Zeitpunkt als die Kinder zwei Jahre alt waren, statt.

Negative Rückmeldungen zu den kindlichen Handlungen wirkten sich signifikant auf das Schamgefühl des Kindes auf. Keine signifikante Beziehung konnte zwischen dem Schamgefühl des Kindes und negativen Rückmeldungen zur kindlichen Person gefunden werden. Korrigierende Rückmeldungen hatten auch keinen signifikanten Einfluss auf das kindliche Schamgefühl. Jedoch konnte nachgewiesen werden, dass sich korrigierende Rückmeldungen auf die spätere kindliche Beharrlichkeit eine Aufgabe zu beenden auswirken. Auch positive Rückmeldungen, sowohl zur Person des Kindes als auch zu den kindlichen Handlungen, konnten das kindliche Durchhaltevermögen in schwierigen Aufgaben vorher bestimmen. Ein – vor allem für Vertreter der positiven Bestärkung – interessanter Befund ist, dass sich negative Rückmeldungen nicht auf die kindliche Beständigkeit in herausfordernden Situationen auswirken. Die Ergebnisse verdeutlichen, dass je nach Art der Rückmeldung die Mutter die Möglichkeit hat, förderlich oder hemmend auf die Kindesentwicklung einzuwirken (Kelley, Brownell & Campbell, 2000).

In einer weiteren Untersuchung, zu den Auswirkungen mütterlicher Rückmeldung auf das Schamgefühl und die Reaktion des Kindes bei schwierigen Aufgaben, rekrutierten Alessandri und Lewis (1996) 84 Kinder im Alter zwischen vier und fünf Jahren. Von den 84 Kindern kamen 42 aus einem misshandelnden Elternhaus und 42 Kinder bildeten die Kontrollgruppe, die aufgrund des Interviews keine Vorgeschichte aufwiesen. Die Geschlechterverteilung war in beiden Gruppen mit 21 Buben und 21 Mädchen ausgeglichen. Des Weiteren waren alle Familien afro-amerikanischer Herkunft. Der Kontakt zu der Gruppe der misshandelten Kinder wurde über Sozialarbeiter der Stadt Philadelphia hergestellt. 50% der Kinder der Experimentalgruppe waren emotionaler Misshandlung, 36% körperlicher Misshandlung, 26% sexuellen Missbrauchs und 52% Vernachlässigung ausgesetzt (Alessandri & Lewis, 1996, S. 1859). Von den 42 Kindern der Experimentalgruppe erlebten 27 mehr als nur eine Form der Misshandlung und das immer ausgeführt durch die Mutter. Alle Kinder lebten zum Zeitpunkt der Studie zu Hause.

Mädchen aus misshandelten Familien zeigten signifikant mehr Schamgefühl und weniger Stolz als die Mädchen der Kontrollgruppe. Buben der Experimentalgruppe (aus misshandelten Familien) zeigten weniger emotionale Ausdrücke als Buben der Kontrollgruppe. Die Mädchen der Experimentalgruppe zeigten am wenigsten Stolz von allen Gruppen, wenn sie eine schwierige Aufgabe erfolgreich bewältigt hatten. Sie zeigten auch am meisten Scham bei Misserfolg in schwierigen Aufgaben. Die Behandlung durch die Familie wirkte sich somit sehr stark auf das Scham- und Stolzgefühl der Kinder sowie auf die Fähigkeit Emotionen zu zeigen, aus (Alessandri & Lewis, 1996).

Ein interessanter Geschlechtereffekt konnte außerdem aufgedeckt werden. Buben erhielten vermehrt positives Feedback, Mädchen hingegen mehr negatives. Dies verdeutlicht sich nochmals in Situationen des Erfolges oder Misserfolges, denn Mädchen erhielten bei Misserfolg mehr negatives Feedback und Burschen bei Erfolg mehr positives. Für die misshandelten Kinder konnte festgestellt werden, dass diese mehr negative Rückmeldungen und weniger positive Rückmeldungen erhielten. Des Weiteren tätigten Mütter, die ihre Kinder misshandelten mehr negative Rückmeldungen, wenn ihre Kinder scheiterten, aber auch weniger positive Rückmeldungen, wenn ihre Kinder Erfolg hatten. Außerdem erhielten Mädchen der Experimentalgruppe die meisten negativen Rückmeldungen von allen vier Gruppen (Alessandri & Lewis, 1996).

Ein negativer Zusammenhang kristallisierte sich zwischen den negativen Rückmeldungen der Mutter und dem kindlichem Stolz heraus und ein positiver Zusammenhang zur kindlichen Scham. Das bedeutet, je mehr negatives Feedback die Mutter ausübte, desto weniger Stolz zeigte das Kind und desto mehr Scham entwickelte es. Einen weiteren negativen Zusammenhang gab es zwischen der kindlichen Äußerung von Scham und positiven Rückmeldungen durch die Mutter. Das Ausmaß der positiven Äußerungen durch die Mutter wirkt also reduzierend auf die kindliche Scham (Alessandri & Lewis, 1996).

## 4 Joint Attention

### 4.1 Definition der Joint Attention

Morales, Mundy, Delgado, Yale, Messinger, Neal und Schwartz (2000) bezeichnen Joint Attention als die Fähigkeit des Kindes seine Aufmerksamkeit zwischen einem Interaktionspartner und einem Gegenstand oder Ereignis zu koordinieren. Dabei können zwei Unterarten der Joint Attention ausfindig gemacht werden: ‚attention switching‘, wobei sich der Interaktionspartner dazu entscheidet den Aufmerksamkeitsfokus des Kindes direkt zu beeinflussen, indem er die kindliche Aufmerksamkeit auf eines neues Objekt oder eine neue Situation lenkt, und ‚attention following‘, hier folgt der Interaktionspartner der Aufmerksamkeit des Kindes und verweilt mit dem Kind bei dessen momentanen Fokus (Saxon, Colombo, Robinson & Frick, 2000).

### 4.2 Studien zur Joint Attention

#### 4.2.1 *Sprache und intellektuelle Fähigkeiten*

In der Längsschnittstudie von Saxon, Colombo, Robinson und Frick (2000) wurden 23 Mutter-Kind-Dyaden während einer Freispiel-Situation im Alter von sechs und acht Monaten beobachtet und die sprachlichen wie intellektuellen Fähigkeiten des Kindes in einem Alter von zwei und drei Jahren untersucht. Für den Bereich der Joint Attention stellten sie fest, dass Kinder, deren Mütter bis zu einem Alter von sechs Monaten aufmerksamkeitslenkend und ab acht Monaten aufmerksamkeitsfolgend vorgingen, höhere Werte in der sprachlichen Entwicklung, im Anpassungsverhalten und in der generellen Intelligenz aufwiesen. Die Entscheidung der Mutter, wie sie mit dem Aufmerksamkeitsfokus des Kindes umgeht und eventuell darauf einwirkt, ist als wichtiger Indikator für die intellektuellen und adaptiven Fähigkeiten im späteren Leben des Kindes zu betrachten. Die Auswirkungen dessen, manifestieren sich erst zu einem späteren Zeitpunkt in der Kindheit. Dabei sind Merkmale der Interaktion wichtige Prädiktoren in der Vorhersage kognitiver und sprachlicher Fähigkeiten des Kindes im Vorschulalter. So konnte gezeigt werden, dass Kinder aus weniger erfolgreichen Mutter-Kind-Dyaden, über

eine weniger optimale Entwicklung im Sprachlichen und Kognitiven während des zweiten und dritten Lebensjahres verfügten.

Morales et al. (2000) beschäftigten sich mit der Fähigkeit eines Kindes zum Zeitpunkt von ein und zwei Jahren auf Joint Attention-Angebote von anderen zu reagieren. Vordergründiges Forschungsinteresse war, zu eruieren, ob das ‚responding to joint attention‘ im ersten und zweiten Lebensjahr mit dem späteren Vokabelerwerb verknüpft ist. Die Ergebnisse sprechen dafür, dass das Erwidern von Joint Attention-Episoden im Alter von 6, 8, 10, 12 und 18 Monaten positiv mit dem späteren Vokabelerwerb korreliert ist.

Carpenter, Nagell und Tomasello (1998) untersuchten in ihrer Studie eine Gruppe von 24 Kindern im Alter zwischen 9 und 15 Monaten. An jedem der sieben monatlichen Besuche maßen sie verschiedene Elemente der Joint Attention-Episoden sowie soziale und kognitive Fähigkeiten. Zu den sozialen Fertigkeiten gehörten Kommunikationsfähigkeit, Sprachproduktion, Sprachverständnis und Gestik. Die Ergebnisse belegen, dass die Zeit, die Mutter und Kind in einer Joint Attention verbringen sowie das Ausmaß, in dem die Mutter aufmerksamkeitsfolgend vorging, die frühe linguistische und Kommunikationskompetenz des Kindes vorhersagen.

Die Studie von Charman, Baron-Cohen, Swettenham, Baird, Cox und Drew (2000) umfasst eine Stichprobe von 13 Kindern. Es wurden Daten zu den Joint Attention-Episoden im Spiel, zur Theory of Mind, Sprachkenntnissen, zum IQ und der Fähigkeit zur Imitation in einem Alter von 20 Monaten und 44 Monaten gesammelt. Die Autoren (Charman et al., 2000) gehen davon aus, dass Joint Attention, die Fähigkeit zum Spiel und zur Imitation sowie die Sprachkenntnisse und die Fähigkeit zur Theory of Mind, möglicherweise Teil eines gemeinsamen sozial-kommunikativen Repräsentationssystems sind, um überhaupt zu Interaktionen mit Menschen und dem Spielen mit Objekten fähig zu sein. Ihre Ergebnisse belegen, dass ein Zusammenhang zwischen dem Verhalten in Joint Attention-Episoden zum Zeitpunkt von 20 Monaten und der Fähigkeit zur Theory of Mind

mit 44 Monaten besteht. Weiters waren die Fähigkeiten zur Imitation in einem Alter von 20 Monaten verknüpft mit der verbalen Ausdrucksweise in einem Alter von 44 Monaten.

### **4.2.2 Spielstil**

Kleinkinder in einem durchschnittlichen Alter von 12 Monaten wurden gemeinsam mit ihren Müttern in einem Labor von Bigelow, MacLean und Proctor (2004) untersucht. Mehrere Spielsachen standen den Mutter-Kind-Dyaden in einem Untersuchungsraum zur Verfügung. Es gab zwei Untersuchungsbedingungen: in der ersten wurde die Mutter instruiert, nicht mit dem Kind zu interagieren und nicht ins Spielgeschehen einzugreifen, außer das Kind initiierte ihre Beteiligung. In der anderen Situation instruierte man die Mutter, mit ihrem Kind zu spielen, wie sie es für gewöhnlich tat. Die Kinder spielten in den beiden Untersuchungsbedingungen gleich viel, hier gab es keinen Unterschied, jedoch änderte sich die Verteilung der verschiedenen Spiel-Arten. Ein fortgeschritteneres Spielverhalten war in der Situation, in der die Mutter sich solange zurück halten sollte bis das Kind die Mutter zum Spielen aufforderte, wahrscheinlicher. Kleinkinder mit Müttern, die im Spiel aufmerksamkeitsfolgend vorgingen, verfügten somit über einen reiferen und fortgeschritteneren Spielstil (Bigelow, MacLean & Proctor, 2004).

### **4.2.3 Soziale Kompetenz**

Vaughan Van Hecke et al. (2007) untersuchten die Beziehung zwischen Joint Attention und der sozioemotionalen Entwicklung im Kleinkindalter. Sie zogen die Daten von 52 Kindern zum Zeitpunkt von 12, 15, 24 und 30 Monaten heran. Es konnte gezeigt werden, dass 12 Monate alte Kinder, die Augenkontakt oder Zeigen häufiger in ihren Joint Attention-Episoden („Initiating Joint Attention“) mit Unbekannten verwendeten, in einem Alter von 30 Monaten bessere Bewertungen zu ihren sozialen Kompetenzen und zu ihrem Verhalten im Generellen erhielten. Des Weiteren bestand ein signifikanter Zusammenhang zwischen höheren Werten der initiiierenden Vorgehensweise in Joint Attention-Episoden von 12 Monate alten Kindern und einem verringerten Berichten der Eltern von externalisierendem Verhalten zum Zeitpunkt von 30 Monaten.

### **4.2.4 Emotionsregulation**

Die Studie von Morales, Mundy, Crowson, Neal und Delgado (2005) beschäftigte sich mit der Beziehung zwischen kindlicher Joint Attention und deren Auswirkungen auf die Emotionsregulation des Kindes. 22 Kinder im Alter von 24 Monaten wurden mit ihren Müttern gemeinsam in die Studie aufgenommen. Es wurde vermutet, dass Kinder, die mit ihren Eltern längere Zeit in Joint Attention Episoden verbringen, mehr selbstgesteuerte Strategien zur Emotionsregulation in Frustrationssituationen verwenden. Gegenstand der zweiten Forschungsfrage war die Beziehung zwischen den bestehenden kindlichen Aufmerksamkeitsfähigkeiten und den Strategien zur Emotionsregulation. Die beiden Vermutungen der Autoren konnten bekräftigt werden. Kinder, die mit ihren Eltern tendenziell mehr Zeit in Joint Attention-Episoden verbrachten, verwendeten tatsächlich häufiger aktivere Strategien zur Emotionsregulation und warteten länger auf das Wiedererlangen eines entfernten Spielzeuges. Die kindliche Fähigkeit der Blickrichtung des Interaktionspartners zu folgen war positiv mit aktiven Spielstrategien assoziiert.

## 5 Systematische Verhaltensbeobachtung

Verhaltensbeobachtung ist ein sehr altes Verfahren, um Informationen über Individuen zu sammeln. Wenn wir sie nicht wissenschaftlich und nicht systematisch betreiben, verwenden wir die wohl älteste Methode des Datensammelns: das, was unsere Erfahrung ausmacht, was wir im täglichen Umgang mit Menschen bewußt oder unbewußt aufnehmen und was wir zu Ansichten, Meinungen, Voreingenommenheiten und Urteilen über »den Menschen« verarbeiten, ist zu einem großen Teil durch unsystematische Verhaltensbeobachtung bestimmt. (Faßnacht, 1979, S. 192)

Kubinger (2009) unterscheidet zwischen Gelegenheitsbeobachtungen und systemischer Verhaltensbeobachtung. Gelegenheitsbeobachtungen spielen nach Kubinger (2009) im Prozess psychologischen Diagnostizierens immer eine begleitende Rolle und können zum Generieren von Hypothesen herangezogen werden. Systematische Verhaltensbeobachtung wird als „zielgerichtete und methodisch kontrollierte Wahrnehmung von Objekten, Ereignissen und Prozessen“ definiert (Häcker & Stapf, 2009, S. 126). Nach Greve und Wentura (1997) ist Systematische Verhaltensbeobachtung eine absichtliche und geplante Unternehmung, die einen gewissen Zweck hat oder ein bestimmtes Ziel verfolgt. Durch Selektion konzentrieren sich die BeobachterInnen auf einen spezifischen Aspekt und blenden andere Elemente aus. Des Weiteren ist die wissenschaftliche Verhaltensbeobachtung immer auf die Auswertung der erhobenen Daten ausgerichtet. Besonders hervorzuheben ist der Umstand, dass eine Theorie oder Hypothese durch die Beobachtung überprüft wird. Außerdem müssen die Ergebnisse der Beobachtung replizierbar sein und verschiedene BeobachterInnen bei demselben vorliegenden Sachverhalt zu denselben Resultaten kommen (Objektivität). Ein großer Vorteil der Systematischen Verhaltensbeobachtung ist nach Kubinger (2009) die hohe ökologische Validität, da sie reales Verhalten erfasst und nicht bloß verbal bekundetes Verhalten sowie den Anforderungen der natürlichen Lebensbedingungen entspricht. Systematische Verhaltensbeobachtung wird nur durch eine passiv teilnehmende, nicht beteiligte BeobachterIn zu bewerkstelligen sein. Eine aktiv teilnehmende BeobachterIn wäre aufgrund von Überforderung höchstens zu einer Gelegenheitsbeobachtung in der Lage. Letzten Endes muss Systemische Verhaltensbeobachtung immer in irgendeine Form der Kategorisierung münden, um die Vielfalt an gewonnenen Informationen zu protokollieren.

## **5.1 Arten der Verhaltensbeobachtung**

Es folgt ein Auszug der, für die hier vorliegende Forschungsarbeit, relevanten Arten der Verhaltensbeobachtung:

### **5.1.1 Unstrukturierte versus strukturierte Beobachtung**

„Der Grad der Systematisierung einer Beobachtung richtet sich nach dem Untersuchungsanliegen ... bzw. nach der Präzision der Vorkenntnisse über den in Frage stehenden Untersuchungsgegenstand. Je genauer man das zu Beobachtende im Prinzip kennt, desto systematischer sollte eine Beobachtung angelegt sein.“ (Bortz & Döring, 2006, S. 267). Die vollstrukturierte oder standardisierte Beobachtung ist eine Methode zur gesicherten und geplanten Datenerhebung, bei der das Vorgehen in einem Beobachtungsplan genau festgelegt ist. Es darf der BeobachterIn nicht überlassen sein, was und wie zu klassifizieren ist. Diese Herangehensweise wird zur Überprüfung von Hypothesen verwendet. Die unstrukturierte oder freie Beobachtung hingegen ist durch ungesteuertes Betrachten ohne spezifische Fragestellung und ohne Vorgabe von Beobachtungsrichtlinien gekennzeichnet. Sie lässt jegliche Form der Reaktion zu und kommt besonderes auf unerforschten Gebieten in Form von induktiven Vorgehensweisen zur Hypothesengenerierung zum Einsatz, hier wäre ein Beobachtungsplan sogar hinderlich. Unsystematische Beobachtungen können auch eingesetzt werden, um in diagnostischen Situationen einen globalen Eindruck zu erhalten (Bortz & Döring, 2006; Jäger & Petermann, 1995).

### **5.1.2 Offene versus verdeckte Beobachtung**

Wenn gegenüber den Versuchspersonen verschwiegen wird, dass sie Teil einer Untersuchung sind und beobachtet werden, so handelt es sich um eine verdeckte Beobachtung. Dieser Umstand ist jedoch ethisch sehr bedenklich und stets kritisch zu hinterfragen. Werden die Personen hingegen über die Beobachtungssituation informiert, so spricht man von einer offenen Beobachtung (Greve & Wentura, 1997). Hierbei ist damit zu rechnen, dass die UntersuchungsteilnehmerInnen über die Beobachtungsintention

spekulieren und sich eventuell konform im Sinne von sozialer Erwünschtheit verhalten (Bortz & Döring, 2006; Greve & Wentura, 1997). Sind reaktive Effekte (siehe Kapitel 5.2.2.1 Reaktivitätseffekte) zu erwarten, so sollte eine verdeckte Beobachtung in Betracht gezogen werden (Bortz & Döring, 2006).

## **5.2 Beobachtungsfehler in der systemischen Verhaltensbeobachtung**

Nach Kubinger (2009) ist das grundsätzliche Problem der Verhaltensbeobachtung darin zu sehen, dass kein Geschehen vollständig erfassbar ist. In der Natur der Sache liegt, dass jeder Rater Verhalten selektiert wahrnimmt und erfasst.

Anschließend ein Auszug möglicher Fehler in der Beobachtung und Beurteilung von Verhalten nach Greve und Wentura (1997):

### **5.2.1 Fehler zu Lasten der BeobachterIn**

Mögliche Fehler liegen in der Verarbeitung der Information durch die BeobachterIn (Greve & Wentura, 1997). Es werden im Anschluss Wahrnehmungsfehler und Interpretationsfehler näher erläutert:

#### **5.2.1.1 Wahrnehmungsfehler**

Greve und Wentura (1997) postulieren, dass das bewusst Wahrgenommene bereits durch Wahrnehmungsfehler verzerrt worden ist. Der Prozess der Beobachtung wird durch sie mitbestimmt. Dazu gehören:

##### ***Konsistenzeffekte***

Darunter versteht man die Tendenz in seinen Urteilen möglichst konsistent und widerspruchsfrei zu sein. Hierzu gehört der *Halo-Effekt*, bei dem eine einzelne Beurteilung vom Gesamteindruck oder einer hervorstechenden Eigenschaften überstrahlt wird und in dessen Abhängigkeit beurteilt wird (Greve & Wentura, 1997; Faßnacht, 1979).

### *Projektion*

Projektion beschreibt die Tendenz von BeobachterInnen im Beobachteten das wiederzuerkennen, was sie von sich selbst kennen. Der Begriff stammt aus der Psychoanalytischen Schule und es muss erwähnt sein, dass für die Bedeutung dieses Beobachtungsfehlers in der Verhaltensbeobachtung keine empirischen Belege bekannt sind (Greve & Wentura, 1997).

### *Erwartungseffekt*

Es herrscht eine unbewusste Theorie bzw. Überzeugung vor, die die Beobachtungen in eine bestimmte Richtung verzerrt. Hierbei neigt die BeobachterIn zu hypothesenkonformen Einschätzungen. Den Erwartungseffekt charakterisiert, dass eine unbefangene BeobachterIn nicht zu denselben Resultaten gelangen würde (Greve & Wentura, 1997).

### *Observer Drift*

Darunter ist die allmähliche Veränderung des Beurteilungsstandards einer BeobachterIn zu verstehen, wie es durch Ermüdung, nachlassende Motivation oder aber auch durch Übung und zunehmende Vertrautheit mit dem Beobachtungsinstrument der Fall sein kann (Greve & Wentura, 1997).

#### *5.2.1.2 Interpretationsfehler*

Interpretationsfehler verzerren das Wahrgenommene, dazu gehören (Greve & Wentura, 1997):

#### *Zentrale Tendenz*

Der Fehler zur zentralen Tendenz kann als Urteilsfehler verstanden werden, bei dem Personen tendenziell extreme Urteile vermeiden und Schätzungen oder Ratings zur Skalenmitte wandern (Greve & Wentura, 1997; Faßnacht, 1995). Nach Bortz und Döring (2006) ist vor allem dann mit der Tendenz zur Mitte zu rechnen, wenn die BeurteilerInnen noch nicht besonders vertraut mit den Objekten sind, eine Situation, die generell zu vermeiden ist. Des Weiteren kann eine Ursache auch darin zu sehen sein, dass es versäumt wurde die Skala an Extrembeispielen zu verankern.

### *Persönliche Tendenzen*

Auf individueller Ebene können persönliche Tendenzen wie die *Tendenz zur Milde* oder die *Tendenz zur Strenge* auftreten. Damit ist eine personenspezifische Tendenz zu besonders milden oder generell eher strengen Urteilen gemeint (Greve & Wentura, 1997). Nach Bortz und Döring (2006) können diese Fehler vermieden werden, indem die BeobachterInnen zuvor auf diese Art der Verzerrungen und Fehler aufmerksam gemacht werden.

### **5.2.2 Fehler zu Lasten der Beobachtung**

#### *5.2.2.1 Reaktivitäts- und Erwartungseffekte*

Reaktivität meint, dass sich Probanden in der Untersuchungssituation anders verhalten, als sie es tun würden, wenn sie keiner Beobachtung ausgesetzt wären. Daraus resultiert das Problem, dass unter Umständen das interessierende Verhalten verändert wird (Greve & Wentura, 1997).

#### *5.2.2.2 Beobachtungsbedingungen*

Greve und Wentura (1997) beschreiben, dass bei einer Untersuchung einer Eltern-Kind-Dyade, die zu einer Therapie angemeldet wurde, typische bzw. problematische Situationen beobachtet werden sollen. Im Labor kann man davon ausgehen, dass diese problematischen interessierenden Verhaltensmuster sehr abgeschwächt oder gar nicht erst auftreten, da es an heimischen Situationen mit Zeitmangel und Stressoren mangelt.

#### *5.2.2.3 Probleme des Beobachtungssystems*

Durch ein fehlerhaftes Beobachtungssystem verliert die Beobachtung an Genauigkeit und Zuverlässigkeit. Artifizielle Kategorien verhindern Differenzierungen, die die BeobachterInnen ausfindig machen könnten. Dies führt dazu, dass wertvolle Informationen nicht aufgedeckt werden können oder verzerrt dargestellt werden. Ein unangemessenes Beobachtungssystem führt zu unreliaablen Urteilen von Seiten der Rater (Greve & Wentura, 1997).

### **5.2.3 Vermeidung von Beobachtungsfehlern**

Nach Greve und Wentura (1997) ist die Vermeidung von Beobachtungsfehlern einerseits in der Vorbeugung von solchen Fehlern zu sehen, was eine sorgfältige und gründliche Planung der Beobachtung bedingt. Dazu gehört laut den beiden Autoren die theoriegeleitete Erstellung eines Beobachtungsschemas, das nicht zu umfangreich ist und in Vorstudien erprobt wurde. Außerdem sollten die Kategorien klar abgetrennt sein und eindeutige Begrifflichkeiten verwendet werden. Die BeobachterInnen sollten vor der Analyse der Untersuchungsbedingung die Möglichkeit haben, das Instrument und die Objekte kennen zu lernen und sich damit vertraut zu machen, beispielsweise in Form einer BeobachterInnenschulung. Die Wahl der Stichprobe spielt auch eine bedeutende Rolle. Greve und Wentura (1997) raten geeignete Objekte für die Beobachtung auszuwählen.

Auf der anderen Seite erwähnen Greve und Wentura (1997) die Auswahl und das Training der eingesetzten BeobachterInnen zur Vermeidung von Beobachtungsfehlern. Sie räumen dem Training der BeobachterInnen einen besonderen Stellenwert ein. Die Kontrolle während des Trainings, ob es sich bei den Ratern um reliable BeobachterInnen handelt, erfolgt über den Vergleich der Bewertungsergebnisse verschiedener Rater. Außerdem sollte genügend Zeit für die Beobachtung zur Verfügung stehen und die Motivation der BeobachterInnen beachtet werden.

## **5.3 Quantifizierungsmethoden**

Nach Faßnacht (1995) gibt es vier Grundtypen der Verhaltensquantifizierung: Quantifizierung von Verhalten als Häufigkeit, Dauer, Intensität und als Ganzes. Er nennt weiters drei Methoden zur Quantifizierung in der Verhaltensbeobachtung:

- Time-sampling
- Event-sampling
- Rating

Die Methode des Time-sampling sowie des Event-sampling quantifizieren beide Verhalten in Dauer und Häufigkeit. Wohingegen das Rating-Verfahren die Intensität von

Verhaltensweisen bestimmt und Verhalten als Ganzes betrachtet (Faßnacht, 1995). Diese drei Möglichkeiten zur Verhaltensquantifizierung seien näher erläutert:

### **5.3.1 *Time-sampling-Verfahren***

Die Methode des Time-sampling wird im Deutschen auch Zeit-Teil-Methode genannt, sie bestimmt die Dauer und die Häufigkeit von Verhaltensweisen. Charakteristisch ist die kurze Zeiteinheit, die abgeschlossen sowie gekennzeichnet ist und als Einheitsintervall verstanden werden kann. Innerhalb dieser Zeiteinheit werden Verhaltensweisen, die im Vorfeld definiert wurden, von den BeobachterInnen registriert und quantifiziert. Pro Zeiteinheit wird für eine Verhaltensweise nur eine einzige Eintragung vorgenommen. Es interessiert also nur ob das Verhalten auftritt oder nicht, nicht aber die Häufigkeit oder Dauer der Verhaltensweise innerhalb dieses Intervalls. Deshalb kommt der Wahl der Länge des Einheitsintervalls eine determinierende Bedeutung zu, genauso wie dem Erstellen von Kodierregeln im Vorfeld. Notationspausen können eingebaut werden, wenn die live-Beobachtung so viel Aufmerksamkeit abverlangt, dass der Rater mit der Beobachtung überfordert ist und aus dem Intervallrhythmus fällt. Dies ist der Fall, wenn die Menge der zu beobachtenden Verhaltensweisen groß ist. Bei Videoaufzeichnungen ist dieses Problem nicht vorhanden, da das Gesehene reproduzierbar ist. In den 70er Jahren waren die Einheitsintervalle durch einen akustischen Piepston gekennzeichnet, der von einem batteriegetriebenen Kassettentonbandgerät kam und in live-Beobachtungen über Kopfhörer abgespielt wurde, um nicht das Umfeld zu irritieren. Des Weiteren gab es Formblätter, auch Checklisten genannt, zum Eintragen der beobachteten Verhaltensweisen in paper-pencil-Format. In den Zeilen waren die möglichen Verhaltensweisen aufgelistet, in den Spalten die Zeitintervalle. Sobald ein Verhalten auftrat wurde in Form eines Strichs eine Verhaltensweise kodiert, jedoch war pro Feld (pro Einheitsintervall und Verhaltensweise) nur ein Strich möglich (Faßnacht, 1979). Heute übernimmt in Videoaufzeichnungen der Computer das Markieren der Zeitintervalle und auch die paper-pencil-Variante ist durch diverse Software-Programme, die auch die Auswertung der Daten erleichtern, abgelöst worden.

### **5.3.2 Event-sampling-Verfahren**

Nach Faßnacht (1979, S. 146) versteht man unter event-sampling „das kontinuierliche genaue zeitliche Erfassen von mindestens einer aber meistens mehreren Verhaltensweisen“. Diese Verhaltensweisen sind im Vorfeld definiert und werden als „natürliche ganze Verhaltenseinheiten in ihrer zeitlichen Ausdehnung gemessen“ (Faßnacht, 1979, S. 146). Auch hier werden Dauer und Häufigkeit in Form von Gesamtzeit, durchschnittliche Zeit, Häufigkeit und relative Zeit einer Verhaltensweise bestimmt. Der charakteristische Unterschied zu time-sampling-Verfahren ist der, dass jedes Auftreten einer Verhaltensweise über eine bestimmte Beobachtungsperiode registriert wird. Die Maße der Häufigkeit und Dauer sind präzise und absolut, wohingegen die Maße der time-sampling-Methode Näherungswerte darstellen, die relativ sind. Event-sampling-Verfahren haben den Vorteil, dass sie – im Gegensatz zur time-sampling-Methode, die Verhaltensweisen in groben Einheitsintervallen abbilden – Verhaltensweisen in kleineren Zeitintervallen erfassen und somit genauer sind. Werden allerdings kleine Einheitsintervalle bei der time-sampling-Methode gewählt, so verblasst dieser Vorteil. Das Event-sampling-Verfahren sollte dann zum Einsatz kommen, wenn die zu erfassenden Verhaltenseinheiten sehr kurz sind und auf Details ausgerichtet sind (Faßnacht, 1979).

### **5.3.3 Rating-Verfahren**

Für das Rating- oder Einstuf-Verfahren ist charakteristisch, dass das Ausmaß einer Verhaltensweise abgeschätzt und auf einer Schätz-Skala eingetragen wird (Faßnacht, 1979). Ratingskalen messen zumindest auf Ordinalskalen-Niveau und häufig auf Intervallskalen-Niveau. Es werden Häufigkeiten und Intensitäten der zu beobachtenden Verhaltensweisen betrachtet. Die BeobachterIn fällt Urteile in ausgedehnten Zeiteinheiten durch Integration mehrerer Verhaltensweisen (Greve & Wentura, 1997). Auch Faßnacht (1979) nennt diesen Umstand, nämlich dass die quantitative Beurteilung von Verhalten oft in Kombination mit Urteilen über andere Eigenschaften gefällt wird, als Nachteil, da es zu Beobachtungsfehlern kommen kann (vgl. Kapitel 5.2.1.1 Wahrnehmungsfehler). Außerdem erläutert er, dass die intersubjektive Übereinstimmung der geschätzten Urteile als gering erachtet werden muss. Bei Rating-Verfahren werden oft Beurteilungen aus dem Gedächtnis getätigt, was zu einer verzerrenden und lückenhaften Erinnerung und in Folge

dessen zu einem fehlerbehafteten Urteil führen kann. Der größte Vorteil von Rating-Verfahren liegt in der ökonomischen Durchführbarkeit, da ein geringer Aufwand technischer und organisatorischer Natur für eine sehr simple Handhabung sorgt (Faßnacht, 1979; Greve & Wentura, 1997). Greve und Wentura (1997) raten aufgrund der erhöhten Wahrscheinlichkeit für mögliche Beurteilungsfehler bei Ratingskalen zu einem intensiven BeobachterInnentraining sowie dem Nachweis der BeobachterInnenübereinstimmung. Weiters fordern sie den Nachweis von Validität und Reliabilität in besonders starkem Ausmaß.

### **5.4 Beobachtungssysteme**

Es kann zwischen vier verschiedenen Beobachtungssystemen unterschieden werden: Verbal-, Nominal-, Dimensional- und Struktural-Systemen, wobei diese sich nicht gegenseitig ausschließen und oftmals zur Beschreibung desselben Untersuchungsgegenstandes eingesetzt werden können (Faßnacht, 1995).

#### **5.4.1 Verbal-Systeme**

Bei Verbal-Systemen handelt es sich um Beschreibungssysteme, die sich der gesprochenen Sprache bedienen, um Beobachtungsinhalte abzubilden. Es können vier Formen innerhalb der Verbal-Systeme unterschieden werden: Tagebuchaufzeichnungen, Verlaufsprotokolle, die Methode der kritischen Vorfälle und die Ereignis-Beschreibung (Faßnacht, 1995).

#### **5.4.2 Nominal-Systeme**

Nominal-Systeme bedienen sich auch verbaler Anzeigesysteme, allerdings mit einem eingeschränkten überblickbaren Zeichenkatalog und dem völligen Verzicht auf Grammatik. Es handelt sich hierbei um eine einfache qualitative Klassifikation von Verhalten, mit der Verwendung vordefinierter semantischer Spracheinheiten. Es kann zwischen zwei Arten von Nominal-Systemen unterschieden werden (Faßnacht, 1995):

### 5.4.2.1 Indexsysteme

Nach Faßnacht (1995) enthält ein Indexsystem Zeichen, die untereinander sowohl kompatibel als auch inkompatibel sein können. Es handelt sich hierbei um ein offenes Beschreibungssystem und es ist möglich, dass die Verhaltenseinheiten das übergeordnete Verhalten nicht vollständig abdecken. Im Indexsystem können jederzeit mehrere Verhaltenseinheiten in einem Beobachtungsfeld bezeichnet werden. Somit können zu jedem beliebigen Zeitpunkt der Beobachtung keine, eine oder mehrere Indizes gewertet werden.

### 5.4.2.2 Kategoriensysteme

In einem Kategoriensystem wird jedes auftretende Verhalten bewertet, das bedeutet, dass jede Verhaltenseinheit in eine vorher festgelegte Kategorie fallen muss. In diesem System muss die BeobachterIn ständig Urteile abgeben und ist somit zu jedem Zeitpunkt beschäftigt, Einheiten zu registrieren und den verschiedenen Kategorien zuzuordnen. Es handelt sich hierbei somit um ein logisch geschlossenes, vollständiges Beschreibungssystem. Übergeordnetes Verhalten wird zu jedem Zeitpunkt durch eine einzelne Verhaltenskategorie vollständig abgedeckt, die einzelnen Kategorien schließen sich gegenseitig aus (Faßnacht, 1995).

### 5.4.3 Dimensional-Systeme

Dimensional-Systeme sind direkte Quantifizierungssysteme, deren Einheiten gleich sind. Die Beschreibung der Beobachtungseinheiten erfolgt entsprechend einer räumlichen, geradlinigen Ordnung in Bezug zu einem Referenzpunkt. Die einzelnen Dimensionen schließen sich gegenseitig aus. Systeme, die ein quantitatives Urteil fällen, nennt man auch Skalen. Dimensional-Systeme können sich drei verschiedener Skalen bedienen: Ordinalskalen, Intervallskalen oder Rationalskalen (Faßnacht, 1995).

#### **5.4.4 Struktural-Systeme**

Das Struktural-System lässt sich durch das Zuordnen der Einheiten, die gleich oder verschieden sein können, charakterisieren. Es kann auch als Verallgemeinerung des dimensional Systems betrachtet werden, da das strukturelle System nicht nur eine Gerade, sondern jegliche Art von räumlichem Gebilde als Bezugsordnung heranziehen kann. Beobachtungsinhalte können somit durch die Anordnung von Einheiten statt auf einer Geraden, in einer figürlichen Bezugsordnung dargestellt werden. Die häufigste figürliche Bezugsordnung des Struktural-Systems ist die Hierarchiestruktur oder das Baumdiagramm. (Faßnacht, 1995).

## 6 Reliabilitätsanalysen

Um festzustellen, ob sich das Beobachtungsinstrument INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) auf Situationen des klinischen Alltags anwenden lässt, wurde eine Reliabilitätsanalyse durchgeführt. Die hierfür verwendeten statistischen Maße waren die Intraklassenkorrelation und Cohens Kappa. Außerdem wurden die Prozentuellen Übereinstimmungen berechnet. Da die Analyse der Interrater-Reliabilität auch immer eine Analyse der Objektivität darstellt, wird im folgenden Kapitel näher darauf eingegangen.

### 6.1 Gütekriterien

#### 6.1.1 Reliabilität

Unter Reliabilität wird die Messgenauigkeit verstanden, mit der ein Test misst, unabhängig davon, was er misst. Hierbei ist die Zuverlässigkeit gemeint, mit der ein Verfahren bei wiederholter Messung und unter gleichen Bedingungen zu demselben Ergebnis gelangt (Rost, 2004). Um eine möglichst hohe Reliabilität erzielen zu können, müssen „den Beurteilern möglichst exakt gleiche, standardisierte Informationen über die Objekte ... bei der Beurteilung vorliegen“ (Wirtz & Caspar, 2002, S. 31). Wirtz und Caspar (2002) geben die Voraussetzung der Standardisierung als eine der wichtigsten für eine reliable Beurteilung an.

$$\text{Reliabilität} = \frac{\text{wahre Varianz}}{\text{beobachtete Varianz}} = \frac{\text{Var}(T_x)}{\text{Var}(X)}$$

Nach dieser Definition (Rost, 2004, S. 38) ist die Reliabilität als Varianzanteil, nämlich als das Verhältnis der wahren Varianz zur beobachteten Varianz zu verstehen. Die wahre Varianz ist stets kleiner als die beobachtete Varianz, sodass die Reliabilität zwischen 0 und 1 liegt.

Wirtz und Caspar (2002) weisen darauf hin, dass es nicht eine beste Methode zur Reliabilitätsbestimmung gibt. Vielmehr raten sie dazu, mehrere Koeffizienten anzuführen, da jeder einen anderen Teilaspekt abbildet und somit ein vollständiges Abbild der Daten

gewährleistet werden kann. Um das tatsächliche Raterverhalten gut beurteilen zu können, empfehlen sie (zusätzlich zur Berechnung der Interrater-Reliabilität) die Angabe einer Übereinstimmungsmatrix.

### **6.1.2 Objektivität**

Nach Rentzsch und Schütz (2009) gibt die Objektivität eines Tests „das Ausmaß an, in dem die Auswertung und die Interpretation eines Tests unabhängig vom Untersucher sind und in dem die Testsituation unter kontrollierten Bedingungen abläuft“ (S. 209). Sie betonen also nicht nur die BeobachterInnenunabhängigkeit, sondern weisen auch auf die kontrollierten Bedingungen einer Standardisierung hin. In nicht-standardisierten Prozessen nennen sie die Möglichkeit der BeobachterInnenschulung zur Erhöhung der Objektivität.

Für die hier vorliegende Arbeit sind vor allem die TestleiterInnenunabhängigkeit und die Verrechnungssicherheit von Bedeutung. TestleiterInnenunabhängigkeit ist dann gegeben, wenn das Testverhalten eines Probanden und damit die Testergebnisse unabhängig von zufälligen oder systematischen Verhaltensvariationen aller möglichen TestleiterInnen sind (Kubinger, 2009). Rentzsch und Schütz (2009) sprechen hierbei von der Durchführungsobjektivität, die vor allem durch Standardisierung der Testsituation gegeben ist. Verrechnungssicherheit beschreibt den Umstand, dass die Reglementierung eines Testverfahrens, wie die einzelnen Testleistungen zu numerischen oder kategorialen Testwerten zu verrechnen sind, exakt festgelegt ist und dass außerdem jede AuswerterIn zu denselben Ergebnissen kommt (Kubinger, 2009).

### **6.1.3 Zusammenhang von Objektivität und Reliabilität**

Genau genommen ist die ‚Interrater-Reliabilität‘ ein Maß der Objektivität, da die Berechnung der Übereinstimmung zwischen zwei Instrumenten auch über die beiden Rater erfolgen kann. Die Reliabilität des Verfahrens wird anhand des Vergleichs der Bewertungen durchgeführt. Je ähnlicher die Ergebnisse, desto größer die Übereinstimmung zwischen den BeobachterInnen, desto reliabler das Instrument (Greve & Wentura, 1997). Auch Faßnacht (1979) gibt an, dass Reliabilitätsbestimmungen rein formal betrachtet, einer

Objektivitätsüberprüfung entsprechen, da sie auf den Vergleich der Übereinstimmungen hinaus laufen. Der Begriff der Interrater-Reliabilität und der der Objektivität können und werden meistens synonym verwendet. Nach Rost (2004) ist die Objektivität eine Voraussetzung für die Reliabilität. Gemeint ist damit, dass ein nicht objektives Verfahren auch keine hohe Messgenauigkeit (Reliabilität) aufweisen kann.

## 6.2 Intraklassenkorrelation

### 6.2.1 Allgemeines zur Intraklassenkorrelation

Die Intraklassenkorrelation (ICC), ein varianzanalytisches Modell, das als Maß zur Aufklärung der Varianz herangezogen wird, gibt Aufschluss über die Stärke des Zusammenhangs zwischen den Beurteilungen zweier Rater. Ein Wert von 0 bedeutet, dass kein Zusammenhang zwischen den Urteilen der Rater besteht. Je mehr sich die ICC dem Wert 1 annähert, desto höher ist die Reliabilität der Urteile. Reliabilität bezeichnet die Zuverlässigkeit oder Genauigkeit einer Messung, was erfüllt ist, wenn verschiedene Rater mit dem gleichen Wissen zu sehr ähnlichen Urteilen kommen (Wirtz & Caspar, 2002).

Reliabilitätsmaße sind Maße, die zuverlässig Unterschiede zwischen den zu beurteilenden Probanden identifizieren. Gibt es keine Unterschiede zwischen den Probanden, so kann das psychologische Verfahren auch keine Unterschiede zuverlässig eruieren. Wo keine Varianz vorliegt, kann auch keine Varianz aufgeklärt werden. Daher wird in der Literatur empfohlen, sowohl Maße der Übereinstimmung (Prozentuelle Übereinstimmung) als auch Maße der Reliabilität (Intraklassenkorrelation) zu bestimmen. Eine hohe Reliabilität bedeutet nicht zwangsläufig, dass die Rater ähnliche Urteile gefällt haben. Sowie eine hohe Übereinstimmung nicht selbstredend an eine hohe Reliabilität geknüpft sein muss. Bei niedriger Reliabilität und hoher Übereinstimmung, besteht die Möglichkeit der geringen Merkmalsvarianz in der zu Grunde liegenden Stichprobe. Nur bei hoher Reliabilität und hohen Übereinstimmungswerten, kann davon ausgegangen werden, dass die Beurteilungen durch die Rater zuverlässig waren. Der Nachteil der Übereinstimmungsmaße ist, dass sie nicht zufallsbereinigt sind. Außerdem fordern sie eine exakte Gleichheit der beurteilten Merkmalsausprägungen, wohingegen Reliabilitätsmaße die relative Lage der Werte zum Mittelwert aller Rater fordern (Wirtz & Caspar, 2002).

Die Reliabilität eines psychologischen Verfahrens ist unter anderem von der Art der Stichprobe abhängig. Je größer die Mittelwertsunterschiede zwischen den beobachteten Personen hinsichtlich eines bestimmten Merkmals, je heterogener also die Stichprobe, desto höher fällt auch die Reliabilität aus. Bestehen keine markanten Unterschiede zwischen den Mittelwerten der Probanden, so wird es für das Messinstrument schwierig, eine bedeutsame Reliabilität zu ermitteln. Denn je weniger Varianz in den wahren Werten der Probanden auftritt, desto schwieriger ist es, diese Varianz aufzuklären. Bei geringfügiger Reliabilität ist daher nicht automatisch die Fähigkeit der Rater verantwortlich zu machen, da dies möglicherweise auf eine homogene Stichprobe zurückzuführen ist. Aus diesem Grund kann eine geringe ICC bei einer homogenen Stichprobe als akzeptabel angesehen werden. Reliabilitätsmaße bilden somit nicht nur ab, wie genau die Rater beurteilt haben, sondern spiegeln indirekt auch die Variabilität eines Merkmals in der untersuchten Stichprobe wieder. Es gilt im Allgemeinen, dass bei kleinen Unterschieden innerhalb eines Probanden (Homogenität der Ratings innerhalb der Personen) und großen Unterschieden zwischen den Probanden (Heterogenität der Mittelwerte zwischen den Personen), von einer guten Reliabilität ausgegangen werden kann (Wirtz & Caspar, 2002).

Nach Wirtz und Caspar (2002) gibt es nicht eine beste Methode zur Ermittlung der Reliabilität. Es bestehe nämlich beim Heranziehen von nur einem Koeffizienten die Gefahr, dass das Datenmaterial verschleiert dargestellt wird. Stattdessen ist es viel aufschlussreicher mehrere Koeffizienten zu eruieren, um die Datenlage möglichst vollständig abzubilden.

### **6.2.2 Modellauswahl in der Intraklassenkorrelation**

Es wird zwischen der einfaktoriellen und der zweifaktoriellen Intraklassenkorrelation (ICC) unterschieden. Wurden alle Probanden von denselben Ratern bewertet, so entscheidet man sich für die zweifaktorielle ICC, wurden nicht alle Personen von denselben Ratern beurteilt, so zieht man die einfaktorielle ICC heran (Wirtz & Caspar, 2002).

Das Modell der unjustierten Intraklassenkorrelation (absolute Übereinstimmung) ist das strengere Maß, da es Mittelwertsunterschiede zwischen den Ratern als Fehlerquelle betrachtet und somit zulasten der Reliabilität verrechnet. Als Übereinstimmungsmaß wird die Ähnlichkeit der absoluten Werte gefordert. Die justierte Intraklassenkorrelation (Konsistenz) hingegen ist vom Effekt unterschiedlicher Mittelwerte bereinigt. Die Mittelwerte der Rater dürfen verschieden sein. Als Konsistenzmaß verlangt die  $ICC_{just}$ , dass sich das Verhältnis der Abweichungen zwischen den einzelnen Objekten ähneln sollte, ohne die Gleichheit der absoluten Werte zu fordern. Aus diesem Grund nimmt die  $ICC_{just}$  tendenziell höhere Werte an als die  $ICC_{unjust}$ . Bei einer fälschlicherweise verwendeten  $ICC_{just}$  wird die Reliabilität überschätzt, hingegen bei der  $ICC_{unjust}$  unterschätzt (Wirtz & Caspar, 2002).

Im nächsten Schritt muss entschieden werden, ob der Rater als ‚zufällig‘ oder ‚gemischt‘ angesehen werden soll. Wenn die Rater eine repräsentative Stichprobe aller Rater darstellen, für die die Reliabilitätsaussage Gültigkeit besitzen soll, handelt es sich um die Modellentscheidung ‚zufällig‘ (‚random‘). Die Forschungsfrage würde sich durch das Heranziehen anderer BeurteilerInnen nicht verändern. Im ‚gemischten‘ (‚fixed‘) Modell bezieht sich die Reliabilitätsaussage lediglich auf die untersuchten Rater und kann nicht verallgemeinert oder auf andere Populationen übertragen werden (Wirtz & Caspar, 2002).

### 6.3 Cohens Kappa

Cohens Kappa dient zur Quantifizierung der BeobachterInnenübereinstimmung und ist das am häufigsten verwendete Maß dafür (Wirtz & Caspar, 2002). Normalerweise sind die Daten, nominalskaliert, sie können aber auch ordinal skaliert sind (Bakeman & Quera, 2011). Bei Cohens  $\kappa$  handelt es sich um ein zufallskorrigiertes Übereinstimmungsmaß, das auf der prozentuellen Übereinstimmung der BeobachterInnen basiert. Allerdings hat  $\kappa$  gegenüber der prozentuellen Übereinstimmung den Vorteil, dass hierbei das Verhältnis der beobachteten zu der bei Zufall erwarteten Übereinstimmung berücksichtigt wird.  $\kappa$  liefert eine standardisierte Maßzahl, die zwischen -1 und +1 liegen kann (Wirtz & Caspar, 2002).

Landis und Koch (1977, S. 165) schlagen folgende Kennwerte zur Gliederung der Stärke der Übereinstimmung mittels Cohens  $\kappa$  vor:

| <i>Kappa</i> | <i>Stärke der Übereinstimmung</i> |
|--------------|-----------------------------------|
| 0.00-0.20    | gering                            |
| 0.21-0.40    | ausreichend                       |
| 0.41-0.60    | mittelmäßig                       |
| 0.61-0.80    | beträchtlich                      |
| 0.81-1.00    | (nahezu) perfekt                  |

Es kann zwischen dem *time-unit kappa* und dem *event-based kappa* unterschieden werden. Bei den Berechnungen zum *event-based kappa* fokussiert sich das Auswertungsprogramm nicht auf die Zeiteinheiten, wie beim *time-unit kappa*, sondern zieht die einzelnen Events für den Vergleich heran. Der time-unit basierte Algorithmus tendiert dazu die wahre Anzahl an Entscheidungen durch die Rater zu überschätzen, der Event basierte Algorithmus hingegen unterschätzt (Bakeman, Quera & Gnisci 2009).

In den Kontingenztabellen bezeichnen Werte der Diagonale die Anzahl der Übereinstimmungen der beiden BeobachterInnen und die Werte abseits der Diagonale die Nichtübereinstimmungen an. BeobachterIn 1 vergibt also eine andere Kodierung als BeobachterIn 2 (Bakeman, Quera & Gnisci, 2009). Die Spalte und Zeile mit der Beschriftung „nil“ in der Event-basierten Vorgehensweise drückt den *commission-omission error* aus. Hierbei wird jegliches Ereignis identifiziert, das nur von einem der beiden BeobachterInnen mit einer Kodierung versehen wurde. Kodiert BeobachterIn 1 ein Ereignis, nicht jedoch BeobachterIn 2, nennt man diesen Umstand *deletion*, es handelt sich um den *omission error*. Wird eine Kodierung jedoch nur von BeobachterIn 2 getätigt, ist dies der *commission error*, was auch als *insertion* bezeichnet wird (Bakeman & Quera, 2011).

## 6.4 Prozentuelle Übereinstimmung

Die Prozentuelle Übereinstimmung ist ein Maß der Übereinstimmung zwischen mehreren Messwertreihen. Sie gibt den prozentuellen Anteil an, in dem zwei oder mehrere Rater das gleiche Urteil abgeben (Wirtz & Caspar, 2002).

$$P\ddot{U} = \frac{\text{Anzahl der beobachteten Übereinstimmungen}}{\text{Anzahl der Objekte}} \cdot 100$$

(Wirtz & Caspar, 2002, S. 48)

Die Prozentuelle Übereinstimmung stellt ein sehr anschauliches Maß zur Interpretation dar, allerdings sind die Resultate nicht zufallsbereinigt (Wirtz & Caspar, 2002).

## **II. EMPIRISCHER TEIL**

### **7 Hintergrund der Untersuchung**

#### **7.1 Zentraler Ausgangspunkt und Forschungsinteresse**

INTAKT (Hirschmann et al., 2011) als videogestütztes Beobachtungsinstrument zur Erfassung der elterlichen Interaktionsqualität wurde bereits bei Mutter-Kind-Dyaden des Pflegekinderwesens eingesetzt. Im Interesse der Autoren dieses Verfahrens steht natürlich auch die Anwendung von INTAKT in psychologischen Praxen und Spitälern, da nun auf Grundlage wissenschaftlicher Kriterien Entscheidungen getroffen und psychologische Diagnostik für Eltern-Kind-Beziehungen im Altersbereich der Kleinkinder und Kindergartenkinder betrieben werden könnte.

Im Zentrum dieser Diplomarbeit steht die Erprobung des Beobachtungsinventars INTAKT auf die Altersgruppe der 1- bis 3-Jährigen, ist es ursprünglich auf den Altersbereich der 3- bis 6-Jährigen zugeschnitten. Außerdem wird evaluiert, ob sich INTAKT, da es für eine standardisierte Bastel- bzw. Spiel-Situation konzipiert wurde, auch auf andere Situationen im klinischen Setting anwenden lässt.

#### **7.2 Kooperation mit der Säuglingspsychosomatik des Wilhelminenspitals**

Im Rahmen einer Kooperation mit der Säuglingspsychosomatik des Wilhelminenspitals sind den beiden DiplomandInnen Valai und Gogl Interaktionsvideos des Pavillons 5 für die Dauer ihrer Arbeit zugänglich gemacht worden. Es handelt sich bei den PatientInnen vornehmlich um Mütter, die gemeinsam mit ihren Kindern stationär aufgenommen wurden. Aufnahmegründe seitens des Kindes können Fütterungsprobleme, exzessives Schreiverhalten, Gedeihstörungen oder Verhaltensprobleme sein. Elternbezogene Aufnahmegründe stellen meistens akute Belastungsreaktionen oder postpartale Depressionen dar. Ebenfalls können Familienprobleme wie es bei Mehrlingen,

minderjährigen Müttern, Adoptionen oder Pflegefamilien der Fall sein kann, Gründe einer stationären Aufnahme darstellen.

Es gibt auf der Station der Säuglingspsychosomatik vier PatientInnenzimmer mit insgesamt acht Betten. Die durchschnittliche Aufenthaltsdauer liegt zwischen zwei und vier Wochen. Es gibt jährlich – gemeinsam mit der Tagesklinik, in der ambulant betreut wird – circa 400 Aufnahmen bei einer jährlichen Auslastung von 100 %. Das Betreuungsteam besteht aus einer Fachärztin für Kinder- und Jugendheilkunde, einer Klinischen Psychologin und Psychotherapeutin, dem Pfllegeteam bestehend aus 9 diplomierten Kinderkrankenschwestern, einer auszubildenden Ärztin und einer Klinischen Psychologin.

Spezifikum der Säuglingsdiagnostik stellt die videogestützte Interaktionsdiagnostik zwischen Elternteil und Kind dar. Im Rahmen der Videodiagnostik wird das Verhalten des Kleinkindes und des Elternteils dokumentiert und in die stationäre Therapie mit einbezogen. Es werden folgende Formen der Psychotherapie praktiziert: Eltern-Kind-Interaktionstherapie, Analytische Gesprächstherapie, Krisenintervention, Nutzung gruppendynamischer Prozesse, verhaltenstherapeutische Maßnahmen und videounterstützte therapeutische Intervention.

Von Seiten der Psychologin Dr. Christine Sonn-Rankl und der Ärztin Dr. Josephine Schwarz-Gerö der Säuglingspsychosomatik von Pavillon 5 besteht reges Interesse an einer neuen Form der Diagnostik zur Einstufung von Eltern-Kind-Interaktionen bzw. der Beziehungsqualität.

### **7.3 Forschungsfragestellung**

Es sei ausdrücklich erwähnt, dass die Untersuchung eine erstmalige Erprobung des Beobachtungsinstruments INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) auf Situationen des Klinischen Alltags darstellt, ist das Inventar ursprünglich auf eine standardisierte Spiel- und Bastelsituation ausgelegt. Die von der Säuglingspsychosomatik des Wilhelminenspitals (Pavillon 5) zur Verfügung gestellten Interaktionsvideos fanden in drei unterschiedlichen Settings statt: Essens-, Pflege- und

Spiel-Situationen. Es gab weder standardisierte Instruktionen noch Settings, da es sich um eine retrospektive Datenanalyse handelt, die mit bereits in der Vergangenheit gedrehten Videos arbeitet. Des Weiteren fand die Untersuchung Anwendung auf den Altersbereich der 1- bis 3-Jährigen, INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) ist jedoch ursprünglich für den Altersbereich der 3- bis 6-Jährigen ausgelegt.

Die hier vorliegende Arbeit vergleicht die Bewertungen zweier unabhängiger BeobachterInnen bezüglich der Skalen *Feinfühligkeit*, *Rückmeldung* und *Joint Attention*. Eine Reliabilitätsanalyse klärt den Grad der Übereinstimmung der beiden Rater pro Kategorie. Es erfolgt außerdem eine genauere Betrachtung und Analyse der Diskrepanzen für jede der drei Skalen. Hieraus leiten sich folgende Überlegungen ab:

- (a) Welche Reliabilitäten ergeben sich aufgrund der BeobachterInnenübereinstimmung für die einzelnen Skalen (*Feinfühligkeit*, *Rückmeldung* und *Joint Attention*) des Instruments?
- (b) Gibt es unterschiedlich hohe Reliabilitäten für die einzelnen Situationen (Essen, Pflege, Spiel) auf den einzelnen Dimensionen (*Feinfühligkeit*, *Rückmeldung*, *Joint Attention*)?

## 8 Untersuchungsinstrument INTAKT

Zur Datenerhebung wurde das Beobachtungsinventar INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) herangezogen. Dieses Instrument wurde entwickelt, um die Beziehungs- bzw. Interaktionsqualität zwischen Mutter und Kind einstufen zu können. Es handelt sich dabei um eine systematische Verhaltensbeobachtung. Die Interaktionssequenz wird anhand von drei Verhaltensdimensionen bewertet: *Feinfühligkeit*, *Rückmeldung* und *Joint Attention*. Es sei erwähnt, dass INTAKT ursprünglich für eine standardisierte Situation mit einer standardisierten Instruktion konzipiert wurde. Die Beobachtung beruht auf einer Bastelaufgabe und einer freien Spielsituation, wobei der Mutter-Kind-Dyade sowohl Bastelutensilien als auch eine Spielkiste zur Verfügung gestellt werden.

### 8.1 Feinfühligkeit

Die mütterliche Feinfühligkeit ist für die Beziehungsqualität maßgebend und stellt die Basis für eine sichere Bindung dar (Ainsworth, Bell & Stayton, 1974; Ainsworth, Blehar, Waters & Wall, 1978). Weiters verstehen Ainsworth et al. (1974) unter Feinfühligkeit die Fähigkeit der Mutter, die Signale des Kindes wahrzunehmen, richtig zu interpretieren und darauf angemessen und prompt zu reagieren.

Die Skala *Feinfühligkeit* wird mittels 7-stufiger Rating-Skala erfasst: (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*. Im Manual werden die Ankerpunkte 1, 3, 5 und 7 genauer definiert, die dazwischenliegenden Abstufungen 2, 4 und 6 werden kodiert, wenn sich das beobachtete Verhalten zwischen zwei definierten Ankerpunkten befindet. Sollte die Mutter und/oder das Kind mind. 50% des Intervalls nicht sichtbar sein, so wird (8) *Unkodierbar – Feinfühligkeit* gewertet. Das Rating wird mit der Methode des Time-Samplings kombiniert, die Einstufung der Feinfühligkeit erfolgt demnach alle 2 Minuten.

Es folgt zum besseren Verständnis der Kategorie *Feinfühligkeit* ein Auszug aus dem Manual (Hirschmann, Aigner, Deimann, Kastner-Koller, 2013), aus dem die beiden äußersten Ankerpunkte geschildert werden:

*Ankerpunkt 7: Sehr hohe Feinfühligkeit*

Die Bezugsperson orientiert sich gänzlich am Kind, stimmt alle Handlungen auf die Bedürfnisse des Kindes ab und stellt ihre eigenen Ideen in den Hintergrund. Ihr Verhalten zeichnet sich durch prompte und angemessene Reaktionen aus. Durch kontinuierliche und motivierende Rückmeldung sowie emotionale Beteiligung werden kindliche Aktivitäten angeregt. Die Bezugsperson setzt Handlungen nur in gemeinsamer Vereinbarung oder mit der offensichtlichen Absicht, dem Kind in der Realisierung seiner Wünsche zu helfen. Der Sprachstil ist dem Entwicklungsstandes des Kindes angepasst, die Interaktion harmonisch und reibungslos (Hirschmann et al., 2013).

*Ankerpunkt 1: Sehr geringe Feinfühligkeit*

Die Bezugsperson zeigt kein Interesse daran, sich einer gemeinsamen Sache mit dem Kind zu widmen oder sich am Spielgeschehen zu beteiligen. Es fällt der Bezugsperson schwer den Blickwinkel des Kindes einzunehmen und ihr Handeln auf das Kind abzustimmen oder eine allgemeine Freude am Spielen zu zeigen. Die Bezugsperson orientiert sich an den eigenen Ideen und verfolgt ihre eigenen Handlungsabsichten. Sie fragt nicht um Erlaubnis, bevor sie ins Spielgeschehen eingreift und regt weder die kindlichen Aktivitäten an, noch greift sie kindlichen Initiativen auf. Nonverbale Signale des Kindes, wie Zeigegesten oder Blickkontakt, werden nicht wahrgenommen. Hat das Kind Schwierigkeiten, muss es sich verbal Aufmerksamkeit verschaffen, was nicht immer gelingt. Die Bezugsperson zeigt keinerlei emotionale Beteiligung und gibt auch keine Rückmeldungen. Der Sprachstil der Bezugsperson ist nicht kindgemäß und sie reagiert oft unangemessen und ungeduldig (Hirschmann et al., 2013).

## 8.2 Rückmeldung

Die mütterliche Rückmeldung dient dem Kind als eine Art Bezugsnorm, durch die es die Angemessenheit seiner Handlungen einstufen kann (Kelley, Brownell & Campbell 2000; Harter, 1978). Je nach Art der Rückmeldung kann die Mutter förderlich oder hemmend auf die Kindesentwicklung wirken (Kelley et al., 2000).

Bei der Kategorie *Rückmeldung* handelt es sich um ein Kategoriensystem, da es ein logisch geschlossenes und vollständiges Beschreibungssystem ist. Die verschiedenen Verhaltenskategorien schließen sich gegenseitig aus und das übergeordnete Verhalten wird zu jedem Zeitpunkt durch einzelne Verhaltenskategorien vollständig abgedeckt. Die Variable *Rückmeldung* wird mittels Event-Sampling-Verfahren bewertet. Es wird fortlaufend jede Verhaltensoption kodiert, sobald sie auftritt.

Es gibt die Möglichkeit im Zuge von INTAKT (Hirschmann et al., 2013), *Rückmeldung* wie folgt zu bewerten:

### *(a) Positive Rückmeldung*

Auf emotional positive Art werden lobende Aussagen bezüglich des Kindes oder dessen Handlungen getätigt. Sie sind als Zustimmung und Verstärkung wahrnehmbar.

### *(b) Korrigierende Rückmeldung*

Es werden Äußerungen in einem neutralen oder freundlichen Tonfall getätigt, die dem Kind Informationen liefern, wie es seine Handlungen verbessern kann. Es werden Informationslücken des Kindes gefüllt, Entscheidungshilfen gegeben und das Kind in seiner Handlung unterstützt. Korrekturen, die bereits vor dem Durchführen einer Handlung getätigt werden, können allerdings nicht als korrigierende Rückmeldung gewertet werden.

### *(c) Negative Rückmeldung*

Hierunter werden alle auf emotional negative Art und Weise hervorgebrachten Äußerungen verstanden, die sich auf das Kind als Person oder dessen Handlungen beziehen. Die Ablehnung oder Unzufriedenheit der Bezugsperson kommt zum Ausdruck. Da hier der Tonfall eine bedeutende Rolle spielt, können beispielsweise sarkastische Aussagen in diese Kategorie fallen.

*(d) Unkodierbar – Rückmeldung*

Die Situation kann nicht eingeschätzt werden, da die Mutter den Raum verlassen hat oder das Gesagte nicht verständlich ist. Auch wenn eine dritte Person den Raum betritt und mit der Mutter spricht, ist dies als unkodierbar zu werten.

*(e) Keine Rückmeldung*

Die Bezugsperson tätigt keine Aussage bezüglich des Kindes als Person, dessen Handlungen oder Ergebnisse.

### **8.3 Joint Attention**

Die Dimension *Joint Attention* stellt die kindliche Fähigkeit dar, seine Aufmerksamkeit zwischen einem sozialen Partner und einem Objekt oder Ereignis zu vereinbaren (Morales et al., 2000). Aufmerksamkeitslenkung erwies sich als besonders förderlich für Kinder unter sechs Monaten, ab acht Monaten zeigte sich eine aufmerksamschaftsfolgende Art der Mutter als besonders dienlich (Saxon, Colombo, Robinson & Frick, 2000).

Bei der Kategorie *Joint Attention* handelt es sich um ein Kategoriensystem, da es ein logisch geschlossenes und vollständiges Beschreibungssystem ist. Die verschiedenen Verhaltenskategorien schließen sich gegenseitig aus und das übergeordnete Verhalten wird zu jedem Zeitpunkt durch einzelne Verhaltenskategorien vollständig abgedeckt. Die Variable *Joint Attention* wird mittels Event-Sampling-Verfahren bewertet. Es wird fortlaufend jede Verhaltensoption kodiert, sobald sie auftritt.

Das Manual (Hirschmann et al., 2013) unterscheidet zwischen folgenden Ausprägungen auf dieser Dimension:

*(h) Aktive Aufrechterhaltung der Joint Attention auf der Handlungsebene*

Die Bezugsperson setzt Handlungen und übernimmt eine handelnde Rolle im gemeinsamen Spielgeschehen. Sie beteiligt sich aktiv an der Beschäftigung des Kindes, um die Joint Attention des Kindes aufrecht zu erhalten, wobei das Interesse und der Aufmerksamkeitsfokus des Kindes berücksichtigt werden.

*(v) Aktive Aufrechterhaltung der Joint Attention auf der verbalen Ebene*

Die Aufrechterhaltung der Joint Attention seitens der Mutter erfolgt rein auf der verbalen Ebene, ohne sich aktiv am Spielgeschehen zu beteiligen.

*(p) Passive Aufrechterhaltung*

Diese Ausprägung ist gekennzeichnet durch bloßes Beobachten des Kindes durch die Bezugsperson. Sie trägt dadurch zur Aufrechterhaltung der kindlichen Joint Attention bei, indem sie sich bewusst gegen ein Eingreifen entscheidet.

*(l) Aufmerksamkeitslenkung im laufenden Spiel*

Die Bezugsperson versucht durch Befehle, Anweisungen, Verbote oder Wünsche die Aufmerksamkeit und das Verhalten des Kindes zu beeinflussen. Dies kann z.B. durch Behaaren auf der Gesprächsebene oder aber durch Manipulation ins Spielgeschehen erfolgen.

*(w) Aufmerksamkeitswechsel zu einem neuen Spiel*

Durch direkte Versuche lenkt die Bezugsperson die Aufmerksamkeit des Kindes von einer Beschäftigung oder einem Objekt auf ein(e) andere(s). Sie versucht einen Wechsel zu einem neuen Spiel herbeizuführen.

*(k) Keine Joint Attention*

Mutter und Kind weisen keinen gemeinsamen Aufmerksamkeitsfokus auf (Beispiel: die Mutter ist noch bei der Bastelaufgabe und malt das Haus an, während das Kind schon die Spielkiste exploriert).

*(u) Unkodierbar – Joint Attention*

Die Situation kann nicht bewertet werden, da Mutter und Kind nicht sichtbar sind, die Mutter den Raum verlässt oder die Interaktion durch eine dritte Person gestört wird. Auch wenn eine oder beide Beteiligten direkt in die Kamera sehen, ist dies als unkodierbar zu werten.

## **9 Untersuchungsdesign und Durchführung**

### **9.1 Design der Untersuchung**

Bei der Untersuchung handelt es sich um ein Quasiexperimentelles Design, da keine Randomisierung der Probanden vorgenommen werden konnte. Die Probanden, sprich die Mutter-Kind-Dyaden (es wurden nur Mütter herangezogen, keine Väter), wurden im Zuge ihres stationären Aufenthalts im Wilhelminenspital mehrmals in ihrem Interaktionsverhalten auf Video aufgenommen. Somit existieren von jeder Mutter-Kind-Dyade mehrere Videos zu unterschiedlichen Zeitpunkten. Es konnten insgesamt drei Untersuchungszeitpunkte der Interaktionsvideos der Station ausfindig gemacht werden: Diagnostikvideos, Interventionsvideos und Abschlussvideos. Unter den Diagnostikvideos sind Videos zu verstehen, die zu Beginn des stationären Aufenthalts gedreht wurden, um den Status Quo von Mutter und Kind dokumentieren bzw. diagnostizieren zu können. Interventionsvideos wurden von der Psychologin und der Ärztin zur Verlaufskontrolle während des Therapieprozesses aufgenommen, da sie als Teil der Intervention fungieren. Die Abschlussvideos dienten kurz vor der Entlassung von Elternteil und Kind als abschließender Beleg für den Fortschritt, den Elternteil und Kind durch den stationären Aufenthalt verzeichnen konnten. Aufgrund einer anonymisierten Zuteilung der Videos, wussten die beiden BeobachterInnen bei der Beurteilung der Videos nicht, ob sie das Interaktionsverhalten vor oder nach der Intervention bewerten. So konnte gewährleistet werden, dass BeobachterInneneffekte (siehe Kapitel 5.2.1) zu Lasten der Objektivität ausgeschaltet wurden.

### **9.2 Durchführung der Datenerhebung**

#### **9.2.1 *Einschulung der BeobachterInnen***

Die beiden BeobachterInnen erhielten eine Einschulung zu INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) und auch zum Umgang mit der Auswertungssoftware ‚INTERACT 9‘ von Mangold durch Frau Dr. Hirschmann. Die Einschulung bezog sich auf Mutter-Kind-Interaktionen in Spiel-Situationen. Die Kinder, die in den Einschulungsvideos

zu sehen waren, befanden sich in einem Alter von 3 bis 6 Jahren. Abschließend musste ein Probevideo bewertet werden, mit dem Ziel eine BeobachterInnenübereinstimmung von mindestens  $\kappa = 0.8$  mit Frau Dr. Hirschmanns Beurteilungen zu erlangen.

Des Weiteren sei erwähnt, dass Frau Dr. Hirschmann während des gesamten Auswertungsprozesses für Rückfragen und eventuell auftretende Unsicherheiten zur Verfügung stand.

### 9.2.2 *Sichtung der Videos*

Die Sichtung, Auswahl und anschließende Protokollierung der Interaktionsvideos erfolgte aufgrund spezieller Kriterien. Die Ausschlusskriterien können wie folgt zusammengefasst werden:

- Es wurden nur Videos mit Mutter-Kind-Interaktionen herangezogen, keine Vater-Kind-Interaktionsvideos.
- In den Videos durften keine Geschwister involviert sein.
- Weiters gab es des Öfteren Familienessen mit anderen Familienmitgliedern, auch diese wurden ausgeschlossen.
- Das Alter des Kindes musste zwischen 1;1 und 5;11 liegen.
- Außerdem konnten keine fremdsprachigen Videos beurteilt werden, bei denen die sprachliche Interaktion nicht verstanden wurde (allerdings wurde eine englischsprachige und eine pakistanische Mutter-Kind-Dyade in die Stichprobe aufgenommen, da diese von mindestens einem Rater verstanden wurde).
- Herangezogen wurden nur die Diagnosevideos zu Beginn eines stationären Aufenthalts (Präsituation) und die Abschlussvideos kurz vor der Entlassung von Mutter und Kind (Postsituation). Alle Videos zur therapeutischen Verlaufskontrolle (sogenannte Interventionsvideos), die zwischen diesen beiden Zeitpunkten aufgenommen wurden, wurden nicht herangezogen.
- Sonderbedingung: Manchmal kam es vor, dass es für eine Mutter-Kind-Dyade kein Abschlussvideo gab, dann wurde das vom Datum späteste Interventionsvideo ausgewählt und als Post-Zeitpunkt herangezogen. Es wurde darauf geachtet, dass zwischen dem Zeitpunkt des Diagnostikvideos und des spätesten

Interventionsvideos mindestens 2 Wochen Abstand lagen. Genau dieselbe Handhabung fand bei einem fehlenden Diagnostikvideo Anwendung. In diesem Fall wurde das vom Zeitpunkt früheste Interventionsvideo herangezogen.

### **9.2.3 Aufteilung der Videos**

Nach Erhalt aller potentiellen Videos wurden diese, für die Auswertung unter den beiden BeobachterInnen aufgeteilt. Bevor jedoch mit der Auswertung der Interaktionsvideos angefangen werden konnte, mussten die ursprünglichen Namen der Videos verschleiert werden, da sie Informationen zu Prä- oder Post-Zeitpunkt gaben und daraus BeobachterInnenfehler hätten resultieren können (vgl. Kapitel 5.2.1.1). Bei der Anonymisierung wurde so vorgegangen, dass der Beobachter oder die Beobachterin die Videos der Kollegin oder des Kollegen anonymisierte und zwar mit den Initialen dieser Kollegin oder des Kollegen und einer zusätzlichen fortlaufenden Nummerierung (z.B. AG09).

Des Weiteren erhielten die beiden KodiererInnen die Gegenkodierungsstichprobe, welche die Grundlage für die hier vorliegende Arbeit darstellt, indem 20% aller Videos gegenkodiert wurden. Jede BeurteilerIn kodierte somit zusätzlich zu den ihr zugewiesenen Aufnahmen 20% der Videos der KollegIn. Hierfür wurde beachtet, dass die Verteilung der verschiedenen Situationen (Essen, Pflege, Spiel) der Gegenkodierungsstichprobe repräsentativ für die Gesamtstichprobe ist (vgl. Kapitel 10.4).

### **9.2.4 Auswertungssoftware ,INTERACT 9‘ von Mangold**

„INTERACT 9“ von Mangold ist ein Auswertungsprogramm mit dessen Hilfe Daten der Verhaltensbeobachtungen quantifiziert und ausgewertet werden können. Ereignisse werden mittels vorher festgelegter Codes durch Tastendruck in Sequenzen mit einer Anfangs- und einer Endzeit im Format Stunden:Minuten:Sekunden:Bilder (HH:MM:SS:FF) abgespeichert. Jedes Video musste vor der Auswertung mit „INTERACT 9“ von Mangold in ein geeignetes Format (.WMV) konvertiert werden, da die Videos der Station ein für das

Auswertungsprogramm nicht kompatibles Format aufwiesen. Hierfür wurde das Programm ‚Any Video Converter‘ verwendet.

### **9.2.5 Vorbereitung der Daten für die statistische Analyse**

Nach der Beurteilung der Videos durch ‚INTERACT 9‘ von Mangold wurden die Datensets der Kategorien *Rückmeldung* und *Joint Attention* mittels der Software ‚ActSds‘ in das Format .SDIS konvertiert. Im Anschluss daran wurden die Daten mit der Software ‚GSEQ 5.1‘ in ein .MDS-Format konvertiert und auch die Berechnungen zu diesen beiden Skalen konnten in ‚GSEQ 5.1‘ angestellt werden. Mangolds ‚INTERACT 9‘ ist ausschließlich in der Lage ein event-basiertes Kappa zu berechnen, wobei hingegen ‚GSEQ 5.1‘ auch ein time-unit basiertes Kappa berechnen kann.

Die Beobachtungen der Skala *Feinfühligkeit* wurden von ‚INTERACT 9‘ von Mangold händisch in IBM SPSS Statistics übertragen. Da es sich um eine Intervallskala handelt und die Ausprägung (8) *Unkodierbar – Feinfühligkeit* nicht berücksichtigt werden konnte, wurden diese Sequenzen von der Analyse ausgeschlossen. Alle anfallenden Berechnungen zur Skala *Feinfühligkeit* wurden in IBM SPSS Statistics getätigt.

## **10 Beschreibung der Stichprobe**

### **10.1 Inhalte der Interaktionsvideos**

Alle Mutter-Kind-Interaktionen wurden von den Psychologinnen, Schwestern oder Ärztinnen der Säuglingspsychosomatik des Wilhelminenspitals in den PatientInnenzimmern aufgenommen. Die Videos zeigten entweder ein gemeinsames Essen, ein Pflege-Szenario oder eine Spiel-Situation. Während der Essens-Situationen wurden Mutter und Kind bei der Nahrungsaufnahme aufgenommen. Die Bandbreite reicht von Müttern, die ihre Kinder mit Löffel oder Flasche füttern, bis hin zu Müttern mit Kindern, die schon selbstständig essen können. Ein Kind wurde über eine Nasensonde gefüttert. Unter Pflege-Szenarios sind jene Situationen zusammengefasst, in denen Mütter ihre Kinder wickeln, baden oder ankleiden. Die Spiel-Situationen waren freie Spielsituationen, in denen Mütter und Kinder frei wählen konnten, mit welchen Spielmaterialien sie sich beschäftigen. Die Videos endeten, sobald die Mutter beschloss die Krankenschwester zum Abschalten des Videos zu holen, oder wenn das Band der Videokamera aufgebraucht war.

Die Aufnahmen fanden außerdem zu zwei verschiedenen Zeitpunkten statt, nämlich zu Beginn des stationären Aufenthalts, vor der gesetzten Interventionsmaßnahme, und am Ende des Aufenthalts, also kurz vor der Entlassung der Mutter-Kind-Dyade, nachdem die Interventionsmaßnahmen abgeschlossen waren.

### **10.2 Beschreibung der Gesamtstichprobe**

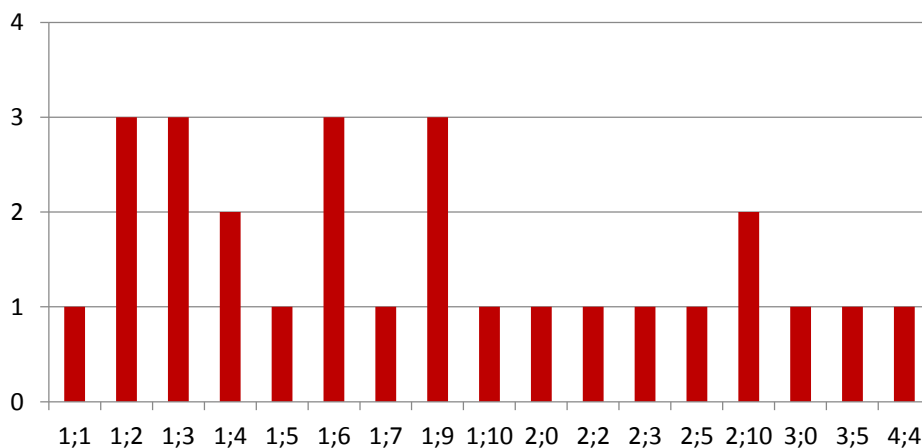
Nach Sichtung des Videomaterials auf der Säuglingspsychosomatik des Wilhelminenspitals konnten aufgrund der Ausschlusskriterien (siehe Kapitel 9.2.2) 27 Mutter-Kind-Dyaden ausfindig gemacht werden, von denen es insgesamt 129 Videos in einem Umfang von 34:14:31 (h:min:sec) gibt. Die durchschnittliche Videodauer beträgt 00:16:09 ( $SD = 00:09:58$ ). Die 129 Interaktionsvideos unterteilen sich in 56 Essens-, 39 Spiel- und 34 Pflege-Situationen sowie 82 Prä- und 47 Post-Situationen (siehe Tabelle 1). Im Anhang (Tabelle A-4) befindet sich eine detaillierte Auflistung aller Videos.

**Tabelle 1:** Kontingenztafel zur Verteilung der einzelnen Untersuchungssituationen und Erhebungszeitpunkte in der Gesamtstichprobe, Angaben in absoluten Häufigkeiten ( $N = 129$  Videos)

|          | Zeitpunkt |      | $\Sigma$ |
|----------|-----------|------|----------|
|          | Prä       | Post |          |
| Essen    | 34        | 22   | 56       |
| Pflege   | 23        | 11   | 34       |
| Spiel    | 25        | 14   | 39       |
| $\Sigma$ | 82        | 47   | 129      |

Der Chi-Quadrat-Unabhängigkeitstest für die beiden Variablen ‚Situation‘ und ‚Zeitpunkt‘ ergab mit einem  $\chi^2 = 0.446 < \chi_{\text{krit}}^2 = 5.991$  ( $df = 2, \alpha = .05$ ), dass die Nullhypothese auf einem Signifikanzniveau von 5% nicht verworfen werden kann. Das bedeutet, dass die beiden Variablen ‚Situation‘ und ‚Untersuchungszeitpunkt‘ aufgrund ihrer Verteilung in der zugrunde liegenden Population als voneinander unabhängig betrachtet werden können (siehe Tabelle 1).

Das jüngste Kind dieser Population war in einem Alter von 1;1 Jahren, das älteste Kind war 4;4 Jahre alt. Der Altersdurchschnitt beträgt 22.6 Monate ( $SD = 9.6$ ). Die Altersverteilung ist in Abbildung 2 graphisch dargestellt. Soziodemographische Daten zu den Müttern liegen nicht vor.



**Abbildung 2:** Darstellung der absoluten Häufigkeiten des Alters in Jahr(en);Monat(en) der Gesamtstichprobe

### 10.3 Beschreibung der Gegenkodierungsstichprobe

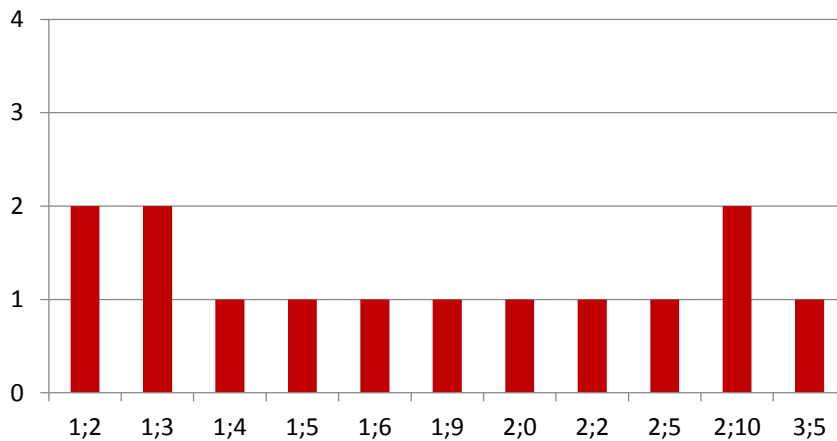
Die Gegenkodierungsstichprobe, 20% der Videos der Gesamtstichprobe wurden von beiden BeobachterInnen beurteilt, umfasst 14 Mutter-Kind-Dyaden und 28 Videos mit einer Gesamtdauer von 06:32:20 (h:min:sec). Der Mittelwert der Videodauer beträgt in der Gegenkodierungsstichprobe 00:14:01 (h:min:sec),  $SD = 00:08:59$ . Die 28 Videos setzen sich aus 13 Essens-, 8 Pflege- und 7 Spiel-Situationen zusammen sowie aus 19 Prä- und 9 Post-Situationen (siehe Tabelle 2).

**Tabelle 2:** Kontingenztafel zur Verteilung der einzelnen Untersuchungssituationen und Erhebungszeitpunkte in der Gegenkodierungsstichprobe, Angaben in absoluten Häufigkeiten ( $N = 28$  Videos)

|           |          | Zeitpunkt |      | $\Sigma$  |
|-----------|----------|-----------|------|-----------|
|           |          | Prä       | Post |           |
| Situation | Essen    | 11        | 2    | 13        |
|           | Pflege   | 4         | 4    | 8         |
|           | Spiel    | 4         | 3    | 7         |
|           | $\Sigma$ | 19        | 9    | <b>28</b> |

Der Chi-Quadrat-Unabhängigkeitstest für die beiden Variablen ‚Situation‘ und ‚Zeitpunkt‘ ergab mit einem  $\chi^2 = 3.212 < \chi_{\text{krit}}^2 = 5.991$  ( $df = 2$ ,  $\alpha = .05$ ), dass die Nullhypothese auf einem Signifikanzniveau von 5% nicht verworfen werden kann. Das bedeutet, dass die beiden Variablen ‚Situation‘ und ‚Untersuchungszeitpunkt‘ aufgrund ihrer Verteilung auch in der Gegenkodierungsstichprobe voneinander unabhängig sind (siehe Tabelle 2).

Das jüngste Kind der Gegenkodierungsstichprobe ist zum Zeitpunkt der Videoaufnahme 14 Monate alt, das älteste 41 Monate. Der Mittelwert beträgt 22.7 Monate ( $SD = 8.8$ ). In Abbildung 3 ist die Altersverteilung graphisch dargestellt.



**Abbildung 3:** Darstellung der absoluten Häufigkeiten des Alters in Jahr(en);Monat(en) der Gegenkodierungsstichprobe

## 10.4 Vergleich der Stichproben

In der Forschungsarbeit des Kollegen Valai (in Vorbereitung) werden die beiden Erhebungszeitpunkte ‚Prä‘ und ‚Post‘ für die Gesamtstichprobe ( $N = 27$ ) analysiert. Die hier vorliegende Arbeit beschäftigt sich mit der Teilstichprobe ( $N = 14$ ), auf deren Grundlage anhand der BeobachterInnenübereinstimmung erörtert wird, ob sich das Instrument auf das klinische Setting und einen jüngeren Altersbereich anwenden lässt.

Tabelle 3 stellt die Kennwerte des Alters der Kinder in den verschiedenen Stichproben dar. Das durchschnittliche Alter der Kinder lag sowohl für die Gesamtstichprobe, die Gegenkodierungsstichprobe, wie auch für die Probanden, die nicht zur Analyse der BeobachterInnenübereinstimmung herangezogen wurden, bei 22 Monaten (1;10).

**Tabelle 3:** Vergleich des Alters (in Monaten) zwischen der Gegenkodierungsstichprobe, jenen Probanden, die nicht zur Reliabilitätsbestimmung herangezogen wurden und der Gesamtstichprobe

|                           | <i>N</i> | <i>M</i> | <i>SD</i> | <i>Min</i> | <i>Max</i> |
|---------------------------|----------|----------|-----------|------------|------------|
| Gegenkodierungsstichprobe | 14       | 22.71    | 8.84      | 14         | 41         |
| Nicht gewählte Probanden  | 13       | 22.46    | 10.77     | 13         | 52         |
| Gesamtstichprobe          | 27       | 22.59    | 9.62      | 13         | 52         |

Es wurde zur Überprüfung, ob die Variablen der Gegenkodierungsstichprobe dieselben Verteilungen aufweisen, wie die nicht gewählten Probanden, der Chi-Quadrat-Homogenitätstest angewandt.

Für die voneinander unabhängigen Variablen ‚Situation‘ (Essen, Spiel, Pflege) und ‚Zeitpunkt‘ (Prä, Post) konnte mit  $\chi^2 = 6.450 < \chi_{\text{krit}}^2 = 11.070$  (df = 5,  $\alpha = .05$ ) ermittelt werden, dass die Nullhypothese beibehalten werden kann. Somit ist anzunehmen, dass für die Untersuchungssituationen und Erhebungszeitpunkte ähnliche Verteilungen in der Gegenkodierungsstichprobe, wie auch für die nicht gewählten Probanden der Gesamtstichprobe, vorliegen (siehe Anhang Tabelle A-5).

Auch das ‚Kindesalter‘ ist sowohl für die Probanden der Gegenkodierungsstichprobe, als auch für die übrigen Probanden der Gesamtstichprobe, die nicht zur Analyse herangezogen wurden, auf die gleiche Weise verteilt. Dies belegte der Chi-Quadrat-Test mit  $\chi^2 = 5.309 < \chi_{\text{krit}}^2 = 12.592$  (df = 6,  $\alpha = .05$ , siehe Tabelle 4).

**Tabelle 4:** Verteilung des ‚Alters‘ in Halbjahresschritten für die Gegenkodierungsstichprobe, für jene Probanden, die nicht für die Reliabilitätsanalyse herangezogen wurden und die Gesamtstichprobe, Angaben in absoluten und relativen Häufigkeiten

|          | Gegenkodierungsstichprobe | Nicht gewählte Probanden | Gesamtstichprobe |
|----------|---------------------------|--------------------------|------------------|
| 1;1-1;6  | 14 (50%)                  | 50 (50%)                 | 64 (50%)         |
| 1;7-2;0  | 4 (14%)                   | 26 (26%)                 | 30 (23%)         |
| 2;1-2;6  | 6 (21%)                   | 8 (8%)                   | 14 (11%)         |
| 2;7-3;0  | 3 (11%)                   | 13 (13%)                 | 16 (12%)         |
| 3;1-3;6  | 1 (4%)                    | 3 (3%)                   | 4 (3%)           |
| 3;7-4;0  | 0 (0%)                    | 0 (0%)                   | 0 (0%)           |
| 4;1-4;6  | 0 (0%)                    | 1 (1%)                   | 1 (1%)           |
| $\Sigma$ | 28 (100%)                 | 101 (100%)               | 129 (100%)       |

Für die Verteilung des Kindesalters auf die verschiedenen Untersuchungssituationen (Essen, Pflege, Spiel) und Erhebungszeitpunkte (Prä, Post) konnte mit  $\chi^2 = 37.659 < \chi_{\text{krit}}^2 = 55.76$  (df = 41,  $\alpha = .05$ ) ebenfalls festgestellt werden, dass für die Gegenkodierungsstichprobe und jene Probanden der Gesamtstichprobe, die nicht zur Analyse der BeobachterInnenübereinstimmung herangezogen wurde, eine ähnliche Verteilung vorliegt (siehe Anhang Tabellen A-6a und A-6b).

## 11 Ergebnisse der Untersuchung

### 11.1 Skala Feinfühligkeit

#### 11.1.1 Betrachtung über alle Situationen

##### 11.1.1.1 Darstellung der Häufigkeiten

Die Gegenkodierungsstichprobe setzt sich aus 28 Videos von insgesamt 14 Mutter-Kind-Dyaden zusammen. Es wurden 169 Intervalle zur *Feinfühligkeit* bewertet. In Tabelle 5 sind die absoluten und relativen Häufigkeiten der Bewertungen zur *Feinfühligkeit* beider BeobachterInnen über alle Situationen und Videos abgebildet. BeobachterIn B hat am Häufigsten (in 48,5% der Fälle) das Rating (4) *mittlere Feinfühligkeit* abgegeben. BeobachterIn A hingegen hat in 33,1% der Fälle eine (5) *eher hohe Feinfühligkeit* angegeben und in 36,7% der Fälle sogar eine (6) *hohe Feinfühligkeit*. Es ist deutlich zu erkennen, dass BeobachterIn B die Probanden generell mit niedrigeren Scores bewertet hat als BeurteilerIn A. 71% der Bewertungen des Raters B entfallen entweder auf die Ausprägung (3) *eher geringe Feinfühligkeit* oder (4) *mittlere Feinfühligkeit*, im Gegensatz zu 19% bei Rater A. In 69,8% der Fälle entschied sich BeurteilerIn A entweder für eine (5) *eher hohe Feinfühligkeit* oder (6) *hohe Feinfühligkeit*, BeobachterIn B hingegen nur in 14,2%. Dies unterstreicht die tendenziell höher angesiedelten Bewertungen von BeobachterIn A.

**Tabelle 5:** absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Situationen ( $N = 169$ )

|                                 | absolute Häufigkeit |         | relative Häufigkeit |         |
|---------------------------------|---------------------|---------|---------------------|---------|
|                                 | Rater A             | Rater B | Rater A             | Rater B |
| (1) sehr geringe Feinfühligkeit | 2                   | 0       | 1,2                 | 0       |
| (2) geringe Feinfühligkeit      | 4                   | 4       | 2,4                 | 2,4     |
| (3) eher geringe Feinfühligkeit | 17                  | 38      | 10,1                | 22,5    |
| (4) mittlere Feinfühligkeit     | 15                  | 82      | 8,9                 | 48,5    |
| (5) eher hohe Feinfühligkeit    | 56                  | 17      | 33,1                | 10,1    |
| (6) hohe Feinfühligkeit         | 62                  | 7       | 36,7                | 4,1     |
| (7) sehr hohe Feinfühligkeit    | 13                  | 21      | 7,7                 | 12,4    |
| Gesamt                          | 169                 | 169     | 100                 | 100     |

Auch Tabelle 6 reflektiert die höher angesiedelten Bewertungen von BeobachterIn A ( $M = 5.11$ ,  $S^2 = 1.57$ ), im Gegensatz zu denen von BeobachterIn B ( $M = 4.28$ ,  $S^2 = 1.63$ ).

**Tabelle 6:** Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Situationen ( $N = 169$ )

|                | <i>N</i> | <i>M</i> | $S^2$ | <i>Min</i> | <i>Max</i> |
|----------------|----------|----------|-------|------------|------------|
| BeobachterIn A | 169      | 5.11     | 1.57  | 1          | 7          |
| BeobachterIn B | 169      | 4.28     | 1.63  | 2          | 7          |

Bei der näheren Betrachtung der Unterschiede in der Bewertung einzelner Feinfühligkeits-Sequenzen zwischen den beiden Ratern (siehe Tabelle 7), ist deutlich zu erkennen, dass in 33 der 169 Ratings beide Rater zu exakt derselben Beurteilung gekommen sind und in 61 Bewertungen liegt eine Differenz von einem Skalenpunkt vor. Nach Wirtz und Caspar (2002) ist die Forderung nach exakter Übereinstimmung bei intervallskalierten Ratingskalen nicht notwendig und auch meist nicht sinnvoll. „Bei Urteilen auf kontinuierlichen Skalen liegt es in der Natur der Sache, dass exakte Übereinstimmung einen Ausnahmefall darstellt“ (Wirtz & Caspar, 2002, S. 34). Sie postulieren weiters, dass Unterschiede zwischen den Urteilen für ein- und dasselbe Objekt zugelassen werden können, wenn diese Unterschiede verhältnismäßig gering sind. Dann kann trotzdem von einer hohen Übereinstimmung gesprochen werden und von einer hohen Zuverlässigkeit ausgegangen werden. Somit entfallen 94 von 169 Bewertungen (55,6%) auf eine perfekte Übereinstimmung oder auf eine annähernde Übereinstimmung mit einer Differenz von einem Skalenpunkt. In 65 von 169 Fällen, unterscheiden sich die Entscheidungen der Rater um zwei Wertpunkte. Neun Intervalle weisen eine Differenz von drei Ausprägungsstufen auf und nur einmal gab es einen Unterschied von vier Stufen (siehe Anhang, Abbildung A-1a).

**Tabelle 7:** Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Situationen ( $N = 169$ )

|                      |   | <i>absolute Häufigkeit</i> | <i>relative Häufigkeit</i> |
|----------------------|---|----------------------------|----------------------------|
| Betrag der Differenz | 0 | 33                         | 19,5                       |
|                      | 1 | 61                         | 36,1                       |
|                      | 2 | 65                         | 38,5                       |
|                      | 3 | 9                          | 5,3                        |
|                      | 4 | 1                          | 0,6                        |
|                      | 5 | 0                          | 0                          |
|                      | 6 | 0                          | 0                          |
| Gesamt               |   | 169                        | 100                        |

#### 11.1.1.2 Darstellung der Intraklassenkorrelation

Es wurde das zweifaktorielle Modell der Intraklassenkorrelation (ICC) gewählt, weil alle Videos der Gegenkodierungs-Stichprobe von beiden BeobachterInnen beurteilt wurden. Da es der Forschungsfrage nicht dienlich ist, die Mittelwertsunterschiede bereinigt zu wissen, denn genau diese sind für die Fragestellung der Arbeit von großem Interesse, wurde die unjustierte Herangehensweise (absolute Übereinstimmung) gewählt. Der Faktor *Rater* wurde als ‚random‘ angesehen, da alle INTAKT-Rater die gleiche Schulung erhalten haben, somit von demselben Wissensstand aller Rater auszugehen ist, und die Zuteilung der Forschungsfragen unter den DiplomandInnen zufällig erfolgt ist (vgl. Kapitel 9.2.3).

Bei der Überprüfung der Voraussetzungen für die ICC über alle Situationen konnte durch Tukeys Additivitätstest ( $F_{nonadd} = .095$ ;  $p = .759 > \alpha = .25$ , siehe Wirtz & Caspar, 2002, S. 210) bekräftigt werden, dass keine signifikanten Interaktionen zwischen den BeobachterInnen und den beurteilten Personen stattgefunden haben. Zusätzlich dazu müssen die Voraussetzungen für varianzanalytische Modelle erfüllt sein, sprich Homogenität der Varianzen, sowie eine angenäherte Normalverteilung der Daten (Wirtz & Caspar, 2002). Der Levene-Test ergab mit einem nicht signifikanten Ergebnis von  $p = .829$  ( $\alpha = .25$ , auf Empfehlung von Wirtz & Caspar, 2002, S. 180), dass die Voraussetzung der Varianzhomogenität als gegeben angesehen werden kann. Die Normalverteilungs-Prüfung mittels Kolmogorov-Smirnov-Test ergab allerdings, dass sich die Verteilung signifikant von einer Normalverteilung unterscheidet ( $p = .000$  für beide Rater,  $\alpha = .05$ , siehe Anhang,

Abbildungen A-2a und A-2b). Nach Wirtz und Caspar (2002) kann eine isolierte Verletzung der Normalverteilungsannahme vernachlässigt werden.

Für die Betrachtung von Intraklassenkorrelation ist prinzipiell festzulegen, dass die ICC ein Maß zur Varianzaufklärung ist, welches Werte zwischen 0 und 1 annehmen kann. Die Ausprägung des Koeffizienten sollte immer in Abhängigkeit des zu messenden Merkmals und der zu Grunde liegenden Stichprobe interpretiert werden (Wirtz & Caspar, 2002, vgl. Kapitel 6.2.1).

Die BeobachterInnenübereinstimmung zwischen den beiden Ratern B und A (siehe Tabelle 8) über alle drei Situationen und Videos (169 Feinfühligkeitssequenzen von insgesamt 28 Videos und 14 Mutter-Kind-Dyaden), ist mit einem signifikanten Reliabilitätskoeffizienten von  $r_{icc} = .361$  ( $F = 2.558$ ;  $p = .000 < \alpha = .05$ ) als eher gering einzustufen. Mit einer Sicherheit von 95% befindet sich der Koeffizient der Intraklassenkorrelation zwischen .126 und .539. Es sei hier außerdem erwähnt, dass aufgrund der Signifikanz des Raterfaktors ( $F = 64.512$ ;  $p = .000 < \alpha = .25$ , siehe Wirtz & Caspar, 2002, S. 188) in der zweifaktoriellen Varianzanalyse die Homogenität der Mittelwerte nicht bestätigt werden konnte. Die Mittelwerte der beiden Rater unterscheiden sich signifikant voneinander, worauf der eher niedrige ICC-Koeffizient unter anderem zurückzuführen ist. Nach Wirtz und Caspar (2002) indiziert „... eine Signifikanz des Faktors ‚Rater‘, dass durch eine Angleichung der Ratermittelwerte eine deutliche Verbesserung der Reliabilität erreicht werden kann“ (S. 187).

**Tabelle 8:** Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Situationen und Videos ( $N = 169$ )

|               | Korrelation<br>innerhalb der<br>Klasse | 95% Konfidenzintervall |            | F-Test mit wahrem Wert 0 |     |     |      |
|---------------|----------------------------------------|------------------------|------------|--------------------------|-----|-----|------|
|               |                                        | Untergrenze            | Obergrenze | Wert                     | df1 | df2 | Sig  |
| Einzelne Maße | .361                                   | .126                   | .539       | 2.558                    | 168 | 168 | .000 |

### 11.1.2 Betrachtung der Essens-Situationen

#### 11.1.2.1 Ergebnisse zum Essens-Kontext

Auch für die Essens-Situationen gilt, dass BeobachterIn A die *Feinfühligkeit* tendenziell höher bewertet hat, als Rater B. Wie aus Tabelle 9 ersichtlich ist, entschied sich BeurteilerIn A in 73,2% der Fälle für eine (5) *eher hohe Feinfühligkeit* oder (6) *hohe Feinfühligkeit*, BeobachterIn B hingegen nur in 11,9% der Zeit. Am Häufigsten, nämlich zu 68,9%, entschied sich B für eine (3) *eher geringe Feinfühligkeit* oder (4) *mittlere Feinfühligkeit*, demgegenüber stehen 16,2% bei BeobachterIn A.

**Tabelle 9:** absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Essens-Situationen ( $N = 93$ )

|                                 | absolute Häufigkeit |         | relative Häufigkeit |         |
|---------------------------------|---------------------|---------|---------------------|---------|
|                                 | Rater A             | Rater B | Rater A             | Rater B |
| (1) sehr geringe Feinfühligkeit | 1                   | 0       | 1,1                 | 0       |
| (2) geringe Feinfühligkeit      | 2                   | 3       | 2,2                 | 3,2     |
| (3) eher geringe Feinfühligkeit | 9                   | 18      | 9,7                 | 19,4    |
| (4) mittlere Feinfühligkeit     | 6                   | 46      | 6,5                 | 49,5    |
| (5) eher hohe Feinfühligkeit    | 34                  | 5       | 36,6                | 5,4     |
| (6) hohe Feinfühligkeit         | 34                  | 6       | 36,6                | 6,5     |
| (7) sehr hohe Feinfühligkeit    | 7                   | 15      | 7,5                 | 16,1    |
| Gesamt                          | 93                  | 93      | 100                 | 100     |

Tabelle 10 reflektiert die durchschnittliche Einschätzung der Feinfühligkeit über alle Sequenzen mit Essens-Kontext. BeobachterIn A vergab eine durchschnittlichen Feinfühligkeitswert von  $M = 5.15$  ( $S^2 = 1.48$ ), BeobachterIn B einen Mittelwert von  $M = 4.41$  ( $S^2 = 1.94$ ).

**Tabelle 10:** Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Essens-Situationen ( $N = 93$ )

|                | $N$ | $M$  | $S^2$ | $Min$ | $Max$ |
|----------------|-----|------|-------|-------|-------|
| BeobachterIn A | 93  | 5.15 | 1.48  | 1     | 7     |
| BeobachterIn B | 93  | 4.41 | 1.94  | 2     | 7     |

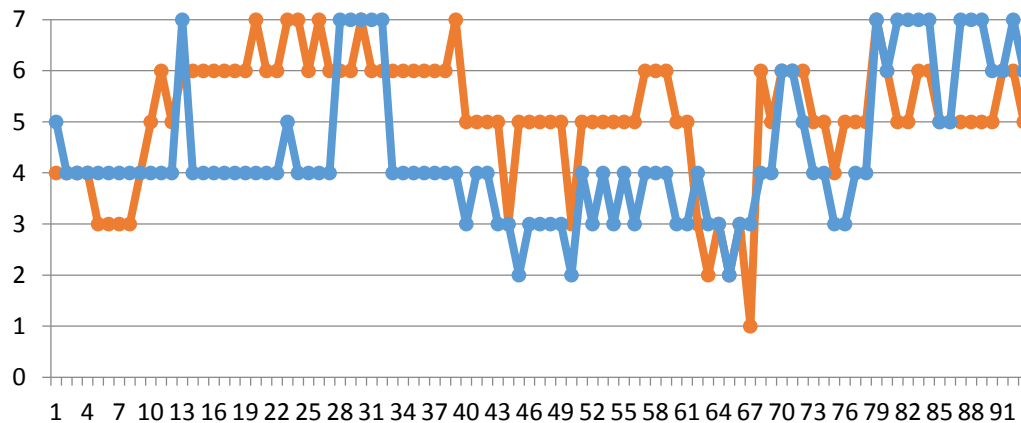
In 16 von 93 Sequenzen aller Videos mit einer Essens-Thematik stimmen die beiden Rater perfekt überein und gaben somit exakt das gleiche Urteil ab (siehe Tabelle 11). Einen Unterschied von einem Skalenpunkt gibt es in 32 Urteilen. Somit kamen die Rater in 51,6% der Fälle zu demselben Schluss oder einem ähnlichen Ergebnis mit einem Wertpunkt Unterschied. Eine Differenz von zwei Abstufungen erfolgte in 40 Bewertungen. Insgesamt fünf Intervalle wurden mit einem Unterschied von drei Ausprägungsstufen bewertet (siehe Anhang, Abbildung A-1b).

**Tabelle 11:** Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Essens-Situationen ( $N = 93$ )

|                      |   | <i>absolute Häufigkeit</i> | <i>relative Häufigkeit</i> |
|----------------------|---|----------------------------|----------------------------|
| Betrag der Differenz | 0 | 16                         | 17,2                       |
|                      | 1 | 32                         | 34,4                       |
|                      | 2 | 40                         | 43,0                       |
|                      | 3 | 5                          | 5,4                        |
|                      | 4 | 0                          | 0                          |
|                      | 5 | 0                          | 0                          |
|                      | 6 | 0                          | 0                          |
| Gesamt               |   | 93                         | 100                        |

In Abbildung 4 ist optisch gut zu erkennen, dass die Beurteilungen von Rater B tendenziell öfter unter denen von Rater A liegen. Von den 93 Essens-Sequenzen gab es 16 exakte Übereinstimmungen (17,2%). In 32 Sequenzen (34,4%) gab es einen Unterschied von einem Wertpunkt und in 40 Sequenzen (43%), gab es eine Differenz von zwei Ausprägungsstufen.

## Ergebnisse der Untersuchung



**Abbildung 4:** Kodierungsverlauf der beiden Rater über alle FFK-Sequenzen der Essens-Situationen ( $N = 93$ )

Bevor der ICC-Koeffizient herangezogen werden konnte, mussten zuerst die Voraussetzungen überprüft werden. Durch Tukeys Additivitätstest ( $F_{nonadd} = 2.050$ ;  $p = .156 < \alpha = .25$ ) konnte nicht bestätigt werden, dass keine signifikanten Interaktionen zwischen Probanden und Ratern stattgefunden haben. Des Weiteren konnte die Voraussetzung der Varianzhomogenität durch den Levene-Test nicht bestätigt werden ( $p = .092 < \alpha = .25$ ). Auch die Berechnung des Kolmogorov-Smirnov-Test zur Überprüfung der Voraussetzung der Normalverteilung ist für die Essens-Videos nicht erfüllt ( $p = .000$  für beide Rater,  $\alpha = .05$ , siehe Anhang Abbildungen A-3a und A-3b).

Aufgrund der schwerwiegenden Verletzungen aller Voraussetzungen ist das Heranziehen der ICC ( $r_{icc} = .354$ , Tabelle 12) der 12 Essens-Videos mit insgesamt 93 FFK-Sequenzen von 10 Mutter-Kind-Dyaden als problematisch anzusehen. Nach Wirtz und Caspar (2002) ist unter diesen Umständen keine sinnvolle Interpretation der Intraklassenkorrelation möglich.

Die zweifaktorielle Varianzanalyse gab einen signifikanten Haupteffekt des Raterfaktors ( $F = 25.348$ ;  $p = .000 < \alpha = .25$ ) aus, was für die Heterogenität der Ratermittelwerte spricht.

**Tabelle 12:** Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Essens-Situationen (N = 93)

|               | Korrelation<br>innerhalb der<br>Klasse | 95% Konfidenzintervall |            | F-Test mit wahrem Wert 0 |     |     |      |
|---------------|----------------------------------------|------------------------|------------|--------------------------|-----|-----|------|
|               |                                        | Untergrenze            | Obergrenze | Wert                     | df1 | df2 | Sig  |
| Einzelne Maße | .354                                   | .130                   | .536       | 2.384                    | 92  | 92  | .000 |

#### 11.1.2.2 Kodierprobleme im Essens-Kontext

Bei der Bewertung der Feinfühligkeit im Essenskontext ist anzumerken, dass Mütter mit sehr jungen Kindern (1;1 - 1;6) oft lenkend vorgehen, weil die Kinder ansonsten die Schüssel mit der heißen Suppe auf den Boden werfen würden oder zu heiße Speisen essen würden und sich somit verbrennen könnten. Dadurch erhalten sie auf der Skala der Feinfühligkeit automatisch schlechtere Werte, als Mütter von größeren Kindern (2;6 - 3;0), bei denen nicht mehr so stark lenkend vorgegangen werden muss. Es ist zu überlegen, den Fokus bei jüngeren Kindern auf andere Verhaltensweisen zu legen. Denn eine Mutter, die ihr Kind vor einer Gefahr schützt, die das Kind noch nicht einschätzen kann, und somit eingreift und ihre Tätigkeit in diesem Moment nicht zurückstellt, ist trotzdem eine feinfühlige Mutter und sollte nicht aufgrund ihres notwendigen Eingreifens schlechtere Bewertungen erhalten. Hier könnten Kriterien wie der Sprachstil, das gemeinsame Arbeiten an einer Sache, das Signale Erkennen und darauf prompt und angemessen zu reagieren oder das gelungene Einnehmen des kindlichen Blickwinkels, in den Vordergrund rücken.

#### 11.1.3 Betrachtung der Pflegesituationen

##### 11.1.3.1 Ergebnisse zum Pflege-Kontext

Auch für die 36 Sequenzen der Pflege-Situationen ergibt sich das bisherige Bild, dass BeobachterIn A etwas höhere Bewertungen abgegeben hat als Rater B. 58,4% der Zeit entschied sich BeobachterIn A entweder für eine (6) *hohe Feinfühligkeit* oder eine (7) *sehr hohe Feinfühligkeit*, im Gegensatz zu Rater B, der nur in 8,3% der Fälle so entschied. Allerdings werteten diesmal die beiden BeurteilerInnen die Ausprägung (5) *eher hohe*

*Feinfühligkeit* mit 25% gleich häufig. Ein weiteres Ungleichgewicht herrscht bei den Ausprägungen (3) *eher geringe Feinfühligkeit* und (4) *mittlere Feinfühligkeit* vor, wo 63,9% des einen Raters 13,9% des anderen gegenüber stehen (siehe Tabelle 13).

**Tabelle 13:** absolute und relative Häufigkeiten der Kodierungen auf der Skala *Feinfühligkeit* über alle *Pflege-Situationen* ( $N = 36$ )

|                                 | <i>absolute Häufigkeit</i> |         | <i>relative Häufigkeit</i> |         |
|---------------------------------|----------------------------|---------|----------------------------|---------|
|                                 | Rater A                    | Rater B | Rater A                    | Rater B |
| (1) sehr geringe Feinfühligkeit | 1                          | 0       | 2,8                        | 0       |
| (2) geringe Feinfühligkeit      | 0                          | 1       | 0                          | 2,8     |
| (3) eher geringe Feinfühligkeit | 3                          | 13      | 8,3                        | 36,1    |
| (4) mittlere Feinfühligkeit     | 2                          | 10      | 5,6                        | 27,8    |
| (5) eher hohe Feinfühligkeit    | 9                          | 9       | 25,0                       | 25,0    |
| (6) hohe Feinfühligkeit         | 15                         | 0       | 41,7                       | 0       |
| (7) sehr hohe Feinfühligkeit    | 6                          | 3       | 16,7                       | 8,3     |
| Gesamt                          | 36                         | 36      | 100                        | 100     |

Wie aus Tabelle 14 ersichtlich wird, liegt der Mittelwert von BeobachterIn A ( $M = 5.42$ ,  $S^2 = 1.79$ ) über dem von B ( $M = 4.08$ ,  $S^2 = 1.51$ ).

**Tabelle 14:** Mittelwert und Varianz der Beobachtungen auf der Dimension *Feinfühligkeit* über alle *Pflege-Situationen* ( $N = 36$ )

|                | $N$ | $M$  | $S^2$ | $Min$ | $Max$ |
|----------------|-----|------|-------|-------|-------|
| BeobachterIn A | 36  | 5.42 | 1.79  | 1     | 7     |
| BeobachterIn B | 36  | 4.08 | 1.51  | 2     | 7     |

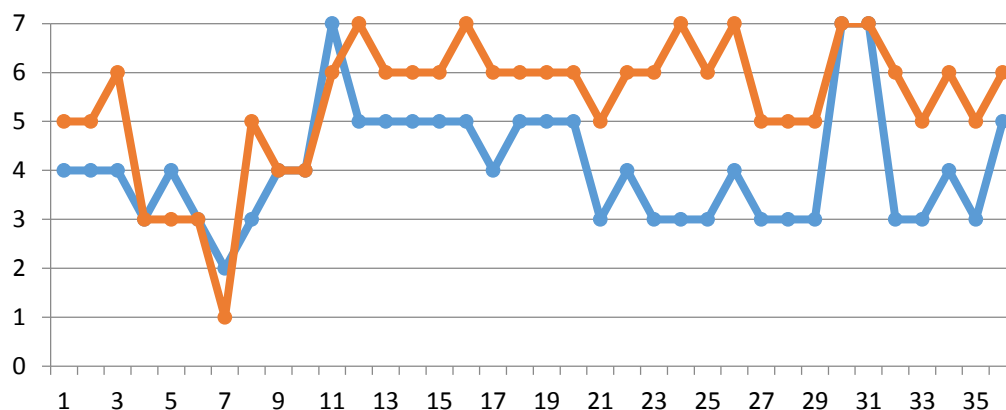
In 6 von 36 *Feinfühligkeits*-Videosequenzen der *Pflege-Situationen* herrscht eine exakte Übereinstimmung vor. Insgesamt 12 Mal gab es eine Differenz von einem Wertpunkt zwischen den Ratern. Somit waren sich die Rater in 50% der *Pflege*-Videos entweder exakt (keine Differenz) oder annähernd (Differenz von einem Wertpunkt) einig. 13 Mal gab es einen Unterschied von zwei Ausprägungen, vier Mal einen Unterschied von drei Skalenpunkten und einmal eine Diskrepanz von vier Merkmalsausprägungen auf der Skala *Feinfühligkeit* (siehe Tabelle 15 und Anhang, Abbildung A-1c).

## Ergebnisse der Untersuchung

**Tabelle 15:** Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Pflege-Situationen ( $N = 36$ )

|                      | <i>absolute Häufigkeit</i> | <i>relative Häufigkeit</i> |
|----------------------|----------------------------|----------------------------|
|                      | 0                          | 6                          |
|                      | 1                          | 12                         |
|                      | 2                          | 13                         |
| Betrag der Differenz | 3                          | 4                          |
|                      | 4                          | 1                          |
|                      | 5                          | 0                          |
|                      | 6                          | 0                          |
| Gesamt               | 36                         | 100                        |

Abbildung 5 verdeutlicht abermals die höher angesiedelten Beurteilungen von BeobachterIn A im Gegensatz zu BeobachterIn B. Von den 36 Pflege-Sequenzen waren sich die Rater bei 6 Intervallen (16,7%) einig. Am Häufigsten gab es eine Differenz von einer (33,3%) oder zwei (36,1%) Ausprägungsstufen.



**Abbildung 5:** Kodierungsverlauf der beiden Rater über alle FFK-Sequenzen der Pflege-Situationen ( $N = 36$ )

Die Ergebnisse von Tukeys Additivitätstest ( $F_{nonadd} = .369$ ;  $p = .547 > \alpha = .25$ ) bestätigen die Voraussetzung der Intraklassenkorrelation, dass keine signifikante Interaktion zwischen den Ratern und den Probanden stattgefunden hat. Durch das nicht signifikante Ergebnis des Levene-Tests ( $p = .718 > \alpha = .25$ ) gilt die Voraussetzung der Varianzhomogenität. Es konnte allerdings keine Normalverteilung für die Bewertungen der

Pflege-Situationen durch die Berechnungen des Kolmogorov-Smirnov-Test ( $p = .000$  für Rater A und  $p = .001$  für Rater B,  $\alpha = .05$ , siehe Anhang Abbildungen A-3c und A-3d) angenommen werden. Da nur die Voraussetzung der Normalverteilung verletzt ist, kann der ICC-Koeffizient ohne Bedenken herangezogen werden (Wirtz & Caspar, 2002).

Die Intraklassenkorrelation für die Pflege-Situationen (36 FFK-Sequenzen und 8 Videos von 6 Mutter-Kind-Dyaden) kann mit einem signifikanten Koeffizienten von  $r_{icc} = .360$  ( $F = 3.442$ ;  $p = .000 < \alpha = .05$ ) als eher gering angesehen werden. Das 95%-Vertrauensintervall besitzt die *Untergrenze* =  $-.075$  und die *Obergrenze* =  $.664$  (siehe Tabelle 16). Auch hier konnte in der Varianzanalyse (Faktor Rater) ein signifikanter Unterschied zwischen den Ratermittelwerten ( $F = 43.077$ ;  $p = .000 < \alpha = .25$ ) eruiert werden. Der eher niedrige ICC-Koeffizient von  $.360$  kann deswegen auch auf die Heterogenität der Mittelwerte zurückgeführt werden. Nach Wirtz und Caspar (2002, S. 187) könnte „durch eine Angleichung der Ratermittelwerte eine deutliche Verbesserung der Reliabilität erreicht werden.“

**Tabelle 16:** Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Pflege-Situationen ( $N = 36$ )

|               | Korrelation<br>innerhalb der<br>Klasse | 95% Konfidenzintervall |            | F-Test mit wahrem Wert 0 |     |     |      |
|---------------|----------------------------------------|------------------------|------------|--------------------------|-----|-----|------|
|               |                                        | Untergrenze            | Obergrenze | Wert                     | df1 | df2 | Sig  |
| Einzelne Maße | .360                                   | -.075                  | .664       | 3.442                    | 35  | 35  | .000 |

### 11.1.3.2 Kodierprobleme im Pflege-Kontext

Unsicherheiten und Schwierigkeiten in der Entscheidungsfindung ergeben sich auf der Verhaltensdimension Feinfühligkeit speziell für Interaktionen im Pflegekontext. Da sich eine Mutter durch das Windel wechseln nicht primär an den Bedürfnissen des Kindes orientiert und auch nicht die eigenen Tätigkeiten in den Hintergrund stellen kann, denn das Kind kann sich gar nicht am Windel wechseln beteiligen, ist besonders die Beurteilung solcher Sequenzen problematisch. Die Formulierung aus Ankerpunkt 5 des Manuals (Hirschmann et al., 2013), dass gemeinsam an einer Sache gearbeitet wird, ist für solche Situationen nicht zutreffend und verlangt nach einer Überarbeitung für diesen speziellen Kontext. Stärker zu betonen sind, ob die Mutter die Signale des Kindes wahrnimmt und

angemessen und prompt darauf reagiert und ob sie liebevoll und würdigend auf die kindlichen Initiativen eingeht, beziehungsweise ob die Interaktion reibungslos und harmonisch verläuft. Weniger feinfühligere Mütter tendieren in Wickelvideos dazu, das Wickeln innerhalb kürzester Zeit sehr rasch und eilig auszuführen, ohne das Kind mit einzubeziehen, es in ein Gespräch zu verwickeln oder als Ablenkung Spielmaterialien anzubieten. Hier war es besonders schwierig, ein Urteil zu fällen, weil sich das Kind nicht wehrte, sich kooperativ zeigte und das Wickeln geduldig über sich ergehen ließ. Somit schien die Interaktion reibungslos und harmonisch zu verlaufen. Für solche Fälle ist klar heraus zu arbeiten, was die Kriterien für die unterschiedlichen Abstufungen *sehr geringe*, *geringe* oder *eher geringe Feinfühligkeit* im Wickelkontext sind. Hier werden Entscheidungsregeln für eine präzise und sichere Entscheidung zwischen den benachbarten Skalenpunkten benötigt. Wickel-Videos sind im Generellen zeitlich sehr kurz, daher ist es schwierig, die Interaktion anhand der Ratingskala zu bewerten. Sie dauern meistens nur fünf Minuten und lassen somit nur zwei Bewertungen zur Feinfühligkeit zu.

Auch für Bade-Situationen zeigt sich das Manual auf der Ebene der Feinfühligkeit als nicht immer passend. Da die Ankerpunkte auf Spielsituationen ausgelegt sind, erschwert es oft die Entscheidung zwischen zwei Skalenpunkten. Ein Großteil des Badens kann zwar wie ein Spiel ablaufen, wenn z.B. mit Kübeln in der Wanne gespielt wird, das ist aber nicht immer der Fall. Wenn die Mutter das Kind wäscht, braucht es andere Kriterien und Entscheidungsregeln. Sehr oft wäscht die Mutter das Kind, während dieses mit Kübeln oder einer Spielzeugente spielt. Hier steht dann nicht im Zentrum der Bewertung, ob Mutter und Kind gemeinsam an einer Sache arbeiten oder ob die Mutter ihre eigenen Tätigkeiten in den Hintergrund stellt, beziehungsweise ob die kindlichen Problemlösefähigkeiten gefördert werden. Es wäre hilfreich und sinnvoll zu erfassen, wie abwechslungsreich und phantasievoll die Mutter die Pflege-Situation gestaltet. Geht sie stumm und eintönig vor und interagiert überhaupt nicht mit dem Kind oder ist sie kreativ und bezieht das Kind mit ein, verwickelt es in ein Gespräch. Diese Marker müssten für Pflege-Situationen neu festgelegt werden.

### 11.1.4 Betrachtung der Spielsituationen

#### 11.1.4.1 Ergebnisse zum Spiel-Kontext

Für die Situationen, in denen ein Spiel Gegenstand der Aufnahme war, kann abermals abgelesen werden (siehe Tabelle 17), dass BeobachterIn A höhere Bewertungen erzielte als B. 65% der Zeit wertete BeobachterIn A entweder eine (6) *hohe Feinfühligkeit* oder eine (5) *eher hohe Feinfühligkeit*, im Gegensatz zu ihrem Kollegen, der sich nur in 10% der Fälle so entschied. Umgekehrt kodierte BeobachterIn B 82,5% der Zeit entweder eine (4) *mittlere Feinfühligkeit* oder eine (3) *eher geringe Feinfühligkeit*, BeobachterIn A hingegen nur in 30%.

**Tabelle 17:** absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Spielsituationen ( $N = 40$ )

|                                 | absolute Häufigkeit |         | relative Häufigkeit |         |
|---------------------------------|---------------------|---------|---------------------|---------|
|                                 | Rater A             | Rater B | Rater A             | Rater B |
| (1) sehr geringe Feinfühligkeit | 0                   | 0       | 0                   | 0       |
| (2) geringe Feinfühligkeit      | 2                   | 0       | 5,0                 | 0       |
| (3) eher geringe Feinfühligkeit | 5                   | 7       | 12,5                | 17,5    |
| (4) mittlere Feinfühligkeit     | 7                   | 26      | 17,5                | 65,0    |
| (5) eher hohe Feinfühligkeit    | 13                  | 3       | 32,5                | 7,5     |
| (6) hohe Feinfühligkeit         | 13                  | 1       | 32,5                | 2,5     |
| (7) sehr hohe Feinfühligkeit    | 0                   | 3       | 0                   | 7,5     |
| Gesamt                          | 40                  | 40      | 100                 | 100     |

Die Mittelwerte für die Skala *Feinfühligkeit* – wie aus Tabelle 18 ersichtlich – nähern sich in den Spiel-Situationen allmählich an. BeobachterIn A stuft Mütter in Spiel-Situationen durchschnittlich mit einem Mittelwert von  $M = 4.75$  ( $S^2 = 1.42$ ) ein, BeobachterIn B mit einem Wert von  $M = 4.18$  ( $S^2 = 1.02$ ).

**Tabelle 18:** Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Spielsituationen ( $N = 40$ )

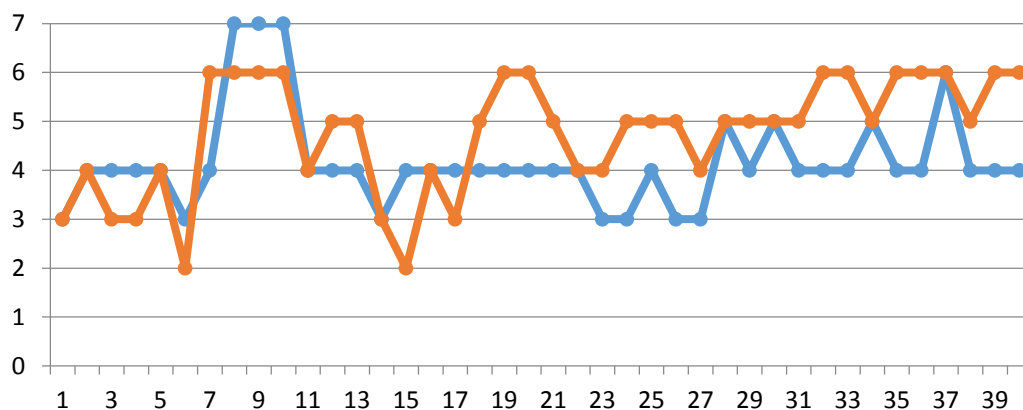
|         | $N$ | $M$  | $S^2$ | $Min$ | $Max$ |
|---------|-----|------|-------|-------|-------|
| Rater A | 40  | 4.75 | 1.42  | 2     | 6     |
| Rater B | 40  | 4.18 | 1.02  | 3     | 7     |

In 28 von 40 Spielsequenzen (70%), gibt es zwischen den Ratern auf der Ebene der *Feinfühligkeit* eine perfekte Übereinstimmung oder eine Differenz von einer Ausprägungsstufe. Tabelle 19 erläutert des Weiteren, dass bei 12 Sequenzen eine Differenz von zwei Skalenpunkten vorliegt (siehe Anhang, Abbildung A-1d).

**Tabelle 19:** Differenzen zwischen den beiden Ratern auf der Dimension *Feinfühligkeit* über alle Spiel-Situationen ( $N = 40$ )

|                      |   | <i>absolute Häufigkeit</i> | <i>relative Häufigkeit</i> |
|----------------------|---|----------------------------|----------------------------|
| Betrag der Differenz | 0 | 11                         | 27,5                       |
|                      | 1 | 17                         | 42,5                       |
|                      | 2 | 12                         | 30,0                       |
|                      | 3 | 0                          | 0                          |
|                      | 4 | 0                          | 0                          |
|                      | 5 | 0                          | 0                          |
|                      | 6 | 0                          | 0                          |
| Gesamt               |   | 40                         | 100                        |

Für die Spiel-Situationen kann eine optische Annäherung der Ratings beider BeurteilerInnen im Gegensatz zu den anderen beiden Situationen verzeichnet werden (siehe Abbildung 6 und Anhang, Abbildungen A-1b bis -1d). Für diese Situation gibt es mit 27,5% (im Vergleich zu 16,7% für Pflege- und 17,2% für Essens-Situationen) den höchsten relativen Anteil an exakter Übereinstimmung zu verzeichnen. 42,4% der Sequenzen wurden mit einer einstufigen Differenz beurteilt und 30% mit einer zweistufigen.



**Abbildung 6:** Kodierungsverlauf der beiden Rater über alle FFK-Sequenzen der Spiel-Situationen ( $N = 40$ )

Tukeys Additivitätstest ( $F_{nonadd} = 1.354$ ;  $p = .252 > \alpha = .25$ ) bekräftigt die Voraussetzung der Intraklassenkorrelation, dass keine signifikante Interaktion zwischen Ratern und beurteilten Dyaden vorliegt. Die Voraussetzung der Varianzhomogenität konnte für die Spielvideos durch den Levene-Test allerdings nicht bestätigt werden ( $p = .041 < \alpha = .25$ ), genauso wenig wie die Voraussetzung der Normalverteilung ( $p = .000$  für beide Rater,  $\alpha = .05$ , siehe Anhang Abbildungen A-3e und A-3f) erfüllt ist.

Da die Voraussetzungen der Varianzhomogenität und der angenäherten Normalverteilung der Spiel-Videos der Gegenkodierungsstichprobe nicht erfüllt sind, kann es deswegen zu einer tendenziell falsch eingeschätzten Reliabilität kommen. Die Intraklassenkorrelation von  $r_{icc} = .407$  (Tabelle 20) für die Spiel-Situationen (40 Feinfühligkeitssequenzen von insgesamt 7 Videos und 6 Mutter-Kind-Dyaden) kann aufgrund dessen nicht herangezogen werden.

Der Raterfaktor in der zweifaktoriellen Varianzanalyse fiel signifikant aus ( $F = 9.962$ ;  $p = .003 < \alpha = .25$ ), was für einen signifikanten Unterschied zwischen den Ratermittelwerten spricht.

**Tabelle 20:** Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Spiel-Situationen ( $N = 40$ )

|               | Korrelation<br>innerhalb der<br>Klasse | 95% Konfidenzintervall |            | F-Test mit wahrem Wert 0 |     |     |      |
|---------------|----------------------------------------|------------------------|------------|--------------------------|-----|-----|------|
|               |                                        | Untergrenze            | Obergrenze | Wert                     | df1 | df2 | Sig  |
| Einzelne Maße | .407                                   | .111                   | .636       | 2.680                    | 39  | 39  | .001 |

#### 11.1.4.2 Kodierprobleme im Spiel-Kontext

Die Beurteilung der Interaktionssequenzen im Spiel-Kontext, fiel insgesamt unproblematisch aus. Erwähnt sei dennoch, dass die Sequenzen in einem unstandardisierten Setting stattfanden und weder Spielmaterialien, noch Instruktionen vorgegeben wurden.

## 11.2 Skala Rückmeldung

Für die Kategorie *Rückmeldung* wurde Kappa mittels Mangolds ‚INTERACT 9‘ errechnet, das immer ein Event-basiertes Kappa ausgibt. Es wurde eine Toleranz von 50% und 3 Sekunden Überlappungsbereich gewählt. Auch in dieser Kategorie wurde die Ausprägung (d) *Unkodierbar – Rückmeldung* von der statistischen Analyse ausgeschlossen. Dies erfolgte über die Erstellung von sogenannten Kontingenztabellen in Microsoft Office Excel. Des Weiteren sei angeführt, dass Landis und Koch (1977, S. 165) die Stärke der Übereinstimmung mit folgenden Kappa-Koeffizienten determinieren:

| <i>Kappa</i> | <i>Stärke der Übereinstimmung</i> |
|--------------|-----------------------------------|
| 0.00-0.20    | gering                            |
| 0.21-0.40    | ausreichend                       |
| 0.41-0.60    | mittelmäßig                       |
| 0.61-0.80    | beträchtlich                      |
| 0.81-1.00    | (nahezu) perfekt                  |

### 11.2.1 Interrater-Reliabilität der Kategorie Rückmeldung

Ein ausreichend großes *event alignment* von  $\kappa = .358^1$  sowie  $\kappa = .359^2$  konnte für die Bewertungen über alle Situationen mit Mangolds ‚INTERACT 9‘ (Toleranzbereich von 50% und 3 Sekunden) eruiert werden. Stellt man diese Berechnung der BeobachterInnenübereinstimmung mittels ‚GSEQ 5.1‘ (Überlappungsbereich von 50% und 5 Events) an, so erhält man, je nach Reihenfolge der Rater, ein mittelmäßiges *event alignment* von  $\kappa = .451$  sowie  $\kappa = .454$  (siehe Tabelle 21).

---

<sup>1</sup> Die Daten von BeobachterIn A werden mit denen von BeobachterIn B verglichen.

<sup>2</sup> Die Daten von BeobachterIn B werden mit denen von BeobachterIn A verglichen.

## Ergebnisse der Untersuchung

**Tabelle 21:** Kappa-Koeffizienten der Beobachtungen der beiden Rater über alle Situationen, in Abhängigkeit des Auswertungsprogramms und der Reihenfolge der BeobachterInnen

|                        | <i>MANGOLD</i><br><i>Kappa</i> | <i>GSEQ</i><br><i>Kappa</i> |
|------------------------|--------------------------------|-----------------------------|
| Rater A versus Rater B | .358                           | .451                        |
| Rater B versus Rater A | .359                           | .454                        |

In den Tabellen 22 und 23 sind die Übereinstimmungen und Nichtübereinstimmungen der beiden Rater in Form von Kontingenztabellen detailliert dargestellt. Die Werte der Diagonale bezeichnen die Anzahl der Übereinstimmungen der beiden BeobachterInnen, die Werte abseits der Diagonale die Nichtübereinstimmungen, die eine BeobachterIn vergibt also eine andere Kodierung als die zweite BeobachterIn (Bakeman, Quera & Gnisci, 2009). Die Spalte und Zeile mit der Beschriftung „nil“ drückt den *commission-omission error* aus. Hierbei wird jegliches Ereignis identifiziert, das nur von einem der beiden BeobachterInnen mit einer Kodierung versehen wurde. Kodiert BeobachterIn 1 ein Ereignis, nicht jedoch BeobachterIn 2, nennt man diesen Umstand *deletion*, es handelt sich um den *omission error*. Wird eine Kodierung jedoch nur von BeobachterIn 2 getätigt, ist dies der *commission error*, was auch als *insertion* bezeichnet wird (Bakeman & Quera, 2011).

Wie aus Tabelle 22 ersichtlich ist, gibt es 320 Übereinstimmungen (man betrachte die färbig unterlegte Diagonale), wovon 250 Übereinstimmungen auf die Verhaltenskategorie *Keine Rückmeldung* entfallen. Für die *Korrigierende Rückmeldung* sind insgesamt 23 Übereinstimmungen zu verzeichnen. Jedoch kodierte BeobachterIn B diese Kategorie 40 Mal nicht, obwohl es KollegIn A tat, und diese wertete 22 Mal nicht, obwohl es ihre KollegIn tat. Bezüglich der *Negativen Rückmeldung* gab es 20 Übereinstimmungen. Hier kodierte Rater B 50 Events nicht, die BeobachterIn A tätigte, umgekehrt waren es 6 Events. Für die *Positive Rückmeldung* konnten 27 Übereinstimmungen ausfindig gemacht werden. 21 Mal kodierte BeobachterIn B eine *Positive Rückmeldung*, die A nicht tätigte, und 5 Mal wertete B nicht, obwohl BeurteilerIn A dies tat.

## Ergebnisse der Untersuchung

**Tabelle 22:** Kontingenztabelle zur Errechnung des event alignment der Kategorie Rückmeldung über alle Situationen, BeobachterIn A im Vergleich zu BeobachterIn B, erstellt mit GSEQ 5.1

|                    | <i>nil</i> | <i>keine RM</i> | <i>korrig. RM</i> | <i>negative RM</i> | <i>positive RM</i> | $\Sigma$ |
|--------------------|------------|-----------------|-------------------|--------------------|--------------------|----------|
| <i>nil</i>         | 0          | 15              | 22                | 6                  | 21                 | 64       |
| <i>keine RM</i>    | 11         | 250             | 2                 | 1                  | 0                  | 264      |
| <i>korrig. RM</i>  | 40         | 3               | 23                | 3                  | 1                  | 70       |
| <i>negative RM</i> | 50         | 1               | 1                 | 20                 | 0                  | 72       |
| <i>positive RM</i> | 5          | 1               | 0                 | 0                  | 27                 | 33       |
| $\Sigma$           | 106        | 270             | 48                | 30                 | 49                 | 503      |

Tabelle 23 zeigt den Vergleich der Bewertungen der KollegIn B mit denen von BeobachterIn A. Insgesamt gab es 321 Übereinstimmungen (Summe der farbigen Diagonale), davon fielen 252 auf die Dimension *Keine Rückmeldung*. Es konnten weiters 22 Übereinstimmungen für die Verhaltenskategorie *Korrigierende Rückmeldung* eruiert werden. Jedoch gab es hier 43 Events, die Rater B nicht tätigte, obwohl es BeobachterIn A tat, umgekehrt waren es 21 Events. Auf der Dimension *Negative Rückmeldung* zeigten sich 20 Übereinstimmungen. 51 Mal kodierte BeurteilerIn B *Negative Rückmeldung* nicht, obwohl es Rater A tat, umgekehrt wertete BeobachterIn A 5 Events nicht, die aber ihre KollegIn B registrierte. Für die *Positive Rückmeldung* konnten 27 Übereinstimmungen ausfindig gemacht werden. Rater B kodierte 5 Events nicht, die A wertete und umgekehrt kodierte BeobachterIn A 21 Mal nicht, wohingegen B die Kodierung *Positive Rückmeldung* tätigte.

**Tabelle 23:** Kontingenztabelle zur Errechnung des event alignment der Kategorie Rückmeldung über alle Situationen, im Vergleich ist BeobachterIn B zu BeobachterIn A, erstellt mit GSEQ 5.1

|                    | <i>nil</i> | <i>keine RM</i> | <i>korrig. RM</i> | <i>negative RM</i> | <i>positive RM</i> | $\Sigma$ |
|--------------------|------------|-----------------|-------------------|--------------------|--------------------|----------|
| <i>nil</i>         | 0          | 11              | 43                | 51                 | 5                  | 110      |
| <i>keine RM</i>    | 10         | 252             | 3                 | 0                  | 1                  | 266      |
| <i>korrig. RM</i>  | 21         | 4               | 22                | 1                  | 0                  | 48       |
| <i>negative RM</i> | 5          | 1               | 4                 | 20                 | 0                  | 30       |
| <i>positive RM</i> | 21         | 1               | 0                 | 0                  | 27                 | 49       |
| $\Sigma$           | 57         | 269             | 72                | 72                 | 33                 | 503      |

Die Prozentuelle Übereinstimmung (PÜ) der beiden Rater, die allerdings nicht zufallsbereinigt ist, lässt sich aus den beiden Kontingenztabellen 22 und 23 errechnen (vgl. Kapitel 6.4). Sie beträgt über alle Situationen und Videos betrachtet, je nach Reihenfolge der BeobachterInnen, 63,6% (Tabelle 22) beziehungsweise 63,8% (Tabelle 23).

Die Kappa-Koeffizienten für die einzelnen Situationen sind aus Tabelle 24 ersichtlich. Sie wurden mit einem Toleranz-Bereich von 50% und 3 Sekunden in Mangolds ‚INTERACT 9‘ berechnet. Für die Essens-Situationen ergibt sich ein *event alignment* von  $\kappa = .305$  bzw.  $\kappa = .307$ , was nach Landis & Koch (1977) als ausreichend bezeichnet werden kann. Die Pflege-Situationen weisen ein ähnlich hohes Kappa mit .308 und .310 auf. Nur die Spiel-Situationen zeigen eine höhere BeobachterInnenübereinstimmung mit  $\kappa = .501$  und  $\kappa = .499$ , was als mittelmäßig eingestuft werden kann.

**Tabelle 24:** Kappa-Koeffizienten für die Kategorie Rückmeldung pro Situation in Abhängigkeit der Reihenfolge der Rater, erstellt mit Mangold INTERACT 9

|                        | <i>Essen</i> | <i>Pflege</i> | <i>Spiel</i> | <i>Gesamt</i> |
|------------------------|--------------|---------------|--------------|---------------|
| Rater A versus Rater B | .305         | .308          | .501         | .358          |
| Rater B versus Rater A | .307         | .310          | .499         | .359          |

Die Prozentuellen Übereinstimmungen, die gegenüber dem Event-basierten Kappa den Nachteil der fehlenden Zufallsbereinigung haben, liegen für die einzelnen Situationen zwischen 63,4% und 72,4% (siehe Tabelle 25).

**Tabelle 25:** Prozentuelle Übereinstimmungen für die Kategorie Rückmeldung pro Situation in Abhängigkeit der Reihenfolge der Rater

|                        | <i>Essen</i> | <i>Pflege</i> | <i>Spiel</i> | <i>Gesamt</i> |
|------------------------|--------------|---------------|--------------|---------------|
| Rater A versus Rater B | 64,30%       | 63,40%        | 72,40%       | 63,60%        |
| Rater B versus Rater A | 64,70%       | 63,70%        | 72,10%       | 63,80%        |

### 11.2.2 Kodierprobleme der Rückmeldung

Laut Manual (Hirschmann et al., 2013) werden unter negativen Rückmeldungen „negative Äußerungen der Bezugsperson bezüglich der Person des Kindes und/oder dessen Ergebnisse oder Handlungen erfasst, d.h. auf emotional negative Art drückt die Bezugsperson ihre Ablehnung/Unzufriedenheit bezüglich des kindlichen Verhaltens aus“ (S. 14). Eine unterschwellig negativ ausgeübte Kritik, die aber in einem (gespielt) freundlichen und neutralen Tonfall getätigt wurde, kann laut Manual nicht als *Negative Rückmeldung* kodiert werden, da der Tonfall nicht emotional negativ gefärbt ist. Immer wieder tauchen Situationen und Aussagen auf, die sich im Grenzbereich zwischen negativem Ton und neutralem Ton bewegen, aber auch nicht als *Korrigierende Rückmeldung* gewertet werden können, da keine Begründung oder Erklärung erfolgt und somit die Hilfestellung fehlt. Aussagen, die eine negative Rückmeldung transportieren und eindeutig die Unzufriedenheit der Mutter ausdrücken, die aber vom Tonfall nicht eindeutig zuordenbar sind, werden von dem einen Rater als *Negative Rückmeldung* gewertet, der andere oder die andere entscheidet sich jedoch für *Keine Rückmeldung*, da der Tonfall für ihn oder sie zu wenig „negativ“ klingt. Es gibt bezüglich der Wertung *Negative Rückmeldung* einen gewissen Ermessens- und Interpretationsspielraum.

Es gab unter den Müttern eine Frau, die ihr Kind heimlich – sie versuchte es vor der Kamera zu verstecken – zwickte und das Kind böse mit den Augen fixierte. Durch diese physische und negative Reaktion drückte sie eindeutig ihre Unzufriedenheit aus. Die Sequenz wurde jedoch als *Keine Rückmeldung* bewertet, obwohl die Reaktion der Mutter zweifellos eine negative Rückmeldung für das Kind darstellte, da es sich um keine verbale Äußerung handelte. Besonders weil Kinder des jüngeren Altersbereichs noch nicht sprechen können, sollte angedacht werden, auch negative Handlungen gegenüber dem Kind als *Negative Rückmeldung* für die Bewertung zuzulassen.

In Tabelle 26 sind die Unterschiede in den Bewertungen der beiden BeurteilerInnen der Kategorie *Rückmeldung* und ihre Häufungen dargestellt. ‚Sehr oft‘ trat 40 bis 50 Mal auf, ‚manchmal‘ 10 bis 20 Mal und ‚selten‘ 5 Mal oder weniger. Die größte Diskrepanz gab es bei der Entscheidung, ob eine Sequenz als *Korrigierende Rückmeldung* oder *Keine Rückmeldung* beziehungsweise *Negative Rückmeldung* oder *Keine Rückmeldung* zu bewerten ist. Es ist ersichtlich, dass die häufigste Nichtübereinstimmung zwischen den

beiden Ratern sich darauf bezog, dass BeobachterIn A entweder *Korrigierende Rückmeldung* oder *Negative Rückmeldung* angab, während sich BeobachterIn B stattdessen für *Keine Rückmeldung* entschied.

**Tabelle 26:** Darstellung der Diskrepanzen zwischen den Ratern unter Angabe der Häufigkeit

|                         | Kodierung<br>BeobachterIn A | Kodierung<br>BeobachterIn B |
|-------------------------|-----------------------------|-----------------------------|
| sehr oft<br>(40-50 Mal) | Korrigierende RM            | Keine RM                    |
|                         | Negative RM                 | Keine RM                    |
| Manchmal<br>(10-20 Mal) | Keine RM                    | Korrigierende RM            |
|                         | Keine RM                    | Positive RM                 |
| Selten<br>(≤ 5 Mal)     | Positive RM                 | Keine RM                    |
|                         | Korrigierende RM            | Negative RM                 |
|                         | Keine RM                    | Negative RM                 |

### 11.3 Skala Joint Attention

Für die Kategorie *Joint Attention* wurde Cohens Kappa mittels ‚GSEQ 5.1‘ errechnet. Es wurde ein Time-Unit basiertes Kappa mit einem toleranten Überlappungsbereich von  $\pm 3$  Sekunden kalkuliert. Auch in dieser Kategorie wurde die Ausprägung (*u*) *Unkodierbar – Joint Attention* von der statistischen Analyse ausgeschlossen. Dies erfolgte über die Erstellung von sogenannten Kontingenztabellen in Microsoft Office Excel. Auch für diese Kategorie relevant, ist die von Landis und Koch (1977, S. 165) vorgeschlagene Gliederung der Stärke der Übereinstimmung:

| Kappa     | Stärke der Übereinstimmung |
|-----------|----------------------------|
| 0.00-0.20 | gering                     |
| 0.21-0.40 | ausreichend                |
| 0.41-0.60 | mittelmäßig                |
| 0.61-0.80 | beträchtlich               |
| 0.81-1.00 | (nahezu) perfekt           |

### 11.3.1 Interrater-Reliabilität der Kategorie Joint Attention

Für die Kategorie *Joint Attention* konnte über alle Situationen ein Time-Unit basiertes Kappa (Toleranz  $\pm 3$  Sekunden) von  $\kappa = .475$  und  $\kappa = .399$  in ‚GSEQ 5.1‘ errechnet werden. Aus Tabelle 27 geht außerdem eine Prozentuelle Übereinstimmung von  $P\ddot{U} = 55,7\%$  bzw.  $63,2\%$  hervor.

**Tabelle 27:** Kappa-Koeffizienten und Prozentuelle Übereinstimmungen für die Kategorie *Joint Attention* über alle Situationen in Abhängigkeit der Reihenfolge der Rater, erstellt mit GSEQ 5.1

|                        | Kappa | $P\ddot{U}$ |
|------------------------|-------|-------------|
| Rater A versus Rater B | .475  | 63,2%       |
| Rater B versus Rater A | .399  | 55,7%       |

Kontingenztafel 28 stellt den Vergleich der Bewertungen von BeobachterIn A mit denen von BeobachterIn B dar. Bei der Betrachtung der farbigen Diagonale ist ersichtlich, dass es 13144 Übereinstimmungen von insgesamt 20770 Events gibt. Die Werte abseits der Diagonale stehen für alle Nichtübereinstimmungen der beiden Rater. Die Werte in den Reihen stehen für Rater A, die Werte in den Spalten stehen für Rater B (Bakeman & Quera, 2011). Es ist zu erkennen, dass die meisten Nichtübereinstimmungen der BeobachterIn B mit den Bewertungen von BeobachterIn A bezüglich der Kodierungen *Aktive Aufrechterhaltung auf der Handlungsebene* auf die Verhaltensdimensionen *Aufrechterhaltung auf der verbalen Ebene* und *Passive Aufrechterhaltung* fallen. Auf der Ebene der *Passiven Aufrechterhaltung* scheinen sich die beiden BeurteilerInnen sehr einig zu sein, es gibt 1392 Übereinstimmungen und nur 104 Nichtübereinstimmungen. Wenn BeobachterIn A *Keine Joint Attention* kodierte, ist die größte Diskrepanz darin zu finden, dass Rater B 346 Mal stattdessen *Passive Aufrechterhaltung* wertete. Nur 79 Übereinstimmungen gibt es bei *Aufmerksamkeitswechsel zu einem neuen Spiel* und 605 Nichtübereinstimmungen. Wenn BeobachterIn A einen *Wechsel* kodierte, so bewertete BeobachterIn B in den meisten Fällen *Aufrechterhaltung auf der verbalen Ebene*, *Passive Aufrechterhaltung* oder *Aufmerksamkeitslenkung im laufenden Spiel*. Die Kategorie *Aufmerksamkeitslenkung im laufenden Spiel* weist 1092 Übereinstimmungen auf, jedoch entfallen 993 Nichtübereinstimmungen auf Kodierungen von Rater B wie *Aufrechterhaltung auf der verbalen Ebene*, *Aufrechterhaltung auf der Handlungsebene*

oder *Aufmerksamkeitswechsel*. Für die Dimension *Aufrechterhaltung auf der verbalen Ebene* gibt es eine hohe Übereinstimmung mit 2279 Events und nur 422 Nichtübereinstimmungen.

**Tabelle 28:** Kontingenztabelle zur Errechnung von Kappa der Kategorie Joint Attention über alle Situationen, im Vergleich ist BeobachterIn A zu BeobachterIn B, erstellt mit GSEQ 5.1

|                    | <i>Handl.ebene</i> | <i>verb. Ebene</i> | <i>Lenkung</i> | <i>Wechsel</i> | <i>keine JA</i> | <i>passive A.</i> | $\Sigma$     |
|--------------------|--------------------|--------------------|----------------|----------------|-----------------|-------------------|--------------|
| <i>Handl.ebene</i> | 7725               | 1728               | 284            | 161            | 470             | 2237              | <b>12605</b> |
| <i>verb. Ebene</i> | 144                | 2279               | 205            | 9              | 8               | 56                | <b>2701</b>  |
| <i>Lenkung</i>     | 283                | 467                | 1092           | 243            | 64              | 117               | <b>2266</b>  |
| <i>Wechsel</i>     | 60                 | 258                | 99             | 79             | 65              | 123               | <b>684</b>   |
| <i>keine JA</i>    | 59                 | 29                 | 6              | 1              | 577             | 346               | <b>1018</b>  |
| <i>passive A.</i>  | 57                 | 18                 | 0              | 0              | 29              | 1392              | <b>1496</b>  |
| $\Sigma$           | <b>8328</b>        | <b>4779</b>        | <b>1686</b>    | <b>493</b>     | <b>1213</b>     | <b>4271</b>       | <b>20770</b> |

Tabelle 29 zeigt die Übereinstimmungen (11609) und Nichtübereinstimmungen (9202) der Bewertungen von BeobachterIn B (Zeilen) mit den Bewertungen von BeobachterIn A (Spalten). Die Übereinstimmung bezüglich der Verhaltenskategorie *Aufrechterhaltung auf der Handlungsebene* ist mit 6311 Events im Vergleich zu 464 Nichtübereinstimmungen sehr hoch. Nur auf der *verbalen Ebene* konnten 2343 Nichtübereinstimmungen aufgedeckt werden, die bei BeobachterIn A auf die Dimension *Aufrechterhaltung auf der Handlungsebene* entfielen und weitere 483 Nichtübereinstimmungen ergeben sich aufgrund der Bewertung *Aufmerksamkeitslenkung im laufenden Spiel*. 971 Mal bewerteten die beiden BeobachterInnen Events mit einer *Aufmerksamkeitslenkung*, im Gegensatz zu 505 Nichtübereinstimmungen, die dadurch zustande kamen, dass BeobachterIn A entweder *Aufrechterhaltung auf der Handlungsebene* oder auf der *verbalen Ebene* kodierte. Für den *Aufmerksamkeitswechsel zu einem neuen Spiel* ergeben sich nur 63 Übereinstimmungen, 188 Events bewertete Rater A stattdessen mit der *Aufrechterhaltung auf der Handlungsebene* oder einer *Aufmerksamkeitslenkung*. 604 Mal entschied sich Rater B für die Bewertung *Keine Joint Attention*, ebenso BeobachterIn A und 541 Mal entschloss sie sich stattdessen für *Aufrechterhaltung auf der Handlungsebene*. Die Kodierungen von Rater B stimmten in 1690 Kodierungen mit denen von BeobachterIn A bezüglich der Kategorie *Passive Aufrechterhaltung* überein. Von den 3794 Nichtübereinstimmungen für

die Dimension *Passive Aufrechterhaltung*, entfielen die meisten auf die Dimension der *Handlungsebene* (in 2902 Fällen).

**Tabelle 29:** Kontingenztafel zur Errechnung von Kappa der Kategorie Joint Attention über alle Situationen, im Vergleich ist Rater B zu Rater A, erstellt mit GSEQ 5.1

|                    | <i>Handl.ebene</i> | <i>verb. Ebene</i> | <i>Lenkung</i> | <i>Wechsel</i> | <i>keine JA</i> | <i>passive A.</i> | $\Sigma$ |
|--------------------|--------------------|--------------------|----------------|----------------|-----------------|-------------------|----------|
| <i>Handl.ebene</i> | 6311               | 158                | 147            | 55             | 48              | 56                | 6775     |
| <i>verb. Ebene</i> | 2343               | 1970               | 483            | 262            | 30              | 23                | 5111     |
| <i>Lenkung</i>     | 279                | 226                | 971            | 86             | 6               | 0                 | 1568     |
| <i>Wechsel</i>     | 188                | 12                 | 235            | 63             | 1               | 0                 | 499      |
| <i>keine JA</i>    | 541                | 60                 | 61             | 59             | 604             | 49                | 1374     |
| <i>passive A.</i>  | 2902               | 333                | 128            | 120            | 311             | 1690              | 5484     |
| $\Sigma$           | 12564              | 2759               | 2025           | 645            | 1000            | 1818              | 20811    |

Betrachtet man die einzelnen Situationen genauer, so ist abermals – wie in der Kategorie *Rückmeldung* – zu erkennen, dass für die Spiel-Situationen das höchste Kappa vorliegt. Tabelle 30 zeigt ein Time-Unit Kappa von .648 und .527 für den Spiel-Kontext, je nach Reihenfolge des Vergleichs. Für den ersten Vergleich (Rater A versus Rater B) erhält man eine beträchtliche Übereinstimmung, für den zweiten Vergleich (Rater B versus Rater A) eine mittelmäßige. Die Werte der Pflege-Situationen ergeben für beide Fälle ein mittelmäßiges Kappa mit  $\kappa = .481$  und  $\kappa = .436$ . Die Essens-Situationen weisen die niedrigste Übereinstimmung mit  $\kappa = .382$  und  $\kappa = .318$  auf, was nach Landis & Koch (1977) als ausreichend zu bezeichnen ist.

**Tabelle 30:** Kappa-Koeffizienten für die Kategorie Joint Attention pro Situation, in Abhängigkeit der Reihenfolge der Rater, erstellt mit GSEQ 5.1

|                        | <i>Essen</i> | <i>Pflege</i> | <i>Spiel</i> | <i>Gesamt</i> |
|------------------------|--------------|---------------|--------------|---------------|
| Rater A versus Rater B | .382         | .481          | .648         | .475          |
| Rater B versus Rater A | .318         | .436          | .527         | .399          |

Die Prozentuellen Übereinstimmungen aus Tabelle 31 spiegeln die erhöhte Einigkeit der beiden Rater für die Spiel-Situationen (76,2% und 66,6%) wieder. Interessanterweise sind die Werte für die Essens- ( $P\ddot{U} = 59,4\%$ ) und Pflege-Situationen ( $P\ddot{U} = 59,0\%$ ) im einen

Vergleich (Rater A versus Rater B) annähernd gleich hoch, obwohl deren Kappa-Koeffizienten einen Unterschied von .100 aufwiesen (siehe Tabelle 30).

**Tabelle 31:** Prozentuelle Übereinstimmung für die Kategorie Joint Attention pro Situation in Abhängigkeit der Rater

|                        | <i>Essen</i> | <i>Pflege</i> | <i>Spiel</i> | <i>Gesamt</i> |
|------------------------|--------------|---------------|--------------|---------------|
| Rater A versus Rater B | 59,40%       | 59,00%        | 76,20%       | 63,20%        |
| Rater B versus Rater A | 51,60%       | 54,60%        | 66,60%       | 55,70%        |

### 11.3.2 Kodierprobleme der Joint Attention

Auf der Ebene der Joint Attention erwies sich folgendes Szenario einer Pflegesituation für die Bewertung als uneindeutig: In einer Badesituation spielte das Kind mit Wasserspielsachen, währenddessen wusch die Mutter dem Kind den Rücken. Führen die beiden, voneinander getrennte Handlungen aus und ist diese Sequenz somit als *Keine Joint Attention* zu beurteilen oder führen die beiden dem Baden untergeordnete Handlungen aus, weswegen es als *Aktive Aufrechterhaltung auf der Handlungsebene* zu werten ist. Hier ist mit Beispielen zu präzisieren, ab wann die Mutter Handlungen setzt, um die Joint Attention des Kindes aufrecht zu erhalten, wobei der Aufmerksamkeitsfokus des Kindes berücksichtigt wird, die auf der *Handlungsebene* zu werten sind, und ab wann solche Handlungen als *Keine Joint Attention* gewertet werden sollen. Hier gab es bei den BeurteilerInnen große Unsicherheiten und Unklarheiten.

Auch für die Essenssituationen war aus dem Manual nicht eindeutig abzuleiten, ob eine Szene, in der Mutter und Kind essen, ohne sich dabei anzusehen oder miteinander zu sprechen, auf der *Handlungsebene* oder als *Keine Joint Attention* zu werten ist. Reicht es aus, da beide an der Aktivität ‚Essen‘ beteiligt sind, dies auf der *Handlungsebene* zu kodieren oder reicht es nicht aus, den gemeinsamen Essensprozess als gemeinsame Beschäftigung anzusehen, an der sich die Bezugsperson beteiligt, um die Joint Attention des Kindes aufrecht zu erhalten und muss aufgrund fehlender Kommunikation *Keine Joint Attention* gewertet werden.

Für die Essens-Szenarios gab es eine weitere Unsicherheit, die sich auf die Abgrenzung *Aufmerksamkeitswechsel* von *Aktiver Aufrechterhaltung auf der Handlungsebene* und

*Aufmerksamkeitslenkung* bezieht. Sehr oft wertete ein Rater eine Sequenz als einen *Wechsel*, der andere oder die andere hingegen – je nach Situation – als *Lenkung* oder *Aktive Aufrechterhaltung auf der Handlungsebene*. Die Mutter-Kind-Dyaden erhielten zu Mittag immer drei Gänge (Suppe, Hauptgang mit Salat und Nachspeise). Manchmal standen alle drei Gänge auf dem Tisch, an dem gegessen wurde und manchmal stand nur eine Speise auf dem Tisch und die beiden anderen Gänge auf der Anrichte. Es war den BeurteilerInnen nicht ganz klar, ob der Übergang von einem Gang zum nächsten als *Aufmerksamkeitswechsel* zu kodieren ist oder nicht. Nicht immer gab es einen klaren Übergang zwischen den einzelnen Gängen, besonders nicht, wenn alle drei Gänge am Tisch platziert waren. Auch hier werden noch klarere Richtlinien benötigt, da jeder Beurteiler oder jede Beurteilerin diese Situation anders auslegen könnte. Ist der Wechsel von einem Gang zum nächsten als ein *Aufmerksamkeitswechsel* zu verstehen, da die Mutter die Aufmerksamkeit des Kindes von einer Beschäftigung, einem Objekt oder einem Ereignis auf etwas anderes lenkt, oder wäre ein *Wechsel* nur dann zu kodieren, wenn sie auf eine ganz andere Beschäftigung (wie z.B. Hände waschen) überleitet. Hier besteht eine große Diskrepanz zwischen den BeurteilerInnen (siehe Tabelle 32). Der *Wechsel* wurde fast nie zeitgleich kodiert. Es gab einen großen „Überlappungsbereich“ zwischen *Aufmerksamkeitswechsel*, *Aufmerksamkeitslenkung* und *Aktive Aufrechterhaltung auf der Handlungsebene*. Dem kann in Zukunft entgegen gewirkt werden, indem das Manual nicht nur auf Spielsituationen, sondern auch auf Essenssituationen eingeht und konkrete Beispiele anführt. Außerdem sollte auf eine standardisierte Vorgabe geachtet werden.

**Tabelle 32:** Muster der Unterschiede in den Bewertungen der beiden Rater auf der Ebene der Joint Attention

|          | Kodierung<br>BeobachterIn A             | Kodierung<br>BeobachterIn B        |
|----------|-----------------------------------------|------------------------------------|
| sehr oft | Handlungsebene<br>Lenkung<br>Wechsel    | verbal, passiv, keine JA<br>/<br>/ |
| manchmal | /<br>unkodierbar<br>keine JA<br>Lenkung | Lenkung<br>/<br>passiv<br>Wechsel  |
| selten   | /<br>Wechsel<br>passiv                  | Wechsel<br>Lenkung<br>keine JA     |

Anmerkung: Der ‚/‘ bedeutet, dass die Verhaltenskategorie nicht gewählt wurde.

Neben den Unstimmigkeiten der BeobachterInnen bezüglich *Lenkung* und *Wechsel*, entfielen die meisten Nichtübereinstimmungen auf die *Handlungsebene*. In vielen Fällen, in denen sich BeurteilerIn A für die *Handlungsebene* entschied, wertete ihre KollegIn B eben nicht auf der *Handlungsebene*, sondern je nach Situation auf der *verbalen Ebene*, *Passive Aufrechterhaltung* oder sogar *Keine Joint Attention*. Dies spiegelt die eingangs erwähnten Unklarheiten z.B. für Essenssituationen wieder, für die fraglich ist, ob eine Mutter, die isst, aber nicht mit dem Kind Blickkontakt hält oder kommuniziert, durch das gemeinsame Essen, auf der *Handlungsebene aktiv aufrechterhält* oder ob *Keine Joint Attention* zu werten ist.

Ein weiterer Punkt, der im Zuge der Essensvideos geklärt werden muss, ist folgender: Wenn Mutter und Kind von ihren eigenen Tellern essen und gerade kauen, während sie Blickkontakt halten, jedoch nicht sprechen, ist dies als *Passive Aufrechterhaltung der Joint Attention* zu kodieren oder auf der *Handlungsebene*, da sich die Mutter durch das Essen und Kauen an der Beschäftigung des Kindes aktiv beteiligt.

Weiters ist genau zu deklarieren, ab wann *Keine Joint Attention* zu werten ist. Ist dies ausschließlich der Fall, wenn zwei unterschiedliche, voneinander getrennte Aktivitäten ausgeführt werden (das Kind isst und die Mutter liest daneben ein Buch), oder ist *Keine Joint Attention* auch zu kodieren, wenn jeder von seinem eigenen Teller isst, es gibt also die gemeinsame Aktivität ‚Essen‘, es findet jedoch kein Blickkontakt statt und es wird auch nicht gesprochen, jegliche Form der Kommunikation fehlt somit.

Es ist darauf hinzuweisen, dass es in der Stichprobe auch Kinder gab, die noch nicht in der Lage waren selbstständig zu essen und somit gefüttert werden mussten oder sogar noch die Flasche bekamen. Für diese Fälle ist oft unklar, wann eine Mutter *aktiv auf der Handlungsebene aufrechterhält* und ab wann sie *passiv aufrechterhält*, da sich das Kind eben nicht selber ernähren kann und die Mutter somit nicht durch Zusehen *passiv aufrechterhalten* kann, wie es im Manual für die Altersgruppe der 3- bis 6-Jährigen festgelegt ist. Hier ist ein explizites Eingehen auf die jüngere Altersgruppe in Essenssituationen notwendig. Ein konkretes Beispiel betrifft das Verhalten einer Mutter, die über 10 bis 15 Minuten hinweg die Flasche gibt, ohne zu sprechen, da das Kind gerade am Einschlafen ist. Ihr Blick wandert vom Kind zum Fenster und sie starrt längere Zeit aus dem Fenster. Ist das Verhalten der Mutter auf der *Handlungsebene* zu kodieren, da sie die

Flasche gibt oder aber als *Keine Joint Attention*, weil der Blickkontakt fehlt. Wann ist das Verhalten einer Mutter, die ihrem Kind gerade die Flasche gibt, auf der *Handlungsebene* zu kodieren, wann findet *keine Joint Attention* statt und wann hält sie den Aufmerksamkeitsfokus des Kindes, das noch gefüttert werden muss, *passiv aufrecht*. Hier sind genaue Richtlinien festzulegen, ab wann diese Dimensionen erfüllt sind.

Eine weitere Sequenz kann als Beispiel genannt werden, um die Verhaltensdimensionen der *Joint Attention* für den Essenskontext zu präzisieren. Ist eine Szene, in der die Mutter, die ihrem Kind mit der einen Hand die Flasche gibt und mit der anderen Hand ‚SMS‘ schreibt, als *Aktive Aufrechterhaltung auf der Handlungsebene* zu bewerten, da sie das Kind zumindest mit der einen Hand füttert, oder als *Keine Joint Attention*, da ihr Aufmerksamkeitsfokus nicht auf dem Kind liegt, sondern der Blick auf dem Handy ruht. Für Essenssituationen ist speziell der Unterschied zwischen *Aktive Aufrechterhaltung auf der Handlungsebene*, *Passive Aufrechterhaltung* und *Keine Joint Attention* heraus zu arbeiten. Hier gab es einige Schwierigkeiten (siehe Tabelle 32).

Für den Fall, in dem das Kind durch eine Sonde gefüttert wird, ist es dem Rater ebenso schwer gefallen eine Entscheidung zwischen der *Handlungsebene*, *Keine Joint Attention* und *Passive Aufrechterhaltung* zu fällen. Die Mutter hat während der Sondenfütterung hin und wieder einen Bissen von ihrem eigenen Essen gemacht und das Kind dabei nicht angesehen, zählt dies als *Aktive Aufrechterhaltung auf der Handlungsebene*, *Passive Aufrechterhaltung* oder *Keine Joint Attention*. Wie die *Passive Aufrechterhaltung auf der Handlungsebene* bei einer Sondenfütterung im Detail aussieht, ist der BeurteilerIn unklar.

Immer wieder gibt es auf der Säuglingspsychosomatik des Wilhelminenspitals gravierende Fälle, wo Zwangsfütterung seitens der Mutter zur Anwendung kommt. Ein Kind weinte und wehrte sich gegen die Fütterung, die Mutter versuchte jedoch sehr brutal und unter Androhungen verzweifelt das Kind zum Essen zu bewegen. Irgendwann war das Kind dann so erschöpft, dass es in eine Art ‚Ohnmacht‘ gefallen ist und regungslos mit geöffneten Mund auf den Fernseher starrte. Die Mutter fütterte weiter, wobei das Kind den Brei nicht schluckte, es rann alles aus dem geöffneten Mund wieder heraus. Hier ergab sich die Frage, ob die Sequenz als *Keine Joint Attention* zu kodieren ist, da das Kind den Fernseher im Aufmerksamkeitsfokus hatte und die Mutter die Fütterung (es gab also keinen gemeinsamen Aufmerksamkeitsfokus) oder aber als *Aufmerksamkeitslenkung*, da

die Mutter noch immer gegen den Willen des Kindes fütterte. Es bedarf hier präziser Entscheidungsregeln, die die Bewertung von Essenssituationen eindeutig machen.

Auch für die Pflegesituationen, im speziellen für die Wickelvideos, ergeben sich auf der Verhaltensdimension der *Joint Attention* Schwierigkeiten. Da sich das Kind am Windel wechseln nicht beteiligen kann, ist die Zuordnung auf der Verhaltensdimension *Joint Attention* manchmal schwierig und nicht eindeutig. Ein Beispiel hierfür, wäre eine Mutter, die ihrem Kind die Windel wechselt, während dieses mit einem Spielzeug beschäftigt ist und das Wickeln kooperativ über sich ergehen lässt. Fraglich ist, ob diese Sequenz auf der *Handlungsebene* oder als *Keine Joint Attention* zu kodieren ist. Wenn letzteres zutrifft, ist genau zu deklarieren, ab wann die *Handlungsebene* erfüllt ist. Darüber hinaus sei angemerkt, dass die Mutter beim Wickeln gar nicht *passiv* werden kann und den Aufmerksamkeitsfokus des Kindes durch bloßes Zusehen aufrechterhalten kann, da sich das Kind nicht selber Wickeln kann.

## 12 Diskussion und Interpretation

Die hier vorliegende Arbeit stellt einen ersten Versuch dar, das Beobachtungsinstrument INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013) auf einen anderen Altersbereich und andere Situationen anzuwenden, sowie unter nicht standardisierten Bedingungen des klinischen Alltags zu erproben. Die Datenerhebung wurde retrospektiv durchgeführt, da durch die Stationsleitung bereits vorhandene Interaktionsvideos für die Dauer der Untersuchung zur Verfügung gestellt wurden. Die Interaktionsvideos der Station des Wilhelminenspitals wurden zu einem anderen Zweck als der wissenschaftlichen Erforschung und Erprobung eines Beobachtungsinstruments aufgenommen, nämlich zur therapeutischen Verlaufskontrolle und weisen daher keinerlei Strukturiertheit oder Standardisierung auf. Zum Zeitpunkt der Videoaufnahmen wurde weder eine Instruktion vorgegeben, noch unter standardisierten Bedingungen gedreht. Zur Bestimmung der Reliabilität der einzelnen Dimensionen (*Feinfühligkeit, Rückmeldung, Joint Attention*) des Beobachtungsinstruments INTAKT wurden die Übereinstimmungen zwischen den beiden Ratern eruiert.

28 Videos mit einer Gesamtdauer von 06:32:20 (h:min:sec) von 14 Mutter-Kind-Dyaden wurden von beiden BeobachterInnen beurteilt. Die 28 Videos setzen sich aus 13 Essens-, 8 Pflege- und 7 Spiel-Situationen sowie aus 19 Prä- und 9 Post-Situationen zusammen. Sowohl das Alter der Kinder als auch die unterschiedlichen Situationen und Erhebungszeitpunkte der Gegenkodierungsstichprobe sind repräsentativ für die Gesamtstichprobe. Des Weiteren ist die Verteilung des Kindesalters auf die verschiedenen Untersuchungssituationen in Gegenkodierungs- und Gesamtstichprobe ähnlich (siehe Kapitel 10.4).

Das INTAKT-Manual (Hirschmann et al., 2013) ist ursprünglich für Kinder in einem Alter zwischen 3 und 6 Jahren entwickelt worden. In der untersuchten Stichprobe wurde der Versuch gemacht, das Instrument auf einen jüngeren Altersbereich (1;2 - 3;5) anzuwenden. Bei der herangezogenen Stichprobe handelt es sich um eine klinische Stichprobe, bei deren Aufnahme auf der Säuglingspsychosomatik des Wilhelminenspitals in Wien akuter Handlungsbedarf bestand. Das jüngste Kind der Gegenkodierungsstichprobe war 14 Monate alt, der Altersdurchschnitt betrug 22,7 Monate (siehe Kapitel 10.3). Da bei jüngeren

Kindern andere Verhaltensweisen im Vordergrund stehen, da sie teilweise noch nicht sprechen oder selbstständig essen können, sollten die Erläuterungen im Manual auf den jüngeren Altersbereich angepasst werden.

Die Definitionen der verschiedenen Kategorien im Manual beziehen sich außerdem auf ein standardisiertes Setting und ausschließlich auf Situationen, in denen gebastelt oder gespielt wurde. Insbesondere die Ankerpunkte der Dimension *Feinfühligkeit* und die Beschreibung der verschiedenen Verhaltensdimensionen der Kategorie *Joint Attention* sind in ihren Formulierungen stark auf Spielsituationen ausgelegt. Die neuen Situationen im Essens- oder Pflegekontext verlangen nach zusätzlichen Definitionen, Richtlinien und Entscheidungsregeln.

### 12.1 Skala Feinfühligkeit

Zur Bestimmung der BeobachterInnenübereinstimmung für die Skala *Feinfühligkeit* wurde der Koeffizient der Intraklassenkorrelation herangezogen. Die Intraklassenkorrelation für die Dimension *Feinfühligkeit* konnte mit einem Wert von  $r_{icc} = .361$  über alle Videos und Situationen berechnet werden. Die eher geringe Übereinstimmung zeigt sich auch im Vergleich der absoluten Häufigkeiten der beiden Rater. BeobachterIn A hat 118 von 169 Sequenzen mit (5) *eher hohe Feinfühligkeit* oder (6) *hohe Feinfühligkeit* bewertet, BeobachterIn B hingegen 120 von 169 Sequenzen mit (3) *eher geringe Feinfühligkeit* oder (4) *mittlere Feinfühligkeit*. Wirtz und Caspar (2002) empfehlen jedoch „bei intervallskalierten Ratingskalen ist diese Forderung nach exakter Übereinstimmung der Werte meist nicht notwendig und auch nicht sinnvoll. ... Bei Urteilen auf kontinuierlichen Skalen liegt es in der Natur der Sache, dass exakte Übereinstimmung einen Ausnahmefall darstellt“ (S. 34). In 94 von 169 Bewertungen (56%) weisen die beiden Rater eine exakte Übereinstimmung oder eine Differenz von einem Skalenpunkt auf. Da es einen signifikanten Unterschied der Ratermittelwerte ( $M = 5.11$  für Rater A,  $M = 4.28$  für Rater B, siehe Kapitel 11.1.1.1) gibt, ist eine Verbesserung der ICC möglich, wenn die Mittelwerte der beiden BeobachterInnen angeglichen werden. Dies kann erzielt werden, indem das Manual in Zukunft auch auf Situationen mit Essens- und Pflegekontext konkret eingeht. Probleme, die sich in der Anwendung herauskristallisiert haben, müssen Eingang ins Manual finden.

Wenn man die einzelnen Situationen (Essen, Pflege, Spiel) näher betrachtet, so fällt auf, dass der höchste Anteil an Übereinstimmungen in den Spiel-Situationen gefunden werden kann. 70% Übereinstimmung (mit einem Toleranzbereich von einem Wertpunkt, siehe Wirtz & Caspar, 2002, S. 34) konnte hier erzielt werden. Für die Pflege- und Essens-Situationen schwankte dieser Wert allerdings um die 50%. Auch optisch war eine Annäherung der Ratings für die Spiel-Videos zu verzeichnen (siehe Abbildung 6, Kapitel 11.1.4.1). Dies spiegelt sich ebenfalls in den Mittelwerten der beiden Rater wieder ( $M = 4.75$ ,  $M = 4.18$ ). Die ICC, die für die Spiel-Situationen errechnet wurde, kann aufgrund mangelnder Voraussetzungen leider nicht interpretiert werden. Die BeobachterInnenübereinstimmung für den Pflege-Kontext kann mit einem Wert von  $r_{icc} = .360$  angegeben werden. Die ICC für den Essens-Kontext ist nicht interpretierbar.

Unbedingt hervorzuheben ist, dass die Intraklassenkorrelation stichproben- und varianzabhängig ist. Ein hoher ICC-Koeffizient kann nur bei einer großen Stichprobe mit einer großen Varianz erzielt werden. Das bedeutet, dass nur bei einer großen Anzahl an Vergleichswerten, der ICC-Koeffizient – bei großer Übereinstimmung zwischen den Ratern – auch hohe Werte erzielen kann. Liegen allerdings nur wenige Vergleichswerte vor, was bei sehr kurzen Videos mit einer geringer Anzahl an *Feinfühligkeits*-Sequenzen der Fall ist, so kann der Wert der ICC nie besonders hoch ausfallen, auch wenn eine nahezu perfekte Übereinstimmung zwischen den beiden BeobachterInnen vorliegt. Auch Zouzoula (Diplomarbeit in Arbeit) zeigte, dass durch das Zusammenfassen von mehreren Videos, was eine Erhöhung der Messwertpaare bedeutet, höhere ICC-Koeffizienten erzielt werden konnten als dies für die einzelnen Videos der Fall war. Dies hängt auch damit zusammen, dass dadurch die Varianz erhöht werden konnte. Die Varianz der Bewertungen für eine Mutter ist nie so groß, wie die Varianz von mehreren Müttern. Es handelt sich also um ein sehr sensibles Maß, das nur mit Vorsicht zu interpretieren ist. Geringe Intraklassenkorrelationen sind also nicht zwangsläufig ein Hinweis auf eine geringe Übereinstimmung zwischen Ratern.

Zur Gruppe der Pflege-Videos zählen auch Wickel-Situationen. Diese dauerten meistens nur fünf Minuten, wiesen daher nur zwei Messwertpaare auf. Wie gerade erläutert, liegt es in der Natur der Intraklassenkorrelation, dass bei wenigen Werten (also bei sehr kurzen Videos) der ICC-Koeffizient besonders gering ausfällt. Der geringe Wert des ICC-Koeffizienten der Pflege-Situationen ( $r_{icc} = .360$ ) könnte unter anderem dadurch erklärt

werden. Es ist anzudenken, Wickel-Szenarios aufgrund dieses Phänomens prinzipiell von der Anwendung des Beobachtungsinstruments INTAKT auszuschließen.

Es konnte beim Wickeln der Umstand, dass sich das Kind gar nicht am Prozess des Wickelns beteiligen kann, für die Bewertung als problematisch genannt werden. Das Manual fordert allerdings ‚das gemeinsame Arbeiten an einer Tätigkeit‘ und das ‚in den Hintergrund Stellen der eigenen Tätigkeiten der Mutter‘, um eine hohe Feinfühligkeit erzielen zu können. Dies ist jedoch beim Windel wechseln nur bedingt möglich. Das Handeln der Mutter ist – im Gegensatz zu einem Spiel – nicht primär an den Bedürfnissen des Kindes orientiert. Die Ankerpunkte sollten daher für Wickelvideos adaptiert werden. Stärker in den Fokus rücken könnten Kriterien, wie das kreative und abwechslungsreiche Vorgehen der Mutter, ob ein Gespräch während des Wickelns aufgebaut wird oder ob Spielmaterialien angeboten werden. Darüber hinaus ist in der Beurteilung stärker zu fokussieren, ob die Mutter Freude an der Aktivität zeigt, ob sie in der Lage ist, den kindlichen Blickwinkel einzunehmen, auf kindliche Initiativen liebevoll und würdigend eingeht und ob sie die Signale des Kindes wahrnimmt beziehungsweise darauf angemessen und prompt reagiert. Das Gegenstück dazu wäre eine Mutter, die stumm und rasch handelt, ohne mit dem Kind zu interagieren oder es miteinzubeziehen. Sie fragt vor ihrem Eingreifen auch nicht um Erlaubnis und reagiert nicht auf Signale des Kindes. Ihr Vorgehen ist eintönig und einfallslos. Besonders die Differenzierung der unteren Skalenpunkte fiel der AnwenderIn bei den Pflegevideos schwer.

Auch die Badesituationen erfordern ein konkretes Entscheidungsschema, das speziell auf diesen Kontext zugeschnitten ist. Kriterien, wie ‚die kindlichen Problemlösefähigkeiten zu fördern‘, stehen hier nicht unbedingt im Vordergrund. Zu definieren ist im Speziellen, ob eine Mutter, die ihrem Kind den Rücken wäscht, während das Kind mit Kübeln spielt, einer gemeinsamen Aktivität nachgeht oder nicht. Wenn nicht, dann ist das Kriterium des Manuals ‚das gemeinsame Arbeiten an einer Sache‘ zu entkräften, da es in dieser Situation kein Kriterium für eine feinfühlige Mutter darstellt.

Des Weiteren sollte explizit überdacht werden, ob die Ankerpunkte der *Feinfühligkeit* die jüngere Altersgruppe (1 - 3 Jahre) systematisch benachteiligen, indem eine lenkend vorgehende Mutter, abgestuft wird. Allerdings – wenn es um das Wohl des Kindes geht – gehen Mütter von jüngeren Kindern tendenziell lenkender vor, um ihre Kinder vor

Gefahren zu schützen und schneiden somit in der Bewertung der *Feinfühligkeit* gegenüber Müttern von älteren Kindern (3 - 6 Jahre) schlechter ab.

Für Interaktionen im Essenskontext ist zu sagen, dass es einen großen Unterschied zwischen Müttern, die ihre Kinder noch füttern müssen, und Müttern von älteren Kindern, die schon selbstständig essen können, gibt. Hier werden unterschiedliche Richtlinien und Beschreibungen der Ankerpunkte benötigt. Mütter von Kindern der jüngeren Altersgruppe, die noch gefüttert werden müssen, können ihre ‚eigenen Tätigkeiten nicht in den Hintergrund stellen‘. Hier benötigt es neue Standards und Kriterien.

### 12.2 Skala Rückmeldung

Die BeobachterInnenübereinstimmung für die Kategorie *Rückmeldung* konnte insgesamt mit einem  $\kappa = .36$  („Interact 9“ von Mangold, Toleranz von 50% und 3 Sekunden) sowie mit  $\kappa = .45$  („GSEQ 5.1“, Überlappungsbereich von 50% und 5 Events) als ausreichend beziehungsweise mittelmäßig identifiziert werden. Die Prozentuelle Übereinstimmung zwischen den Ratern liegt bei 64%, dieser Wert ist jedoch nicht zufallsbereinigt.

Die Videos wurden auch nach ihrer spezifischen Untersuchungsbedingung ausgewertet. Es konnte ein ausreichend großes Kappa mit  $\kappa = .31$  für die Essens- sowie Pflege-Situationen und ein mittelmäßiges Kappa von  $\kappa = .50$  für die Spielvideos errechnet werden. Der Anteil der prozentuellen Übereinstimmung zwischen den BeobachterInnen liegt zwischen 63 und 65% für die Essens- und Pflege-Videos, sowie bei 72% für die Spielvideos. Es ist ersichtlich, dass die Spielsituation – INTAKT wird normalerweise in einer standardisierten Spiel- und Bastelsituation angewendet – die höchste Reliabilität erzielt.

Die häufigste Unstimmigkeit bezog sich darauf, dass BeobachterIn A entweder *Negative Rückmeldung* oder *Korrigierende Rückmeldung* wertete, während sich BeurteilerIn B für *Keine Rückmeldung* entschied (siehe Kapitel 11.2.1). Der Tonfall der Rückmeldung spielt für die Bewertung eine entscheidende Rolle. Des Öfteren ist der Tonfall einer vermeintlich *negativen Rückmeldung* nicht eindeutig negativ und lässt somit einen gewissen Ermessens- und Interpretationsspielraum zu. Für die oder den eine(n) ist ein negativer Unterton herauszuhören, für die oder den andere(n) wiederum nicht. Dies erschwert natürlich das

Erzielen einer hohen Übereinstimmung. Hinzu kommen jene Aussagen, die eindeutig die Unzufriedenheit der Mutter ausdrücken, allerdings keinen eindeutig negativen Unterton vorweisen und auch nicht als *Korrigierende Rückmeldung* zu werten sind, da die Erklärung oder Hilfestellung ausbleibt. Solche Aussagen, werden nach Manual als *Keine Rückmeldung* klassifiziert. Der Rater steht somit vor dem Konflikt eine Aussage, die dem Kind Unzufriedenheit und Kritik übermittelt nicht zu registrieren, da sie zu wenig emotional negativ gefärbt ist.

### 12.3 Skala Joint Attention

Die Berechnungen zur Kategorie *Joint Attention* ergaben laut der Einteilung von Landis und Koch (1977) ein mittelmäßiges Kappa von  $\kappa = .48$  (A versus B) und ein ausreichend großes Kappa von  $\kappa = .40$  (B versus A). Die Prozentuelle Übereinstimmung zwischen den Ratern beträgt 63% (A versus B) und 56% (B versus A).

Für die spezifischen Untersuchungsbedingungen wurde ebenfalls die Reliabilität in Form der BeobachterInnenübereinstimmung ermittelt. Es konnte abermals die größte Übereinstimmung zwischen den Ratern für den Spielkontext gefunden werden. Nach Landis und Koch (1977) kann hier von einer beträchtlichen Übereinstimmung mit  $\kappa = .65$  (A versus B) und einer mittelmäßigen Übereinstimmung von  $\kappa = .53$  (B versus A) gesprochen werden. Die Reliabilität für die Pflege-Situationen kann als mittelmäßig ( $\kappa = .48$  und  $\kappa = .44$ ) angegeben werden. Die Essens-Situationen weisen die niedrigste Übereinstimmung mit  $\kappa = .38$  und  $\kappa = .32$  auf, was nach Landis und Koch (1977) als ausreichend zu bezeichnen ist. Die Prozentuelle Übereinstimmung für den Spielkontext betrug 76% beziehungsweise 67%. Interessanterweise sind die Werte für die Essens- und Pflege-Situation mit 59% (für den Vergleich A versus B) gleich hoch, obwohl es beim Kappa-Koeffizienten einen Unterschied von .10 gab.

Mehr als die Hälfte aller Nichtübereinstimmungen zwischen den beiden Ratern (4435 von 7626 Nichtübereinstimmungen, siehe Tabelle 28, Kapitel 11.3.1) entfallen auf die Problematik, dass BeobachterIn A *Aktive Aufrechterhaltung auf der Handlungsebene* kodiert, während sich BeobachterIn B in den meisten Fällen für *Passive Aufrechterhaltung* oder *Verbale Aufrechterhaltung* entscheidet, dicht gefolgt von der Verhaltensdimension

*Keine Joint Attention.* In Essensvideos taucht immer wieder die Frage auf, wie eine Sequenz zu bewerten ist, in der Mutter und Kind von ihren eigenen Tellern essen, ohne miteinander zu sprechen oder Blickkontakt zu halten. Ist es auf der *Handlungsebene* zu bewerten, da durch das gemeinsame Essen eine gemeinsame Aktivität ausgeübt wird, die einen gemeinsamen Aufmerksamkeitsfokus innehat, oder aber *Keine Joint Attention*, da es hier an Kommunikation fehlt und jeder für sich handelt. Ein anderes Beispiel bezieht sich auf die Frage, ob das Verhalten einer Mutter, die gerade kaut und ihrem Kind beim Essen zusieht, als *Passive Aufrechterhaltung* zu bewerten ist oder aber, da sie kaut und sich somit aktiv an der gemeinsamen Aktivität ‚Essen‘ beteiligt, auf der *Handlungsebene* zu werten ist. Hier verlangt es nach Kriterien und genauen Richtlinien sowie Definitionen, die auf die neuen Situationen zugeschnitten sind. Für Badevideos, die in den Pflegekontext eingeordnet werden können, ist oft unklar, ob das Verhalten einer Mutter, die ihrem Kind den Rücken wäscht, während dieses mit Kübeln und dem Badewasser spielt, als *Keine Joint Attention* oder *Aktive Aufrechterhaltung auf der Handlungsebene* zu kodieren ist. Ist die Beteiligung der Mutter an der Aktivität ‚Baden‘, auch wenn die beiden zwei voneinander getrennte Handlungen ausführen, ausreichend, um sie auf der *Handlungsebene* zu kodieren, da diese eine dem ‚Baden‘ untergeordnete Handlung darstellt. Auch für die Wickelsituation, die dem Pflegekontext angehört, ist nicht ganz klar, wie mit dem Verhalten einer Mutter umzugehen ist. Da sich das Kind nicht aktiv am Windel wechseln beteiligen kann, kann die Mutter auch keine passive Haltung einnehmen. Hier benötigt es klare Definitionen der verschiedenen Verhaltensdimensionen und klare Richtlinien für die Entscheidung vor allem für die Abgrenzung *Handlungsebene*, *Passive Aufrechterhaltung* und *Keine Joint Attention*.

Des Weiteren kann die Unstimmigkeit genannt werden, dass BeobachterIn A *Aufmerksamkeitslenkung* kodiert, während sich BeurteilerIn B für *Verbale Aufrechterhaltung* oder für die *Handlungsebene* entscheidet (750 von 7626 Nichtübereinstimmungen). Es geschieht auch umgekehrt, dass Rater B die Verhaltensdimension *Aufmerksamkeitslenkung* angibt, während Rater A die *Handlungsebene* oder *Verbale Ebene* auswählt (505 von 9202 Nichtübereinstimmungen, Tabelle 29). Bezüglich der Verhaltensdimension *Aufmerksamkeitslenkung* dürfte es besonders für die Anwendung auf die jüngere Altersgruppe und andere Situationen, Unklarheiten in der Entscheidungsfindung geben.

Es ist außerdem das Muster zu erkennen, dass in manchen Situationen, in denen BeobachterIn A *Keine Joint Attention* angibt, BeobachterIn B *Passive Aufrechterhaltung* kodiert (346 von 7626 Nichtübereinstimmungen, Tabelle 28). Auch hier gab es Schwierigkeiten in der Übertragung des Manuals auf andere Situationen und Untersuchungsbedingungen. Die Schwierigkeit bezog sich darauf abzugrenzen, ab wann das Verhalten einer Mutter passiv ist und ab wann kein gemeinsamer Aufmerksamkeitsfokus mehr vorliegt. Hierfür dürften die Essensvideos eine zentrale Rolle gespielt haben. Ist eine Sequenz, in der Mutter und Kind von ihren eigenen Tellern essen, gerade nicht miteinander sprechen und auch keinen Blickkontakt halten, da sie gerade mit dem Essen auf ihren Tellern beschäftigt sind, als *Keine Joint Attention* oder als *Passive Aufrechterhaltung* zu kodieren. Für die Verhaltenskategorie *Keine Joint Attention* spricht, dass sie keinen gemeinsamen Aufmerksamkeitsfokus aufweisen, da jeder für sich handelt, jegliche Kommunikation zum anderen fehlt, sogar der Blickkontakt. Für die Bewertung *Passive Aufrechterhaltung* spricht, dass die Mutter durch Kauen und ihre Beteiligung am Essensprozess zur Aufrechterhaltung der kindlichen Joint Attention beiträgt, allerdings ist das bloße Zusehen hier nicht erfüllt.

Insgesamt 605 Mal kodierte Rater B den *Aufmerksamkeitswechsel* nicht, den BeurteilerIn A kodiert hat, und 436 Mal bewertete Rater A den *Aufmerksamkeitswechsel*, den B angegeben hat, mit einer anderen Verhaltensdimension. Auch die Anwendung der Dimension *Aufmerksamkeitswechsel* auf neue Situationen (Essens- oder Pflegekontext) bereitete den BeobachterInnen Schwierigkeiten. Findet in Essenssituationen schon ein Wechsel statt, wenn die Mutter von einem Gang zum nächsten überleitet oder erst dann, wenn eine komplett neue Aktivität, die dem Essensprozess nicht untergeordnet ist, angeregt wird. Auch für Badevideos ist fraglich, ob der Übergang vom Spielen mit Kübel zum Waschen des Körpers einen Wechsel darstellt, oder erst wenn z.B. die Mutter das Baden beendet.

## 12.4 Conclusio

Das Beobachtungsinstrument INTAKT zur Erfassung der Mutter-Kind-Interaktion wurde im Zuge der hier vorliegenden Arbeit auf eine klinische Stichprobe in der Retrospektive in einem unstrukturierten Setting angewendet. Anhand einer jüngeren Altersgruppe und

neuen Situationen wurde INTAKT erprobt und die Reliabilität der einzelnen Dimensionen *Feinfühligkeit, Rückmeldung, Joint Attention* erörtert.

Die Schulung der BeobachterInnen wurde anhand von Spielsituationen vorgenommen, da INTAKT eigentlich auf eine Bastel- und Spielsituation in standardisierten Setting und einen Altersbereich von 3 bis 6 Jahren angewendet wird. Somit ist es nicht verwunderlich, dass sich in der Übertragung auf andere Situationen (Pflege- und Essenskontext), sowie eine jüngere Altersgruppe (1;2 - 3;5) Probleme ergeben und Unklarheiten auftun. Es verlangt daher in Zukunft nach der Überarbeitung des Manuals und der Ergänzung von Entscheidungsregeln für die neuen Situationen. Des Weiteren sollten die Rater auf die neuen Untersuchungssituationen und auch auf den jüngeren Altersbereich geschult werden. Nach Greve und Wentura (1997) können Beobachtungsfehler durch das gezielte Einsetzen von BeobachterInnenschulungen vermieden werden. Sie räumen dem Training einen besonderen Stellenwert ein. Rentzsch & Schütz (2009) nennen die BeobachterInnenschulung als eine Möglichkeit zur Erhöhung der Objektivität in nicht-standardisierten Prozessen.

Wenn die Klärung, der offenen Fragen, die im Zuge dieser Arbeit aufgekommen sind, erfolgt ist und das Manual um diese Antworten erweitert wurde, sollte die Reliabilität deutlich höher ausfallen. Es ist anzuraten ebenfalls zu bedenken, dass solange das Setting, in dem das Erhebungsinstrument angewendet wird, unstrukturiert ist, die Koeffizienten nie so hoch sein werden, wie die, die für INTAKT im standardisierten Setting gefunden werden konnten. Wirtz und Caspar (2002) geben die Voraussetzung der Standardisierung als eine der wichtigsten für eine reliable Beurteilung an.

## 13 Zusammenfassung

Diese Untersuchung stellt eine erste Erprobung des Beobachtungsinventars INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013), einem Instrument zur Erfassung elterlichen Erziehungsqualität, auf Situationen des klinischen Alltags dar. Hierfür wurden im Zuge der Kooperation mit der Säuglingspsychosomatik des Wilhelminenspitals in Wien Interaktionsvideos von auffälligen Mutter-Kind-Dyaden anhand von INTAKT analysiert. Eine Reliabilitätsanalyse klärt den Grad der Übereinstimmung der beiden Rater für die einzelnen Skalen (*Feinfühligkeit*, *Rückmeldung*, *Joint Attention*) des Instruments.

Die Schulung der BeobachterInnen erfolgte anhand von Spiel- und Bastelsituationen auf den Altersbereich der 3- bis 6-jährigen Kinder. Die Anwendung des Instruments erfolgte jedoch auf einen jüngeren Altersbereich (1;2 - 3;5) und auf Situationen mit Essens-, Pflege- und Spielkontext. Die Situationen der Studie fanden in einem nicht-standardisierten Setting statt, da es sich um eine retrospektive Datenanalyse handelt, die mit bereits in der Vergangenheit gedrehten Videos arbeitete.

Für die Kategorie *Feinfühligkeit* konnte eine Intraklassenkorrelation von  $r_{icc} = .361$  berechnet werden. In 56% der Fälle erzielten die beiden BeurteilerInnen eine exakte oder eine annähernde Übereinstimmung mit einer Differenz von einem Skalenpunkt. Der höchste Anteil an Übereinstimmung konnte mit 70% (Toleranz von einem Skalenpunkt) für die Spiel-Situationen gefunden werden kann. Für die Pflege- und Essens-Situationen schwankte dieser Wert allerdings um die 50%.

Die BeobachterInnenübereinstimmung für die Dimension *Rückmeldung* fiel mit einem Kappa-Koeffizienten von  $\kappa = .45$  mittelmäßig aus, die prozentuelle Übereinstimmung betrug 64%. Auch hier waren höhere Werte für die Spiel-Situationen ( $\kappa = .50$ ) im Gegensatz zu denen der Essens- und Pflege-Videos ( $\kappa = .31$ ) zu verzeichnen.

Für die Kategorie *Joint Attention* konnte je nach Reihenfolge des Vergleichs der Rater ein mittelmäßiges ( $\kappa = .48$ ) und ein ausreichend großes ( $\kappa = .40$ ) Kappa gefunden werden. Die Reliabilität für die Spielsituationen stellte sich mit  $\kappa = .65$  und  $\kappa = .53$  als beträchtlich und mittelmäßig heraus. Die Reliabilität für die Pflege-Situationen kann als mittelmäßig

( $\kappa = .48$  und  $\kappa = .44$ ) angegeben werden. Die Essens-Situationen wiesen die niedrigste Übereinstimmung mit  $\kappa = .38$  und  $\kappa = .32$  auf.

Da Instrument und Manual auf einen älteren Altersbereich und eine standardisierte Bastel- sowie Spielsituation ausgelegt sind, tauchten einige Schwierigkeiten und Unklarheiten in der Übertragung auf andere Situationen auf. Deshalb konnten in der Untersuchung nur ausreichend große bis mittelmäßige Reliabilitäten erzielt werden. Es ist daher notwendig das Manual zu überarbeiten und auf die neuen Situationen und die jüngere Altersgruppe auszulegen. Die BeurteilerInnen müssen zukünftig auf die neuen Situationen geschult werden, denn nur so kann eine reliable Beurteilung gewährleistet werden.

## 14 Abstract

This research study tests the applicability of INTAKT (Hirschmann, Aigner, Deimann & Kastner-Koller, 2013), an instrument for assessing the quality of mother-child interactions, to clinical situations. In the course of cooperation with 'Säuglingspsychosomatik' of Wilhelminenspital in Vienna, recorded interaction-sequences of mother-child dyads were provided. Two independent observers assessed those sequences with INTAKT and the inter-rater agreement for every single scale (*Maternal Sensitivity*, *Feedback*, *Joint Attention*) was calculated, in order to obtain the reliability for the instrument.

For the observer-training, videos with play- and handicraft-context were used. What has to be highlighted is that all children in the training videos were between 3 and 6 years old, in contrast to the participants of the study, who were between 14 and 41 months of age. The interactions of the study took place in different situations, which can be categorized in food-, care- and play-situations. Finally interactions between mother and child were not standardized, nor were there any instructions or materials used. This study is retrospective, because the mother-child interactions were already recorded.

An inter-rater reliability of  $r_{icc} = .361$  could be found for *Maternal Sensitivity*. 56% of the time the two raters achieved a perfect agreement or rather an agreement with a difference of one point on scale. The highest observer agreement turned out to be 70% (with a tolerance of one point on scale) for play-situations. The agreement for food- and care-situations laid around 50%

A value of  $\kappa = .45$  could be calculated for *Feedback*, which represents an average inter-rater reliability. The observer agreement turned out to be 64%. The highest score could be reached for play-context ( $\kappa = .50$ ), situations with food- and care-topics only achieved an inter-rater reliability of  $\kappa = .31$ .

According to the order of the comparison of the raters, moderate ( $\kappa = .48$ ) and fair ( $\kappa = .40$ ) Kappa-values were calculated for *Joint Attention*. The inter-rater reliability for play-context was identified as substantial and moderate ( $\kappa = .65$  and  $\kappa = .53$ ). For care-situations a moderate ( $\kappa = .48$  and  $\kappa = .44$ ) reliability could be found. The lowest agreement was identified with  $\kappa = .38$  and  $\kappa = .32$  for videos with food-context.

The described instrument was designed for an older age-group and a standardized play- and handicraft-situation. That's the reason why there were difficulties and uncertainties in the adaption to new situations and a younger age-group. Furthermore only sufficient or mediate inter-rater reliabilities could be identified. As a result it is necessary to rework and expand the manual. Prospectively the observers must be trained and instructed to these new situations and to the younger age-group, in order to achieve reliable evaluations.

## 15 Literatur

- Ainsworth, M.D.S. & Bell, S.M. (1970). Attachment, exploration and separation: Illustrated by the behavior of one-year olds in a strange situation. *Child Development*, 41, 49-67.
- Ainsworth, M.D.S., Bell, S.M. & Stayton, D.J. (1974). Infant-mother attachment and social development: "Socialization" as a product of reciprocal responsiveness to signals. In P.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, NJ: Erlbaum.
- Alessandri, S.M. & Lewis, M. (1996). Differences in Pride and Shame Maltreated and Nonmaltreated Preschoolers. *Child Development*, 67, 1857-1869.
- Bakeman, R., Quera, V. & Gnisci, A. (2009). Observer agreement for timed-event sequential data: A comparison of timed-based and event-based algorithms. *Behavior Research Methods*, 41 (1), 137-147.
- Bakeman, R. & Quera, V. (2011). *Sequential Analysis and Observational Methods for the Behavioral Sciences*. New York: Cambridge University Press.
- Bigelow, A.E., MacLean, K. & Proctor, J. (2004). The role of joint attention in the development of infants' play with objects. *Developmental Science*, 7, 518-526.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer Medizin Verlag.
- Bowlby, J. (2008). *Bindung als sichere Basis. Grundlagen und Anwendungen der Bindungstheorie*. München: Ernst Reinhardt.
- Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63, 1-174.

- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A. & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, 15, 481-498.
- Dallaire, D.H. & Weinraub, M. (2005). Predicting children's separation anxiety at age 6: The contributions of infant-mother attachment security, maternal sensitivity, and maternal separation anxiety. *Attachment & Human Development*, 7, 393-408.
- Faßnacht, G. (1979). *Systematische Verhaltensbeobachtung*. München: Ernst Reinhardt Verlag.
- Faßnacht, G. (1995). *Systematische Verhaltensbeobachtung*. München: Ernst Reinhardt Verlag.
- Greve, W. & Wentura, D. (1997). *Wissenschaftliche Beobachtung: Eine Einführung*. Weinheim: Psychologie Verlags Union.
- Grossmann, K. & Grossmann, K.E. (2008). *Bindungen. Das Gefüge psychischer Sicherheit* (4. durchgesehene Aufl.). Stuttgart: Klett-Cotta.
- Harter, S. (1978). Effectance motivation reconsidered: Toward a developmental model. *Human Development*, 21, 34-64.
- Häcker, H.O. & Stapf, K.-H. (2009). *Dorsch. Psychologisches Wörterbuch*. Bern: Verlag Hans Huber.
- Hänggi, Y., Benz-Fragnière, C., Haberkorn, K., Furler, K. & Perrez, M. (2013). Determinanten elterlicher Sensitivität. *Kindheit und Entwicklung*, 22 (1), 23-30.
- Hirschmann, N., Aigner, N., Deimann, P. & Kastner-Koller, U. (2013). *INTAKT. Ein Video-Beobachtungsinstrument zur Erfassung der Mutter-Kind-Interaktion – Manual*. Unveröffentlichtes Manuskript.
- 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*, Volume 53, 2011 (3), 295-311.

- Jäger, R.S. & Petermann, F. (1995). *Psychologische Diagnostik. Ein Lehrbuch*. Weinheim: Beltz, Psychologie Verlags Union.
- 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, 1061-1071.
- Kubinger, K.D. (2009). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Landis, J.R. & Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33, 159-174.
- Main, M. & Solomon, J. (1986). Discovery of an insecure-disorganized/disoriented attachment pattern. In T.B. Brazelton & M.W. Yogman (Eds.), *Affective development in infancy* (pp. 95-124). Norwood, NJ: Ablex
- Morales, M., Mundy, P., Delgado, C.E.F., Yale, M., Messinger, D., Neal, R. & Schwartz, H.K. (2000). Responding to Joint Attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology* 21(3), 283-298.
- Morales, M., Mundy, P., Crowson, M.M., Neal, A.R. & Delgado, C.E.F. (2005). Individual differences in infant attention skills, joint attention, and emotion regulation behavior. *International Journal of Behavioral Development*, 29 (3), 259-263.
- NICHD Early Child Care Research Network (2006). Infant-mother attachment: Risk and protection in relation to changing maternal caregiving quality over time. *Developmental Psychology*, 42, 38-58.
- Raikes, H.A., & Thompson, R.A. (2008). Attachment security and parenting quality predict children's problem-solving, attributions, and loneliness with peers. *Attachment & Human Development*, 10, 319-344.
- Rentzsch, K., & Schütz, A., (2009). *Psychologische Diagnostik. Grundlagen und Anwendungsperspektiven*. Stuttgart: Verlag W. Kohlhammer.

- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Verlag Hans Huber.
- Saxon, T.F., Colombo, J., Robinson, E.L. & Frick, J.E. (2000). Dyadic interaction profiles in infancy and preschool intelligence. *Journal of School Psychology, 38*, 9-25.
- Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment & Human Development, 7*, 349-367.
- Stams, G.-J. J. M., Juffer, F. & van Ijzendoorn, M. H. (2002). Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: The case of adopted children and their biologically unrelated parents. *Developmental Psychology, 38*, 806-821.
- Vaughan Van Hecke, A., Mundy, P.C., Acra, C.F., Block, J.J., Delgado, C.E.F., Parlade, M.V. et al. (2007). Infant joint attention, temperament, and social competence in preschool children. *Child Development, 78*, 53-69.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- [http://www.wienkav.at/kav/wil/medstellen\\_anzeigen.asp?ID=592](http://www.wienkav.at/kav/wil/medstellen_anzeigen.asp?ID=592) (17. Dezember 2013)

## 16 Abbildungsverzeichnis

|                                                                                                                             |    |
|-----------------------------------------------------------------------------------------------------------------------------|----|
| Abbildung 1: Wirkungsmodell der kindlichen Trennungsangst (Dallaire & Weinraub, 2005, S. 404)<br>.....                      | 7  |
| Abbildung 2: Darstellung der absoluten Häufigkeiten des Alters in Jahr(en);Monat(en) der<br>Gesamtstichprobe .....          | 52 |
| Abbildung 3: Darstellung der absoluten Häufigkeiten des Alters in Jahr(en);Monat(en) der<br>Gegenkodierungsstichprobe ..... | 54 |
| Abbildung 4: Kodierungsverlauf der beiden Rater über alle FFK-Sequenzen der Essens-Situationen<br>( $N = 93$ ) .....        | 62 |
| Abbildung 5: Kodierungsverlauf der beiden Rater über alle FFK-Sequenzen der Pflege-Situationen<br>( $N = 36$ ) .....        | 65 |
| Abbildung 6: Kodierungsverlauf der beiden Rater über alle FFK-Sequenzen der Spiel-Situationen<br>( $N = 40$ ) .....         | 69 |

## 17 Tabellenverzeichnis

|                                                                                                                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabelle 1: Kontingenztafel zur Verteilung der einzelnen Untersuchungssituationen und Erhebungszeitpunkte in der Gesamtstichprobe, Angaben in absoluten Häufigkeiten (N = 129 Videos) .....                                                                  | 52 |
| Tabelle 2: Kontingenztafel zur Verteilung der einzelnen Untersuchungssituationen und Erhebungszeitpunkte in der Gegenkodierungsstichprobe, Angaben in absoluten Häufigkeiten (N = 28 Videos) .....                                                          | 53 |
| Tabelle 3: Vergleich des Alters (in Monaten) zwischen der Gegenkodierungsstichprobe, jenen Probanden, die nicht zur Reliabilitätsbestimmung herangezogen wurden und der Gesamtstichprobe .....                                                              | 54 |
| Tabelle 4: Verteilung des ‚Alters‘ in Halbjahresschritten für die Gegenkodierungsstichprobe, für jene Probanden, die nicht für die Reliabilitätsanalyse herangezogen wurden und die Gesamtstichprobe, Angaben in absoluten und relativen Häufigkeiten ..... | 55 |
| Tabelle 5: absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Situationen (N = 169) .....                                                                                                                            | 56 |
| Tabelle 6: Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Situationen (N = 169) .....                                                                                                                                  | 57 |
| Tabelle 7: Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Situationen (N = 169) .....                                                                                                                                    | 58 |
| Tabelle 8: Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Situationen und Videos (N = 169) .....                                                                                                                            | 59 |
| Tabelle 9: absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Essens-Situationen (N = 93) .....                                                                                                                      | 60 |
| Tabelle 10: Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Essens-Situationen (N = 93) .....                                                                                                                           | 60 |
| Tabelle 11: Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Essens-Situationen (N = 93) .....                                                                                                                             | 61 |

## Tabellenverzeichnis

---

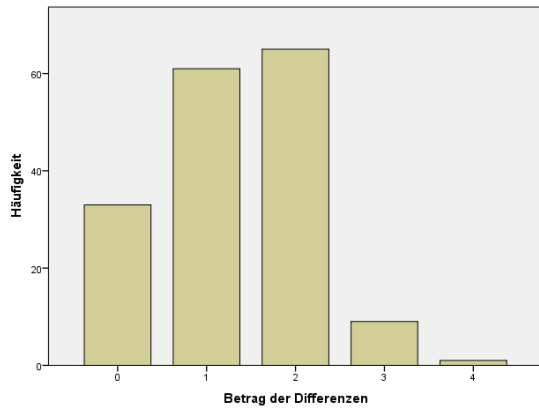
|                                                                                                                                                                                                  |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabelle 12: Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Essens-Situationen ( $N = 93$ ) .....                                                                 | 63 |
| Tabelle 13: absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Pflege-Situationen ( $N = 36$ ) .....                                                      | 64 |
| Tabelle 14: Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Pflege-Situationen ( $N = 36$ ) .....                                                            | 64 |
| Tabelle 15: Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Pflege-Situationen ( $N = 36$ ) .....                                                              | 65 |
| Tabelle 16: Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Pflege-Situationen ( $N = 36$ ) .....                                                                 | 66 |
| Tabelle 17: absolute und relative Häufigkeiten der Kodierungen auf der Skala Feinfühligkeit über alle Spiel-Situationen ( $N = 40$ ) .....                                                       | 68 |
| Tabelle 18: Mittelwert und Varianz der Beobachtungen auf der Dimension Feinfühligkeit über alle Spiel-Situationen ( $N = 40$ ) .....                                                             | 68 |
| Tabelle 19: Differenzen zwischen den beiden Ratern auf der Dimension Feinfühligkeit über alle Spiel-Situationen ( $N = 40$ ) .....                                                               | 69 |
| Tabelle 20: Intraklassenkorrelation der Feinfühligkeits-Bewertungen beider Rater über alle Spiel-Situationen ( $N = 40$ ) .....                                                                  | 70 |
| Tabelle 21: Kappa-Koeffizienten der Beobachtungen der beiden Rater über alle Situationen, in Abhängigkeit des Auswertungsprogramms und der Reihenfolge der BeobachterInnen .....                 | 72 |
| Tabelle 22: Kontingenztabelle zur Errechnung des event alignment der Kategorie Rückmeldung über alle Situationen, BeobachterIn A im Vergleich zu BeobachterIn B, erstellt mit GSEQ 5.1 .....     | 73 |
| Tabelle 23: Kontingenztabelle zur Errechnung des event alignment der Kategorie Rückmeldung über alle Situationen, im Vergleich ist BeobachterIn B zu BeobachterIn A, erstellt mit GSEQ 5.1 ..... | 73 |
| Tabelle 24: Kappa-Koeffizienten für die Kategorie Rückmeldung pro Situation in Abhängigkeit der Reihenfolge der Rater, erstellt mit Mangold INTERACT 9 .....                                     | 74 |

## Tabellenverzeichnis

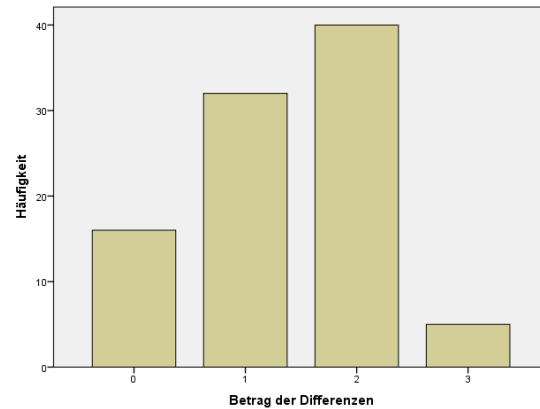
---

|                                                                                                                                                                                                   |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabelle 25: Prozentuelle Übereinstimmungen für die Kategorie Rückmeldung pro Situation in Abhängigkeit der Reihenfolge der Rater .....                                                            | 74 |
| Tabelle 26: Darstellung der Diskrepanzen zwischen den Ratern unter Angabe der Häufigkeit .....                                                                                                    | 76 |
| Tabelle 27: Kappa-Koeffizienten und Prozentuelle Übereinstimmungen für die Kategorie Joint Attention über alle Situationen in Abhängigkeit der Reihenfolge der Rater, erstellt mit GSEQ 5.1 ..... | 77 |
| Tabelle 28: Kontingenztabelle zur Errechnung von Kappa der Kategorie Joint Attention über alle Situationen, im Vergleich ist BeobachterIn A zu BeobachterIn B, erstellt mit GSEQ 5.1 .....        | 78 |
| Tabelle 29: Kontingenztabelle zur Errechnung von Kappa der Kategorie Joint Attention über alle Situationen, im Vergleich ist Rater B zu Rater A, erstellt mit GSEQ 5.1 .....                      | 79 |
| Tabelle 30: Kappa-Koeffizienten für die Kategorie Joint Attention pro Situation, in Abhängigkeit der Reihenfolge der Rater, erstellt mit GSEQ 5.1 .....                                           | 79 |
| Tabelle 31: Prozentuelle Übereinstimmung für die Kategorie Joint Attention pro Situation in Abhängigkeit der Rater .....                                                                          | 80 |
| Tabelle 32: Muster der Unterschiede in den Bewertungen der beiden Rater auf der Ebene der Joint Attention .....                                                                                   | 81 |

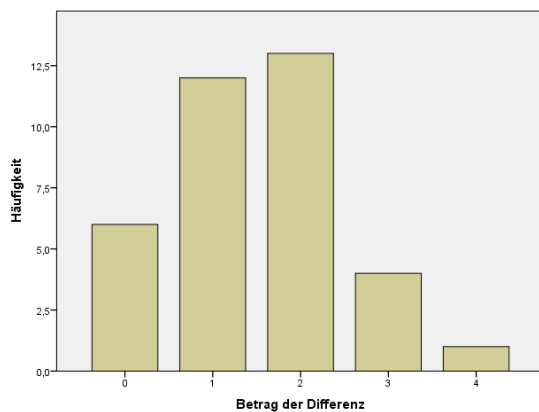
## 18 Anhang



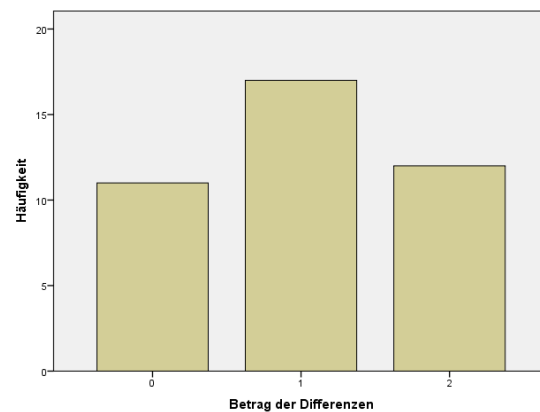
**Abbildung A-1a:** absolute Häufigkeiten der Differenzen zwischen den FFK-Bewertungen der beiden Rater für alle Situationen ( $N = 169$ )



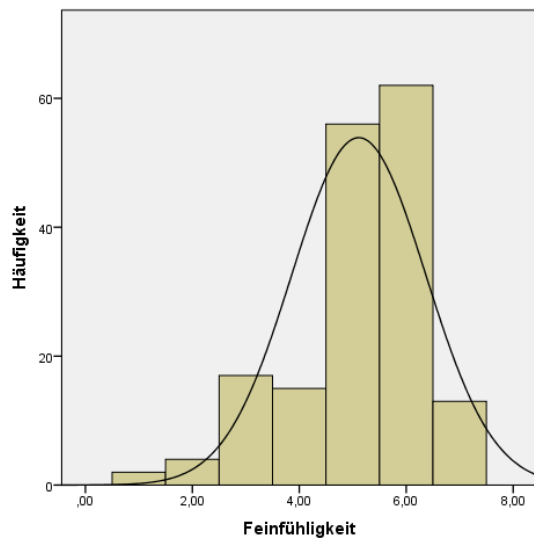
**Abbildung A-1b:** absolute Häufigkeiten der Differenzen zwischen den FFK-Bewertungen der beiden Rater für alle Essens-Situationen ( $N = 93$ )



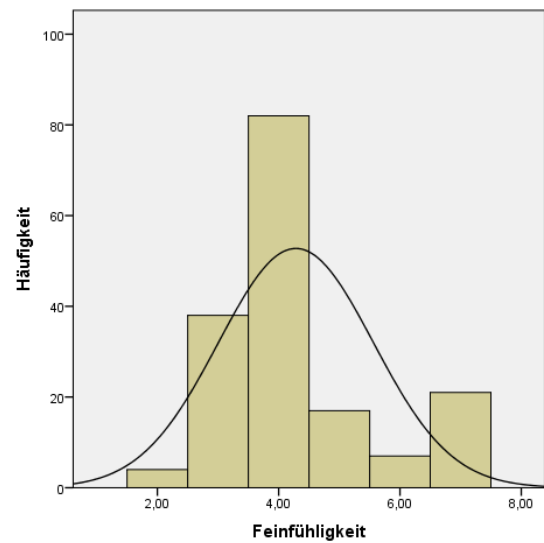
**Abbildung A-1c:** absolute Häufigkeiten der Differenzen zwischen den FFK-Bewertungen der beiden Rater für alle Pflege-Situationen ( $N = 36$ )



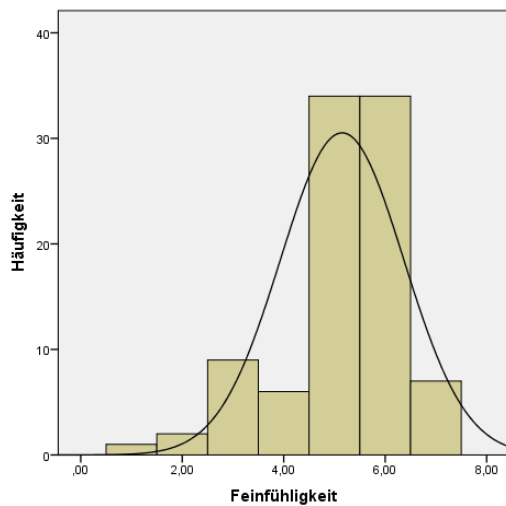
**Abbildung A-1d:** absolute Häufigkeiten der Differenzen zwischen den FFK-Bewertungen der beiden Rater für alle Spiel-Situationen ( $N = 40$ )



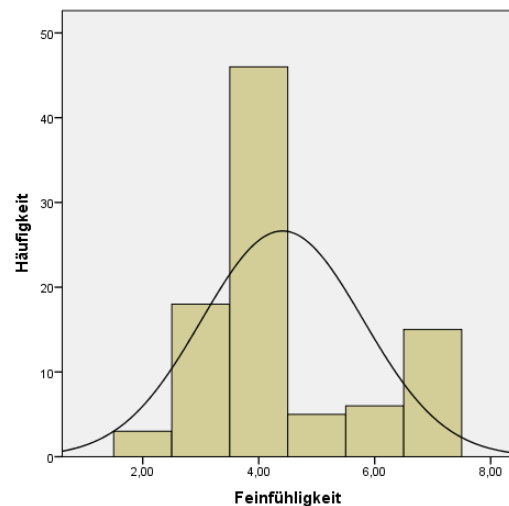
**Abbildung A-2a:** Verteilung der Bewertungen zur Feinfühligkeit von Rater A über alle Videos ( $N = 169$ ,  $M = 5.11$ ,  $S^2 = 1.57$ )



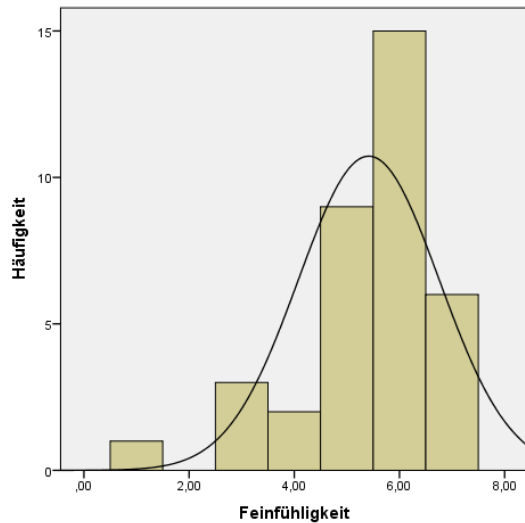
**Abbildung A-2b:** Verteilung der Bewertungen zur Feinfühligkeit von Rater B über alle Videos ( $N = 169$ ,  $M = 4.28$ ,  $S^2 = 1.63$ )



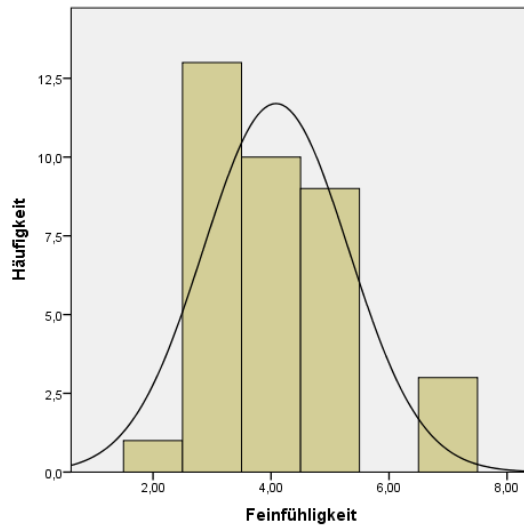
**Abbildung A-3a:** Verteilung der Bewertungen zur Feinfühligkeit von Rater A über alle Essens-Videos ( $N = 93$ ,  $M = 5.15$ ,  $S^2 = 1.48$ )



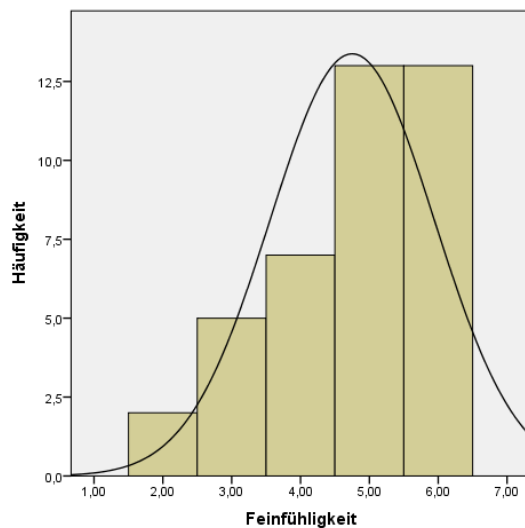
**Abbildung A-3b:** Verteilung der Bewertungen zur Feinfühligkeit von Rater B über alle Essens-Videos ( $N = 93$ ,  $M = 4.41$ ,  $S^2 = 1.94$ )



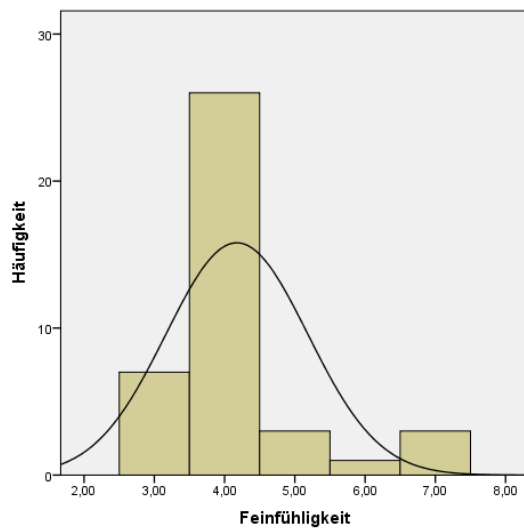
**Abbildung A-3c:** Verteilung der Bewertungen zur Feinfühligkeit von Rater A über alle Pflege-Videos ( $N = 36$ ,  $M = 5.42$ ,  $S^2 = 1.79$ )



**Abbildung A-3d:** Verteilung der Bewertungen zur Feinfühligkeit von Rater B über alle Pflege-Videos ( $N = 36$ ,  $M = 4.08$ ,  $S^2 = 1.51$ )



**Abbildung A-3e:** Verteilung der Bewertungen zur Feinfühligkeit von Rater A über alle Spiel-Videos ( $N = 40$ ,  $M = 4.75$ ,  $S^2 = 1.42$ )



**Abbildung A-3f:** Verteilung der Bewertungen zur Feinfühligkeit von Rater B über alle Spiel-Videos ( $N = 40$ ,  $M = 4.18$ ,  $S^2 = 1.02$ )

**Tabelle A-4:** Auflistung aller Videos

|    | <i>Kind</i> | <i>Videocode</i> | <i>Rater</i> | <i>Alter</i> | <i>Zeitpunkt</i> | <i>Situation</i> | <i>Videodauer</i> | <i>Gegenkodiert</i> |
|----|-------------|------------------|--------------|--------------|------------------|------------------|-------------------|---------------------|
| 1  | A.C.        | AG 36            | Gogl         | 1;3          | Prä              | Spiel            | 00:15:16          |                     |
| 2  | A.C.        | AG 37            | Gogl         | 1;3          | Prä              | Pflege           | 00:18:53          |                     |
| 3  | A.C.        | AG 35            | Gogl         | 1;3          | Post             | Essen            | 00:24:59          | x                   |
| 4  | A.C.        | HV 31            | Valai        | 1;3          | Prä              | Pflege           | 00:02:13          | x                   |
| 5  | A.C.        | HV 69            | Valai        | 1;3          | Prä              | Essen            | 00:14:08          |                     |
| 6  | A.C.        | HV 70            | Valai        | 1;3          | Post             | Pflege           | 00:02:44          |                     |
| 7  | A.F.        | AG 68            | Gogl         | 1;3          | Prä              | Spiel            | 00:12:08          | x                   |
| 8  | A.F.        | AG 65            | Gogl         | 1;3          | Post             | Pflege           | 00:14:48          | x                   |
| 9  | A.F.        | AG 67            | Gogl         | 1;3          | Prä              | Pflege           | 00:03:51          | x                   |
| 10 | A.F.        | AG 66            | Gogl         | 1;3          | Post             | Essen            | 00:28:21          |                     |
| 11 | A.F.        | HV 15            | Valai        | 1;3          | Prä              | Essen            | 00:12:58          | x                   |
| 12 | A.F.        | HV 30            | Valai        | 1;3          | Post             | Essen            | 00:01:32          |                     |
| 13 | A.F.        | HV 39            | Valai        | 1;3          | Post             | Spiel            | 00:13:59          | x                   |
| 14 | A.G.        | AG 49            | Gogl         | 2;2          | Post             | Spiel            | 00:26:46          |                     |
| 15 | A.G.        | AG 51            | Gogl         | 2;2          | Prä              | Spiel            | 00:12:36          |                     |
| 16 | A.G.        | AG 50            | Gogl         | 2;2          | Post             | Pflege           | 00:28:23          |                     |
| 17 | A.G.        | HV 27            | Valai        | 2;2          | Prä              | Essen            | 00:23:48          | x                   |
| 18 | A.G.        | HV 45            | Valai        | 2;2          | Post             | Pflege           | 00:22:08          |                     |
| 19 | A.G.        | HV 51            | Valai        | 2;2          | Prä              | Pflege           | 00:21:13          |                     |
| 20 | A.K.        | AG 26            | Gogl         | 2;3          | Prä              | Essen            | 00:23:35          |                     |
| 21 | A.K.        | HV 16            | Valai        | 2;3          | Post             | Pflege           | 00:08:02          |                     |
| 22 | A.K.        | HV 47            | Valai        | 2;3          | Prä              | Pflege           | 00:13:19          |                     |
| 23 | A.M.M.      | AG 15            | Gogl         | 4;4          | Prä              | Spiel            | 00:12:51          |                     |
| 24 | D.W.        | AG 16            | Gogl         | 1;6          | Post             | Spiel            | 00:12:25          |                     |
| 25 | D.W.        | AG 17            | Gogl         | 1;6          | Prä              | Spiel            | 00:10:01          |                     |
| 26 | D.W.        | AG 18            | Gogl         | 1;6          | Prä              | Essen            | 00:28:19          |                     |
| 27 | D.W.        | HV 05            | Valai        | 1;6          | Prä              | Pflege           | 00:10:32          |                     |
| 28 | D.W.        | HV 25            | Valai        | 1;6          | Post             | Pflege           | 00:10:40          |                     |
| 29 | D.W.        | HV 43            | Valai        | 1;6          | Prä              | Essen            | 00:04:06          |                     |
| 30 | E.V.        | AG 43            | Gogl         | 1;5          | Prä              | Pflege           | 00:07:25          |                     |
| 31 | E.V.        | AG 44            | Gogl         | 1;5          | Prä              | Essen            | 00:11:24          | x                   |
| 32 | E.V.        | AG 45            | Gogl         | 1;5          | Prä              | Essen            | 00:18:10          |                     |
| 33 | E.V.        | HV 18            | Valai        | 1;5          | Post             | Essen            | 00:26:43          |                     |
| 34 | E.V.        | HV 41            | Valai        | 1;5          | Prä              | Essen            | 00:28:22          | x                   |
| 35 | E.V.        | HV 68            | Valai        | 1;5          | Prä              | Spiel            | 00:16:19          |                     |
| 36 | Ed.W.       | AG 52            | Gogl         | 1;2          | Post             | Pflege           | 00:04:21          | x                   |
| 37 | Ed.W.       | AG 53            | Gogl         | 1;2          | Prä              | Essen            | 00:13:29          |                     |
| 38 | Ed.W.       | HV 01            | Valai        | 1;2          | Prä              | Pflege           | 00:12:26          |                     |
| 39 | Ed.W.       | HV 56            | Valai        | 1;2          | Prä              | Spiel            | 00:10:58          |                     |
| 40 | F.J.        | AG 28            | Gogl         | 3;5          | Prä              | Spiel            | 00:10:29          |                     |
| 41 | F.J.        | AG 27            | Gogl         | 3;5          | Post             | Essen            | 00:36:40          |                     |
| 42 | F.J.        | HV 10            | Valai        | 3;5          | Prä              | Essen            | 00:25:14          | x                   |
| 43 | F.J.        | HV 28            | Valai        | 3;5          | Prä              | Pflege           | 00:04:46          |                     |

## Anhang

|    |      |       |       |      |      |        |          |   |
|----|------|-------|-------|------|------|--------|----------|---|
| 44 | H.A. | AG 11 | Gogl  | 1;9  | Prä  | Pflege | 00:05:50 | x |
| 45 | H.A. | AG 13 | Gogl  | 1;9  | Prä  | Pflege | 00:18:28 |   |
| 46 | H.A. | AG 10 | Gogl  | 1;9  | Post | Essen  | 00:15:35 |   |
| 47 | H.A. | AG 12 | Gogl  | 1;9  | Prä  | Essen  | 00:18:21 |   |
| 48 | H.A. | HV 11 | Valai | 1;9  | Prä  | Spiel  | 00:12:04 |   |
| 49 | H.A. | HV 19 | Valai | 1;9  | Post | Pflege | 00:15:17 | x |
| 50 | H.A. | HV 54 | Valai | 1;9  | Prä  | Pflege | 00:02:43 |   |
| 51 | J.L. | AG 02 | Gogl  | 2;10 | Prä  | Spiel  | 00:12:40 | x |
| 52 | J.L. | AG 01 | Gogl  | 2;10 | Post | Essen  | 00:27:50 |   |
| 53 | J.L. | AG 03 | Gogl  | 2;10 | Prä  | Essen  | 00:28:00 |   |
| 54 | J.L. | HV 08 | Valai | 2;10 | Prä  | Spiel  | 00:05:54 |   |
| 55 | J.L. | HV 20 | Valai | 2;10 | Prä  | Essen  | 00:28:04 |   |
| 56 | J.L. | HV 29 | Valai | 2;10 | Prä  | Essen  | 00:02:38 | x |
| 57 | K.D. | AG 09 | Gogl  | 1;4  | Post | Spiel  | 00:11:02 | x |
| 58 | K.D. | HV 60 | Valai | 1;4  | Prä  | Spiel  | 00:10:24 |   |
| 59 | K.S. | AG 46 | Gogl  | 1;6  | Post | Essen  | 00:15:46 |   |
| 60 | K.S. | AG 47 | Gogl  | 1;6  | Prä  | Essen  | 00:12:11 | x |
| 61 | K.S. | AG 48 | Gogl  | 1;6  | Prä  | Essen  | 00:04:23 |   |
| 62 | K.S. | HV 07 | Valai | 1;6  | Prä  | Essen  | 00:06:51 |   |
| 63 | K.S. | HV 34 | Valai | 1;6  | Prä  | Spiel  | 00:13:54 |   |
| 64 | K.S. | HV 57 | Valai | 1;6  | Post | Spiel  | 00:15:07 | x |
| 65 | K.S. | HV 63 | Valai | 1;6  | Prä  | Essen  | 00:10:35 |   |
| 66 | L.A. | AG 42 | Gogl  | 3;0  | Prä  | Pflege | 00:17:59 |   |
| 67 | L.A. | AG 41 | Gogl  | 3;0  | Post | Essen  | 00:21:45 |   |
| 68 | L.A. | HV 22 | Valai | 3;0  | Post | Spiel  | 00:21:31 |   |
| 69 | L.A. | HV 52 | Valai | 3;0  | Prä  | Spiel  | 00:10:40 |   |
| 70 | L.H. | AG 30 | Gogl  | 1;3  | Prä  | Spiel  | 00:10:02 |   |
| 71 | L.H. | AG 29 | Gogl  | 1;3  | Post | Essen  | 00:34:55 |   |
| 72 | L.H. | HV 37 | Valai | 1;3  | Prä  | Essen  | 00:20:20 |   |
| 73 | L.H. | HV 53 | Valai | 1;3  | Prä  | Spiel  | 00:12:14 |   |
| 74 | L.O. | AG 21 | Gogl  | 2;0  | Post | Spiel  | 00:13:06 |   |
| 75 | L.O. | AG 22 | Gogl  | 2;0  | Prä  | Pflege | 00:02:14 |   |
| 76 | L.O. | AG 23 | Gogl  | 2;0  | Prä  | Essen  | 00:16:47 | x |
| 77 | L.O. | HV 03 | Valai | 2;0  | Post | Pflege | 00:05:42 | x |
| 78 | L.O. | HV 40 | Valai | 2;0  | Post | Essen  | 00:19:04 |   |
| 79 | L.O. | HV 59 | Valai | 2;0  | Prä  | Spiel  | 00:13:02 |   |
| 80 | M.A. | AG 07 | Gogl  | 1;2  | Post | Essen  | 00:34:18 |   |
| 81 | M.A. | AG 08 | Gogl  | 1;2  | Prä  | Essen  | 00:15:41 | x |
| 82 | M.A. | HV 35 | Valai | 1;2  | Prä  | Spiel  | 00:12:17 |   |
| 83 | M.A. | HV 66 | Valai | 1;2  | Prä  | Essen  | 00:19:21 |   |
| 84 | M.B. | AG 38 | Gogl  | 2;10 | Post | Spiel  | 00:57:22 |   |
| 85 | M.B. | AG 40 | Gogl  | 2;10 | Prä  | Spiel  | 00:09:59 | x |
| 86 | M.B. | AG 39 | Gogl  | 2;10 | Post | Essen  | 00:15:40 |   |
| 87 | M.B. | HV 02 | Valai | 2;10 | Post | Spiel  | 00:03:24 |   |
| 88 | M.B. | HV 46 | Valai | 2;10 | Post | Essen  | 00:46:47 |   |

## Anhang

---

|     |         |           |       |      |      |        |          |   |
|-----|---------|-----------|-------|------|------|--------|----------|---|
| 89  | M.B.    | HV 64     | Valai | 2;10 | Prä  | Pflege | 00:06:01 |   |
| 90  | M.D.    | AG69/HV58 | Valai | 2;5  | Prä  | Spiel  | 00:12:44 | x |
| 91  | M.D.    | HV06/AG34 | Gogl  | 2;5  | Prä  | Pflege | 00:38:49 | x |
| 92  | M.D.    | AG32      | Gogl  | 2;5  | Post | Essen  | 00:07:54 | x |
| 93  | M.D.    | AG33/HV12 | Gogl  | 2;5  | Prä  | Essen  | 00:03:28 | x |
| 94  | M.D.    | HV14/AG70 | Valai | 2;5  | Prä  | Essen  | 00:28:26 | x |
| 95  | N.M.    | HV 24     | Valai | 1;4  | Post | Essen  | 00:33:08 |   |
| 96  | S.H.    | AG 31     | Gogl  | 1;9  | Post | Essen  | 00:20:42 |   |
| 97  | S.H.    | HV 23     | Valai | 1;9  | Prä  | Essen  | 00:43:11 |   |
| 98  | S.Q.    | AG 59     | Gogl  | 1;1  | Prä  | Pflege | 00:04:59 |   |
| 99  | S.Q.    | AG 60     | Gogl  | 1;1  | Prä  | Essen  | 00:20:51 |   |
| 100 | S.Q.    | AG 61     | Valai | 1;1  | Prä  | Essen  | 00:28:26 |   |
| 101 | S.Q.    | HV 17     | Valai | 1;1  | Post | Essen  | 00:18:18 |   |
| 102 | S.Q.    | HV 36     | Valai | 1;1  | Prä  | Essen  | 00:26:39 |   |
| 103 | S.Q.    | HV 50     | Valai | 1;1  | Prä  | Pflege | 00:15:31 |   |
| 104 | S.Q.    | HV 55     | Valai | 1;1  | Prä  | Spiel  | 00:19:13 |   |
| 105 | S.T.    | AG 55     | Gogl  | 1;2  | Prä  | Spiel  | 00:10:11 |   |
| 106 | S.T.    | AG 54     | Gogl  | 1;2  | Post | Essen  | 00:36:40 |   |
| 107 | S.T.    | HV 21     | Valai | 1;2  | Prä  | Pflege | 00:23:56 |   |
| 108 | S.T.    | HV 67     | Valai | 1;2  | Prä  | Essen  | 00:19:31 |   |
| 109 | Sa.Li.  | AG 05     | Gogl  | 1;7  | Post | Spiel  | 00:18:18 |   |
| 110 | Sa.Li.  | AG 04     | Gogl  | 1;7  | Post | Essen  | 00:11:18 |   |
| 111 | Sa.Li.  | AG 06     | Gogl  | 1;7  | Prä  | Essen  | 00:19:21 |   |
| 112 | Sa.Li.  | HV 04     | Valai | 1;7  | Prä  | Spiel  | 00:13:02 |   |
| 113 | Sa.Li.  | HV 42     | Valai | 1;7  | Prä  | Pflege | 00:06:01 |   |
| 114 | Sa.Li.  | HV 49     | Valai | 1;7  | Post | Spiel  | 00:01:05 |   |
| 115 | Sa.Sch. | AG 57     | Gogl  | 1;10 | Prä  | Pflege | 00:02:01 |   |
| 116 | Sa.Sch. | AG 56     | Gogl  | 1;10 | Post | Essen  | 00:12:20 |   |
| 117 | Sa.Sch. | AG 58     | Gogl  | 1;10 | Prä  | Essen  | 00:11:05 |   |
| 118 | Sa.Sch. | HV 44     | Valai | 1;10 | Prä  | Spiel  | 00:10:39 |   |
| 119 | Sa.Sch. | HV 62     | Valai | 1;10 | Post | Spiel  | 00:08:02 |   |
| 120 | Sa.St.  | AG 19     | Gogl  | 1;9  | Post | Pflege | 00:17:45 |   |
| 121 | Sa.St.  | AG 20     | Gogl  | 1;9  | Prä  | Pflege | 00:27:13 |   |
| 122 | Sa.St.  | HV 13     | Valai | 1;9  | Prä  | Essen  | 00:17:17 |   |
| 123 | Sa.St.  | HV 32     | Valai | 1;9  | Post | Spiel  | 00:19:01 |   |
| 124 | Si.Li.  | AG 62     | Gogl  | 1;6  | Post | Spiel  | 00:13:21 |   |
| 125 | Si.Li.  | AG 63     | Gogl  | 1;6  | Post | Pflege | 00:05:03 |   |
| 126 | Si.Li.  | AG 64     | Gogl  | 1;6  | Prä  | Essen  | 00:21:07 |   |
| 127 | Si.Li.  | HV 09     | Valai | 1;6  | Post | Essen  | 00:22:24 |   |
| 128 | Si.Li.  | HV 33     | Valai | 1;6  | Prä  | Pflege | 00:06:24 |   |
| 129 | Si.Li.  | HV 61     | Valai | 1;6  | Prä  | Spiel  | 00:12:20 |   |

---

## Anhang

**Tabelle A-5:** Verteilung der Variablen ‚Situation‘ und ‚Zeitpunkt‘ in der Gegenkodierungsstichprobe, bei jenen Probanden, die nicht zur Analyse herangezogen wurden und der Gesamtstichprobe

|               | Gegenkodierungsstichprobe | Nicht gewählte Probanden | Gesamtstichprobe |
|---------------|---------------------------|--------------------------|------------------|
| Prä x Essen   | 11                        | 23                       | 34               |
| Prä x Pflege  | 4                         | 19                       | 23               |
| Prä x Spiel   | 4                         | 21                       | 25               |
| Post x Essen  | 2                         | 20                       | 22               |
| Post x Pflege | 4                         | 7                        | 11               |
| Post x Spiel  | 3                         | 11                       | 14               |
| $\Sigma$      | <b>28</b>                 | <b>101</b>               | <b>129</b>       |

**Tabelle A-6a:** Verteilung des ‚Alters‘ in Halbjahresschritten auf die verschiedenen Untersuchungssituationen und Erhebungszeitpunkte für die nicht herangezogenen Probanden der Gesamtstichprobe

|          | Essen x Prä | Spiel x Prä | Pflege x Prä | Essen x Post | Spiel x Post | Pflege x Post | $\Sigma$   |
|----------|-------------|-------------|--------------|--------------|--------------|---------------|------------|
| 1;1-1;6  | 15          | 12          | 8            | 10           | 2            | 3             | 50         |
| 1;7-2;0  | 5           | 4           | 6            | 5            | 5            | 1             | 26         |
| 2;1-2;6  | 1           | 1           | 2            | 0            | 1            | 3             | 8          |
| 2;7-3;0  | 2           | 2           | 2            | 4            | 3            | 0             | 13         |
| 3;1-3;6  | 0           | 1           | 1            | 1            | 0            | 0             | 3          |
| 3;7-4;0  | 0           | 0           | 0            | 0            | 0            | 0             | 0          |
| 4;1-4;6  | 0           | 1           | 0            | 0            | 0            | 0             | 1          |
| $\Sigma$ | 23          | 21          | 19           | 20           | 11           | 7             | <b>101</b> |

**Tabelle A-6b:** Verteilung des ‚Alters‘ in Halbjahresschritten auf die verschiedenen Untersuchungssituationen und Erhebungszeitpunkte in der Gegenkodierungsstichprobe

|          | Essen x Prä | Spiel x Prä | Pflege x Prä | Essen x Post | Spiel x Post | Pflege x Post | $\Sigma$  |
|----------|-------------|-------------|--------------|--------------|--------------|---------------|-----------|
| 1;1-1;6  | 5           | 1           | 2            | 1            | 3            | 2             | 14        |
| 1;7-2;0  | 1           | 0           | 1            | 0            | 0            | 2             | 4         |
| 2;1-2;6  | 3           | 1           | 1            | 1            | 0            | 0             | 6         |
| 2;7-3;0  | 1           | 2           | 0            | 0            | 0            | 0             | 3         |
| 3;1-3;6  | 1           | 0           | 0            | 0            | 0            | 0             | 1         |
| 3;7-4;0  | 0           | 0           | 0            | 0            | 0            | 0             | 0         |
| 4;1-4;6  | 0           | 0           | 0            | 0            | 0            | 0             | 0         |
| $\Sigma$ | 11          | 4           | 4            | 2            | 3            | 4             | <b>28</b> |

## Lebenslauf

### Persönliche Daten

Name: Alexandra Gogl

Geburtsdatum/-ort: 20.8.1987, Wien

Staatsbürgerschaft: Österreich

Kontakt: [alexandra.gogl@gmx.at](mailto:alexandra.gogl@gmx.at)

### Universitäre Ausbildung

seit 10/2005 Diplomstudium Psychologie, Universität Wien

*Schwerpunkte:* Angewandte Kinder- und Jugendpsychologie  
Klinische und Gesundheitspsychologie

### Schulbildung

09/1997 - 06/2005 Akademisches Gymnasium Wien

09/1993 - 06/1997 Volksschule Stubenbastei Wien

### Studienbezogene Tätigkeiten

09/2013 - 11/2013 Volontariat an der Psychologischen Fakultät Wien, Psychologische Diagnostik, unter der Leitung von Ass.-Prof. Dr. Pia Deimann und Ass.-Prof. Dr. Ursula Kastner-Koller, Beurteilen von Interaktionsvideos zur Überprüfung der Beobachterübereinstimmung

09/2012 - 10/2012 Praktikum am Landesklinikum Mödling, Abteilung für Kinder- und Jugendheilkunde auf der Psychosomatischen Station

### Sprachen

Englisch: Verhandlungssicher

Italienisch: Grundkenntnisse

Französisch: Grundkenntnisse

Latein: Großes Latinum