



Deutsches
Jugendinstitut

Kurzbericht

Noemi Eberlein, Regine Schelle

Methodische Herausforderungen bei der Bewertung der Kita-Qualität

Ergebnisse einer qualitativen Studie

Impressum

© 2019 Deutsches Jugendinstitut

Abt. Kinder und Kinderbetreuung

Nockherstraße 2

81541 München

Telefon +49 89 62306-0

Fax +49 89 62306-162

E-Mail info@dji.de

www.dji.de

ISBN 978-3-86379-310-4

Forschung zu Kindern, Jugendlichen und Familien an der Schnittstelle von Wissenschaft, Politik und Fachpraxis

Das Deutsche Jugendinstitut e.V. (DJI) ist eines der größten sozialwissenschaftlichen Forschungsinstitute Europas. Seit über 50 Jahren erforscht es die Lebenslagen von Kindern, Jugendlichen und Familien, berät Bund, Länder und Gemeinden und liefert wichtige Impulse für die Fachpraxis. Träger des 1963 gegründeten Instituts ist ein gemeinnütziger Verein mit Mitgliedern aus Politik, Wissenschaft, Verbänden und Einrichtungen der Kinder-, Jugend- und Familienhilfe. Die Finanzierung erfolgt überwiegend aus Mitteln des Bundesministeriums für Familie, Senioren, Frauen und Jugend und den Bundesländern. Weitere Zuwendungen erhält das DJI im Rahmen von Projektförderungen vom Bundesministerium für Bildung und Forschung, der Europäischen Kommission, Stiftungen und anderen Institutionen der Wissenschaftsförderung.

Aktuell arbeiten und forschen 360 Mitarbeiter/innen (davon 225 Wissenschaftler/innen) an den beiden Standorten München und Halle (Saale). Die qualitative Beobachtungsstudie wurde im Rahmen des Projekts „Methodenstudie: Qualität in der Kindertageseinrichtung“ in der Fachgruppe „Pädagogische Konzepte für die Kindheit“ der Abteilung Kinder und Kinderbetreuung durchgeführt.

Inhalt

	Vorwort	5
1	Das Projekt „Methodenstudie: Qualität in der Kindertageseinrichtung“	6
2	Die qualitative Studie	8
	2.1 Bausteine der Datenerhebung	8
	2.2 Auswertungsstrategie	12
3	Ergebnisse der qualitativen Studie	19
4	Rückschlüsse für die Weiterentwicklung	30
	4.1 Empfehlungen zur Weiterentwicklung	30
	4.2 Ausblick	32
	Literatur	33

Vorwort

In der Fachabteilung „Kinder und Kinderbetreuung“ des Deutschen Jugendinstituts e.V. wurde zwischen Dezember 2016 bis Dezember 2018 die „Methodenstudie: Qualität in der Kindertageseinrichtung (MS-Kita)“ durchgeführt, die vom Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) gefördert wurde. Die Studie bestand aus zwei Projektbausteinen: der erste Baustein befasste sich mit pädagogischer Qualität aus Eltern- und Fachkräftesicht, der zweite Baustein mit der Erhebung pädagogischer Qualität unter Anwendung standardisierter Beobachtungsinstrumente für den frühpädagogischen Bereich. In der vorliegenden Publikation werden die zentralen Projektergebnisse des zweiten Bausteins der Methodenstudie, einer qualitativen Studie, dargestellt.

Ziel war es herauszufinden, wie Qualität in Kitas mit wissenschaftlichen Methoden erhoben wird und welche methodischen Herausforderungen sich zeigen, wenn Qualität unter Anwendung verschiedener standardisierter Instrumente bewertet wird. Dazu wurde eine Dokumentenanalyse einiger Instrumente vorgenommen, teilnehmende Beobachtungen in Kindergartengruppen durchgeführt sowie Filmsequenzen von Fachkraft-Kind-Interaktionen aufgenommen, die anschließend von Expertinnen und Experten für verschiedene standardisierte Instrumente zur Qualitätsbewertung ausgewertet wurden.

Die Studienergebnisse zeigen, dass mit der Bewertung pädagogischer Qualität unterschiedliche Schwierigkeiten verbunden sind, die sich sowohl auf methodische, als auch inhaltliche Aspekte der standardisierten Beobachtungsinstrumente beziehen. So zeigt sich, dass für die Interpretation eines Bewertungsergebnisses das zugrundeliegende Qualitätsverständnis des jeweiligen Instruments entscheidend ist. Insbesondere methodische Vorgaben und Anforderungen stellen zentrale Herausforderungen bei der Bewertung der pädagogischen Qualität dar. Mit diesen sind wiederum Weiterentwicklungsbedarfe verbunden. Zum einen ist es wichtig, dass bei der Bewertung des pädagogischen Handelns in der Kita inhaltliche Schwerpunkte und methodische Herausforderungen transparenter werden. Zum anderen sollte das kindliche Wohlbefinden Bestandteil einer Qualitätsbewertung sein.

Danksagung

An der Realisierung der qualitativen Studie waren verschiedene Personen und Organisationen beteiligt. Wir danken zunächst dem Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) für die finanzielle Förderung. Auch möchten wir uns bei den Akteurinnen und Akteuren bedanken, die uns beim Feldzugang unterstützt haben. Darüber hinaus bedanken wir uns für das große Vertrauen, das uns die pädagogischen Fachkräfte und Eltern entgegengebracht und so einen wertvollen Einblick in die Praxis ermöglicht haben. Unser größter Dank gilt den Kindern, die uns an ihrem Alltag in der Kita teilhaben ließen. Ohne ihre Offenheit wäre die Durchführung unserer Studie nicht möglich gewesen.

1.

Das Projekt „Methodenstudie: Qualität in der Kindertages- einrichtung“

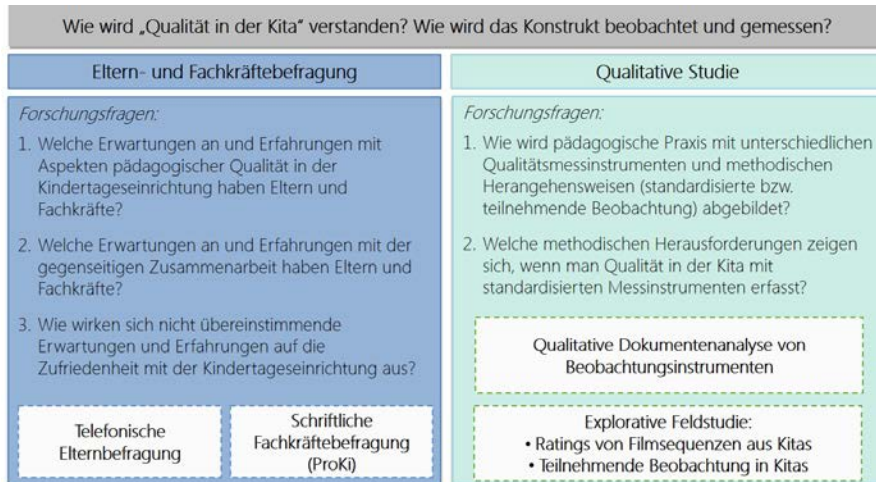
Mit dem Ausbau der Kinderbetreuungsinfrastruktur sind gestiegene Anforderungen an die pädagogische Qualität in den Einrichtungen verbunden. In der Wissenschaft wie auch in Praxis und Politik wird allerdings kontrovers diskutiert, wie Qualität in frühpädagogischen Einrichtungen zu definieren, zu messen und weiterzuentwickeln ist: Schwierigkeiten zeigen sich vor allem dann, wenn die Qualität solcher pädagogischen Konzepte erfasst werden soll, die individuell auf die Bedürfnisse jedes Kindes eingehen und Bildungsprozesse im Alltag der Kinder unterstützen wollen.

Die „Methodenstudie: Qualität in der Kindertageseinrichtung (MS-Kita)“¹ zielte daher auf eine theoretische Präzisierung des Konzepts der pädagogischen Qualität in der frühkindlichen Bildung, Betreuung und Erziehung. Zu diesem Zweck wurde einerseits die Eltern- und Fachkräfteperspektive auf pädagogische Qualität sowie deren Kooperation analysiert. Andererseits wurden in einem explorativen Vorgehen methodische Herausforderungen bei der standardisierten Bewertung als auch die Bestandteile pädagogischer Qualität in unterschiedlichen Instrumenten identifiziert.

Das Gesamtprojekt MS-Kita bestand aus zwei Modulen. Während im quantitativ ausgerichteten Modul eine telefonische und schriftliche Befragung von Eltern und pädagogischen Fachkräften durchgeführt wurde, stand im Zentrum des qualitativ ausgerichteten Moduls eine Dokumentenanalyse sowie eine explorative Feldstudie in Kitas (Abb. 1). In diesem Bericht wird das qualitative Modul behandelt. Der analog aufgebaute Bericht zum quantitativen Modul findet sich bei Michl und Geier (2019).

¹ Die Studie wurde am DJI in der Fachabteilung „Kinder und Kinderbetreuung“ von Dezember 2016 bis Dezember 2018 durchgeführt und vom Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) gefördert.

Abbildung 1: Das Projekt MS-Kita und seine zwei Module



2.

Die qualitative Studie

Längsschnittstudien konnten zeigen, dass der Besuch einer vorschulischen Einrichtung kurz- und mittelfristig positive Effekte auf die kognitive und sozial-emotionale Entwicklung der betreuten Kinder haben kann – allerdings nur, wenn das Betreuungsangebot von hoher Qualität ist (im Überblick: Anders/Roßbach 2014). Diese Ergebnisse verdeutlichen die Notwendigkeit, die Qualität von Kindertageseinrichtungen sicherzustellen und vor allem weiterzuentwickeln. Insbesondere mit dem Ausbau der Betreuungsinfrastruktur, dem Rechtsanspruch auf Kindertagesbetreuung für Kinder ab dem Alter von einem Jahr und der damit stark zunehmenden Teilhabe der Ein- bis Dreijährigen an frühkindlicher Bildung, Betreuung und Erziehung (FBBE) wächst auch die öffentliche Verantwortung zur Entwicklung und Sicherung von Qualität. Eine Weiterentwicklung der pädagogischen Qualität liegt dabei auch im Interesse der Eltern. So zeigte eine aktuelle Elternbefragung, dass ein Großteil der Eltern bereit wäre, für eine bessere Qualität in ihrer Kita auch höhere Beiträge zu bezahlen – und dies unabhängig von ihrem Einkommen (Bertelsmann Stiftung 2018).

Was aber eine Kindertageseinrichtung von hoher Qualität kennzeichnet, wird im öffentlichen Diskurs oft unterschiedlich definiert. Auch in der Wissenschaft gibt es verschiedene Ansätze, die Qualität von Kitas zu bestimmen und zu bewerten. Um pädagogische Qualität zielgerichtet weiterzuentwickeln, gilt es zunächst zu erkennen, welche Vorstellung von einer „guten Kita“ zugrunde gelegt wird. Die qualitative Studie des Projekts MS-Kita trägt zur Klärung des Konstrukts „Qualität“ bei und richtete dafür ihren Blick auf standardisierte Beobachtungsinstrumente, die pädagogische Qualität in Kitas erfassen sollen.

2.1 Bausteine der Datenerhebung

Die Datenerhebung der qualitativen Studie des Projekts bestand aus den folgenden zwei Bausteinen, die anschließend detaillierter dargestellt werden:

1. einer Dokumentenanalyse von standardisierten Beobachtungsinstrumenten zur Qualitätsbewertung,
2. einer explorativen Feldstudie, die sich aus einer Videografie als auch teilnehmenden Beobachtungen in Kindergartengruppen zusammensetzte.

Dokumentenanalyse von Instrumenten zur Qualitätsbewertung

In einer Dokumentenanalyse wurden die Inhalte ausgewählter standardisierter Beobachtungsinstrumente für die Qualitätsbewertung systematisch analysiert. Leitfragen waren dabei unter anderem, wie Qualität im jeweiligen Instrument definiert wird und welche Aspekte erfasst werden.

Insgesamt wurden fünf Instrumente detailliert analysiert, die unter anderem in der Forschung häufig zum Einsatz kommen (CLASS-PreK, KES-RZ), denen jeweils eine spezifische Fokussierung zugrunde liegt (LIS, ASOS), aber auch ein Instrument, welches jüngst in Deutschland entwickelt wurde (GInA).

- Das „Classroom Assessment Scoring System“ (CLASS-PreK) ist ein Instrument, das im Jahr 2008 erstmals veröffentlicht wurde. Es basiert auf Erkenntnissen der Entwicklungspsychologie und Forschungsergebnissen, die Interaktionen zwischen Kindern und Erwachsenen als entscheidenden Faktor für die kindliche Entwicklung und das kindliche Lernen betrachten (Pianta/La Paro/Hamre 2011, S. 1). Entwickelt wurde die CLASS-PreK für die Altersgruppe von 3 bis 5 Jahren, sodass das Instrument in Deutschland in Kindergartengruppen zum Einsatz kommt. Die Qualität in „classrooms“ wird in diesem Instrument durch drei unterschiedliche Bereiche dargestellt: „Emotional support“ (Emotionale Unterstützung der Kinder), „Classroom organization“ (Organisation der Gruppe) und „Instructional support“ (Lernunterstützung der Kinder). Diese Bereiche teilen sich wiederum auf, sodass insgesamt zehn unterschiedliche Kategorien entlang einer Skala von 1 bis 7 abgefragt werden. Die Werte 1 und 2 stehen hier für eine niedrige, 3 bis 5 für eine mittlere sowie 6 und 7 für eine hohe Qualität. Anzumerken ist, dass die Bewertung des „Negative Climate“, einer Subskala des „Emotional Support“, einer umgekehrten Logik folgt: So entspricht die Vergabe des Wertes 1 einer 7, 2 einer 6, usw. Wenn also beispielsweise mit einer 1 bewertet wurde, entspricht dies einem sehr guten Ergebnis (Beispiel: „Die Fachkraft und Kinder sind im Umgang miteinander nicht sarkastisch oder respektlos.“). Für die Einschätzung einer Kindergruppe werden an einem Vormittag über einen Zeitraum von mindestens zwei Stunden mehrere Beobachtungsdurchgänge von etwa 20 Minuten realisiert, anschließend werden basierend auf den Notizen der Beobachterinnen und Beobachtern Bewertungen vergeben. Möglich ist auch eine Bewertung anhand videografierter Filmsequenzen.
- Die Instrumente „Leuener Engagiertheitsskala“ (LIS) und „Adult Style Observation Schedule“ (ASOS) wurden beide am belgischen Centre for Experiential Education (CEGO) der Universität Leuven von Ferre Laevers und Kolleginnen bzw. Kollegen entwickelt. Grundlage beider Instrumente ist der „Experiential Education“-Ansatz (EXE), welcher das Ziel eines prozessorientierten Monitorings für die Weiterentwicklung der Qualität in der Kita verfolgt. Mit dem Instrument LIS werden das Wohlbefinden und die Engagiertheit einzelner Kinder in einer Kindergruppe erfasst. Diese Aspekte werden

auf einer Skala von 1 (niedrig) bis 5 (hoch) eingeschätzt. Hierfür gibt es zwei unterschiedliche Verfahren, das „Scanning“ und das „Screening“. Im Rahmen des „Scanning“-Verfahrens werden die Kinder in der Gruppe jeweils zwei Minuten beobachtet und anschließend Werte für die Aspekte Wohlbefinden und Engagiertheit vergeben. Dieses Verfahren wurde auch für die Feldstudie angewandt, wobei hier ausschließlich Bewertungen für die in den Filmsequenzen sichtbaren Kinder vergeben wurden. Für das „Screening“-Verfahren werden die Kinder dreimal im Jahr beobachtet, um fundierte Aussagen zu ihrem Wohlbefinden und ihrer Engagiertheit treffen zu können (Laevers/Moons/Declercq 2012).

- Das Instrument ASOS legt den Schwerpunkt hingegen auf die Fachkraft und ihre Handlungskompetenzen in der Interaktion mit dem Kind. Dabei wird bewertet, wie gut es der Fachkraft gelingt, sensitiv mit dem Kind zu interagieren (Sensitivity), Impulse zur Erweiterung kindlicher Denkprozesse zu geben (Stimulation) und die kindliche Autonomie zu unterstützen und wertzuschätzen (Autonomy). Für die Feldstudie wurde für jede 20-minütige Filmsequenz jeweils ein Wert für die genannten Aspekte vergeben und diese Einschätzung schriftlich in kurzen Ausführungen begründet. Die Einschätzungen erfolgen auf einer Skala von 1 bis 7. Zu diesem Instrument gibt es noch kein offizielles Handbuch; es wurde daher auf Materialien aus einer Schulung zurückgegriffen.
- Auch das in Deutschland entwickelte und videogestützte Evaluationsinstrument „Gestaltung von Interaktionen im Alltag“ (GInA) stellt das Interaktionsverhalten bzw. interaktionsbezogene Handlungskompetenzen in den Fokus (Weltzien u.a. 2017, S. 15f.). Da es erst 2015 entwickelt bzw. 2017 veröffentlicht wurde, ist es in der Forschung bisher selten zum Einsatz gekommen. Drei Skalen werden bewertet: „Beziehung gestalten“, „Denken und Handeln anregen“ und „Sprechen und Sprache anregen“. Diese enthalten zusammen 22 Subskalen, wobei sich alleine elf auf den Aspekt „Beziehung gestalten“ beziehen. Zur Bewertung der Filmsequenzen wird ein mikroanalytisches Vorgehen durchgeführt, indem jeweils 4- bis 6-minütige Filmsequenzen betrachtet werden. Die Werte reichen von 1 bis 7, wobei 1 die geringste, 4 die mittlere und 7 die höchste Ausprägung darstellt.
- Mit der „Kindergarten Skala“ (KES-RZ) (Tietze u.a. 2017) ist ein Instrument zur Analyse hinzugezogen worden, welches ebenfalls ursprünglich in den Vereinigten Staaten entwickelt wurde (Harms/Clifford/Cryer 2015). 1997 erschien das Instrument unter dem Namen „Kindergarten-Einschätz-Skala“ erstmals auch für den deutschsprachigen Raum, 2018 die Weiterentwicklung dieser Skala, die KES-RZ. Unter den fünf ausgewählten Instrumenten ist die KES-RZ mit insgesamt acht Bereichen (1. Räume und Ausstattung, 2. Pflege und Routinen, 3. Sprachliche und kognitive Anregungen, 4. Aktivitäten, 5. Interaktionen, 6. Strukturierung der pädagogischen Arbeit, 7. Eltern und pä-

dagogische Fachkräfte, 8. Übergänge) das umfangreichste. Die Einschätzungen beruhen auf einer drei- bis vierstündigen Beobachtung in einer Kindergartengruppe und einem Interview mit einer Fachkraft zu Aspekten, die sich für die Beobachtung weniger eignen, etwa die Zusammenarbeit mit Eltern. Dementsprechend kann die KES-RZ ausschließlich vor Ort und nicht über Filmsequenzen angewandt werden. Die Skalenstufen reichen von 1 bis 7, wobei 1 für eine unzureichende, 3 für eine minimale, 5 für eine gute und 7 für eine ausgezeichnete Qualität steht (Tietze u.a. 2017).

Explorative Feldstudie

Videografie in Kindergartengruppen

Im Zeitraum von Oktober bis Dezember 2017 besuchte ein Forscherinnenteam sechs Kindergärten unterschiedlicher Trägerschaft an verschiedenen Standorten im Großraum München. Dabei wurde darauf geachtet, dass ausschließlich Gruppen mit Kindern im Alter von drei bis sechs Jahren Teil der Stichprobe waren. So konnten jeweils die gleichen Instrumente, die zum Teil zielgruppenspezifisch entwickelt wurden, angewendet werden. In einem drei- bis viertägigen Feldaufenthalt wurden in jeder Kindergartengruppe jeweils sechs etwa 20-minütige Filmsequenzen erstellt, die Interaktionen zwischen pädagogischen Fachkräften und Kindern fokussieren. So entstanden insgesamt 36 Filmsequenzen. Die Aufnahmen zeigten den Morgenkreis, das Mittagessen sowie pädagogische Angebote und Spielsituationen, bei denen die Fachkraft mit einer Kleingruppe agiert (z. B. beim Basteln, bei einem Singspiel oder beim gemeinsamen Puzzeln). Bei der Erstellung der Filmsequenzen wurde darauf geachtet, dass die Anforderungen an Länge und Szenerie aller einbezogener Beobachtungsinstrumente erfüllt wurden.

Die Filmsequenzen wurden durch externe Expertinnen und Experten für vier der Instrumente (CLASS-PreK, GInA, LIS, ASOS) unabhängig voneinander bewertet. Da für das Rating unter Anwendung der CLASS-PreK eine Gesamtbeobachtungsdauer von mindestens zwei Stunden pro Kindergartengruppe festgelegt ist, wurden alle 36 Filmsequenzen anhand dieses Verfahrens beurteilt. Für 18 von 36 Filmsequenzen liegen auch Ratings der Instrumente LIS, ASOS und GInA vor. Zusätzlich führte eine geschulte Expertin für jede Kita jeweils vor Ort ein Qualitätsrating nach der Kindergarten-Skala (KES-RZ) durch.

Beobachtungen im Kita-Alltag

In vier der sechs Kindergartengruppen wurde der Kita-Alltag ergänzend – und unabhängig von den Videografien – von einer Forscherin des Teams beobach-

tet. Ziel war es, die Situation und den Kontext in der Gruppe stärker zu beleuchten und herauszufinden, welcher Unterschied oder Mehrwert entsteht, wenn das pädagogische Handeln vor Ort und nicht standardisiert beobachtet wird. Das Augenmerk lag auf den sozialen Interaktionen zwischen Handelnden, die hinsichtlich ihrer Muster rekonstruiert wurden (Lamnek 2005). Neben den videografierten Sequenzen stand es der Forscherin frei, weitere Interaktionssituationen im Gruppenalltag, aber auch einzelne Kinder über einen gewissen Zeitraum zu beobachten und dies zu protokollieren. Zusätzlich wurden Protokolle erstellt, die sich allgemein mit der Einrichtung befassen, beispielsweise mit dem Gebäude, den täglichen Abläufen oder der Gruppenstruktur. Insgesamt entstanden so 33 Beobachtungsprotokolle aus vier Kindergartengruppen.

2.2 Auswertungsstrategie

Auswertung der Dokumente

Durch tiefgehende inhaltliche Analysen der Handbücher standardisierter Beobachtungsinstrumente als auch der einzelnen Items, die anhand von Beobachtungen bewertet werden sollen, konnten Unterschiede und Gemeinsamkeiten zwischen den Instrumenten erkannt werden (Kuckartz 2016). Darüber hinaus wurden das pädagogische Grundverständnis sowie der Qualitätsbegriff, der dem jeweiligen Instrument zugrunde liegt, herausgearbeitet (detaillierter zum methodischen Vorgehen siehe Eberlein/Schelle 2018).

Zu diesem Zweck wurden Textstellen in den Manualen als auch die einzelnen Items in den Instrumenten anhand eines Kategoriensystems, das durch das Forscherinnenteam aufbauend auf dem Datenmaterial entwickelt wurde, kodiert. So war es möglich, die Inhalte der Instrumente thematisch zu bündeln und in einer Themenmatrix überblicksartig abzubilden. Darauf aufbauend wurden insbesondere Kategorien tiefgehend analysiert, die in allen Instrumente vorkommen und die vielfältigen Aspekte einer gelungenen Interaktion zwischen pädagogischen Fachkräften und Kindern beschreiben (kognitiv anregende Interaktionen, Sensitive Responsivität, Autonomie, Beziehungsgestaltung). Auf dieser Basis entstanden Teilauswertungen bestimmter Aspekte, wie zur Abbildung einer Sensitiven Responsivität in den Instrumenten (Schelle/Eberlein/Durand, i. E.) sowie zum kindlichen Wohlbefinden (Eberlein/Schelle 2018).

Auswertung der Beobachtungsergebnisse

Ziel der Auswertung war es, durch eine interpretierende Auseinandersetzung mit dem Datenmaterial methodische Herausforderungen für die standardisierte Qualitätsbewertung zu identifizieren.

Drei Datengrundlagen wurden dafür genutzt:

- standardisierte Beobachtungsergebnisse in Form von Ratings von 18 Filmsequenzen (drei pro Kindergartengruppe)
- standardisierte Qualitätsbewertung jeder teilnehmenden Kindergartengruppe aufbauend auf Beobachtungen vor Ort mit der KES-RZ
- teilnehmende Beobachtungen in vier der sechs Kitas, die in Form von Beobachtungsprotokollen festgehalten wurden.

Die im Folgenden beispielhaften Ratingergebnisse einer Filmsequenz als auch die Ergebnisse des KES-RZ Ratings vor Ort geben ein Einblick in einen Teil der Datengrundlage.

Ratingergebnisse eines pädagogischen Kreises – ein Beispiel

Beispielhaft für die Datengrundlage der Auswertung zeigen die folgenden Tabellen die Einschätzungen der unterschiedlichen Instrumente zu ein und derselben Filmsequenz eines pädagogischen Kreises in einer der teilnehmenden Kitas. Ebenso sind die Ergebnisse der KES-RZ abgebildet, mit der die pädagogische Qualität der Kindergruppe vor Ort eingeschätzt wurde. Das Beispiel zeigt die unterschiedlichen Qualitätseinschätzungen der 20-minütigen Sequenz.

Tabelle 1: Ratingergebnisse des Instruments „Classroom Assessment Scoring System“ (CLASS-PreK)

Emotional Support		Classroom Organization		Instructional Support	
Positive Climate	6	Behavior Management	5	Concept development	3
Negative Climate	3	Productivity	4	Quality of feedback	2
Teacher Sensitivity	5	Instructional learning formats	2	Language modeling	4
Regard for student Perspective	2				
Gesamtwert	4,5	Gesamtwert	4,3	Gesamtwert	3

Tabelle 1 fasst die Ratingergebnisse zusammen, die unter Anwendung des CLASS-PreK für den pädagogischen Kreis entstanden sind. In der ersten Spalte finden sich die Bewertungen für die Subskalen des „Emotional Support“, in der zweiten für die „Classroom Organization“ und in der dritten Spalte die Bewertungen für die Subskalen des „Instructional Support“. Für jede Subskala wurde anhand der 20-minütigen Filmsequenz ein Durchschnittswert entlang der Stufen 1 bis 7 vergeben. Aus diesen Einzelwerten wurde der Gesamtwert für die drei zentralen Skalen errechnet (4,5; 4,3; 3,0). Die Bewertungen für den pädagogischen Kreis liegen damit im Sinne des CLASS-PreK im mittleren Qualitätsbereich.

Tabelle 2: Ratingergebnisse des Instruments „Gestaltung von Interaktionen im Alltag“ (GInA)

GInA Items	Zeitraum 1 (Min. 0-4)	Zeitraum 2 (Min. 8-12)	Zeitraum 3 (Min. 16-20)
Skala 1: Beziehung gestalten (11 Items)	1,91	2,82	2,73
Skala 2: Denken und Handeln anregen (7 Items)	1,67	2,57	2,29
Skala 3: Sprechen und Sprache anregen (4 Items)	1,75	3,50	2,50
Gesamtwert pro 4-Minuten Sequenz	1,81	2,68	2,55
Gesamtwert	2,35		

In Tabelle 2 werden die Ergebnisse der GInA-Bewertungen des pädagogischen Kreises festgehalten. Diese wurden in drei vierminütige Beobachtungszeiträume gegliedert. Zeitraum 1 umfasst die Minuten 0 bis 4, Zeitraum 2 die Minuten 8 bis 12 und Zeitraum 3 die Minuten 16 bis 20 der Filmsequenz. So wurde für jeden Zeitraum und für die einzelnen Items jeder Skala („Beziehung gestalten“, „Denken und Handeln anregen“, „Sprechen und Sprache anregen“) jeweils ein Wert von 1 bis 7 festgelegt. Bei den in der Tabelle angeführten Werten handelt es sich entsprechend um Durchschnittswerte für die einzelnen Zeiträume. Die Bewertungen befinden sich durchgehend im unteren Qualitätsbereich, es zeigt sich aber zum Ende der Sequenz hin eine (leichte) Verbesserung.

Tabelle 3: Ratingergebnisse des Instruments „Adult Style Observation Schedule“ (ASOS)

	Erläuterung (Auszug)	Gesamtwert
Stimulation	„(...) because the E. [Anm.: die pädagogische Fachkraft] is rather steering in the kind of answers she wants to hear and the fact that children can only answer when sitting still and raising hands, a lot of chances for meaningful communication are taken away.“	3
Sensitivity	„The first part of the clip there is no sensitivity at all. Children are spoken to as a group and have to practice 'sitting quietly in a circle'. For several children it is not at all clear what the E. exactly wants. The E. regularly commands children (...). (...) in general, children are not really been heard or seen.“	2
Autonomy	„Autonomy of children is highly restricted. It is the adult that decides the pace of the activity. (...) The E. asks a lot of questions, but the answers of children are mostly ignored.“	1

Tabelle 3 zeigt ebenfalls die Bewertung des Interaktionsverhaltens der Fachkraft, die den pädagogischen Kreis leitet, in diesem Falle unter Anwendung des Instruments ASOS. Die drei Dimensionen (Stimulation, Sensitivity, Autonomy), eine Begründung der Bewertung (in der Tabelle gekürzt) sowie der Gesamtwert, welcher von 1 bis 7 vergeben werden kann, wird in der Tabelle dargestellt. Die Werte liegen im unteren bis mittleren Qualitätsbereich (1-3).

Tabelle 4: Ratingergebnisse des Instruments „Leuven Involvement Scale“ (LIS)

Kind	Minute	Well-Being	Involvement
1	00:30	2,5	1,5
1	15:30	3,0	2,0
2	02:00	2,5	2,0
2	17:00	2,0	1,0
3	03:30	2,0	1,5
4	05:00	2,0	1,5
4	18:30	2,5	1,5
5	06:30	2,5	2,0
5	20:00	2,5	2,0
6	08:00	2,5	2,5
7	09:30	3,0	3,0
8	11:00	3,0	3,0
9	12:30	2,5	2,0
10	14:00	3,0	2,0
Gesamtwert		2,54	1,96

Tabelle 4 stellt die Ratingergebnisse für den pädagogischen Kreis unter Anwendung des Instrumentes LIS dar. Die Werte für Wohlbefinden (Well-Being) und Engagiertheit (Involvement) der einzelnen Kinder (Kind 1 bis 10) finden sich in den beiden rechten Spalten. Da einige der Kinder pro Filmsequenz mehrfach beobachtet wurden, werden entsprechend bei einigen Kindern – hier Kind 1 und 2 sowie 4 und 5 – zwei Beobachtungszeitpunkte angegeben. Die entsprechende Minute im Film ist ebenfalls in der Tabelle vermerkt. Die Werte des kindlichen Wohlbefindens und der Engagiertheit werden entlang einer Skala von 1 bis 5 vergeben, hier liegen sie also im unteren bis mittleren Bereich. Individuelle Unterschiede sind bei den einzelnen Kindern zu erkennen.

Tabelle 5: Ratingergebnisse des Instruments „Kindergarten Skala“ (KES-RZ)

Platz für Ausstattung	Betreuung & Pflege	Sprachliche & kognitive Anregung	Aktivitäten	
6,13	3,50	2,75	3,00	
Interaktionen	Strukturierung	Eltern & Erzieher	Pädagogische Interaktionen	Ressourcen
2,40	2,67	4,00	3,00	3,70
KES R Gesamtwert	KES Z Gesamtwert			
3,79	4,33			

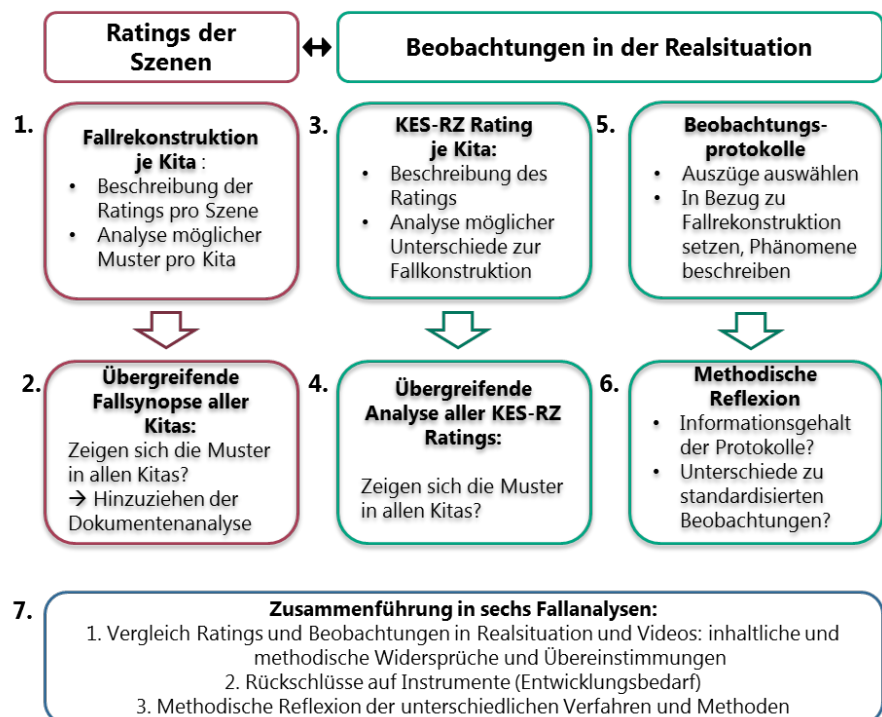
Tabelle 5 zeigt die Ratingergebnisse der KES-RZ, die vor Ort erhoben wurden. Sie beziehen sich somit nicht allein auf den pädagogischen Kreis, sondern auf die gesamte Lernumgebung in der Kindergartengruppe. Für die Auswertung wurde neben der KES-R die Forschungsversion der KES-Z angewandt („Z“ steht für Zusatzmerkmale). So ergeben sich zwei Gesamtwerte. Auch hier werden Bewertungen von 1 bis 7 vergeben. Die Gesamtwerte für die Gruppe liegen im mittleren Skalenbereich und damit zwischen dem minimalen bzw. knapp unter dem guten Qualitätsbereich.

Auswertungsschema

Diese Bewertungen wurden in qualitativen Fallrekonstruktionen detailliert beschrieben, einander gegenübergestellt und mögliche Unterschiede bzw. Übereinstimmungen erschlossen. Dabei war es kein Anliegen, statistische Übereinstimmungen oder die Güte der Instrumente zu überprüfen, sondern interpretierend die Ratings zu derselben Filmsequenz nachzuvollziehen und zu analysieren, wie es zu den Einschätzungen und ggf. Unterschieden kam.

Da sich die Studie durch ein innovatives Design auszeichnet, wurde im Projekt für die qualitative Interpretation der gewonnenen Daten ein spezifisches Auswertungsschema entwickelt. Die folgende Abbildung stellt dieses differenzierte Vorgehen überblickartig dar.

Abbildung 2: Entwickeltes Auswertungsschema der im Rahmen der explorativen Feldstudie generierten Daten



Für jede Kindergartengruppe wurde zunächst anhand der vorliegenden Ratingergebnisse eine Fallrekonstruktion erstellt (1.). Diese fasst Besonderheiten bei der Bewertung zwischen den einzelnen Ratings der drei Filmsequenzen pro Gruppe zusammen. Dabei wurde u.a. nach Auffälligkeiten (Phänomenen), Irritationen sowie Unterschieden und Überschneidungen in den Ratings gesucht und – falls vorhanden – Feldnotizen der Raterinnen und Rater gesichtet, um die Ratings nachvollziehbar werden zu lassen. Abschließend wurde untersucht, inwiefern sich Muster in den Bewertungen einer Kita finden ließen, z.B.: „Das Instrument GInA bewertet den letzten Zeitraum der Filmsequenz tendenziell besser als den ersten“ oder „Der Gesamtwert für die Skala *Emotional Support* fällt immer besser aus als der des *Instructional Support* (CLASS-PreK)“. Insbesondere wurde das Datenmaterial dahingehend analysiert, ob sich Unregelmäßigkeiten zwischen den Ratings durch die Instrumente zeigen, wie etwa, dass eine Filmsequenz im vergleichbaren Qualitätsbereich bewertet wird, eine andere allerdings deutlich unterschiedlich.

In einer Zusammenschau der Ratings der 18 Filmsequenzen (Fallsynopse) (2.) wurden solche (Un)Regelmäßigkeiten kitaübergreifend identifiziert. Bewertet beispielsweise das Instrument ASOS das Handeln der Fachkraft im Morgenkreis immer besser, als es die anderen Instrumente tun? Oder fallen die Werte für die Engagiertheit der Kinder beim Mittagessen immer deutlich schlechter aus, als es in den anderen Sequenzen der Fall ist (LIS)? Ziel dieser Fallsynopse war es aber nicht, die Ratings der Instrumente zu prüfen oder in Frage zu stellen. Vielmehr sollte die Frage beantwortet werden, inwiefern sich daraus fallunabhängige Muster ableiten lassen, die auf Herausforderungen bei der Beurteilung der pädagogischen Qualität in Kitas hindeuten.

Da sich die Instrumente sowohl in ihrem Aufbau als auch in ihrer inhaltlichen Schwerpunktsetzung unterscheiden, konzentrierte sich die Auswertung auf bestimmte inhaltliche Aspekte, die von allen Instrumenten als Qualitätsmerkmal thematisiert werden. Unter Rückgriff auf die Auswertungsschritte bei der Dokumentenanalyse waren vier Qualitätsmerkmale für die Auswertung leitend: „Sensitive Responsivität“ im Interaktionsverhalten der Fachkräfte, „Kognitiv anregende Interaktionen“ zwischen Fachkraft und Kind(ern), Unterstützung der kindlichen „Autonomie“ im Kindergartenalltag sowie die „Beziehungsgestaltung“ zwischen Fachkräften und Kindern.

Abbildung 3: Definitionen von Sensitiver Responsivität, Kognitiv anregenden Interaktionen, Autonomie und Beziehungsgestaltung entsprechend des MS-Kita Kategorienhandbuchs



Parallel zu diesen Bewertungen der Filmsequenzen wurden ebenso die Ergebnisse des KES-RZ Ratings analysiert und in Bezug zu den anderen Datenquellen gesetzt (3.-4.).

Die Protokolle der teilnehmenden Beobachtung, wurden gesichtet, im Team reflektiert, besonders eindrückliche Ausschnitte ausgewählt und in Bezug zu den Rekonstruktionen der Ratingergebnisse gesetzt. Dieses Vorgehen diente vorrangig dazu, den Mehrwert der Informationen, die man aus den Protokollen ziehen kann, zu reflektieren. Darüber hinaus konnte der Kontext der jeweiligen Kita zur Interpretation hinzugezogen werden (5.-6.). Beispiele für solche Ausschnitte aus den Protokollen finden sich bei der Darstellung der Ergebnisse (Kapitel 3).

Die Zusammenführung dieser Daten stellte den letzten Schritt der Auswertung dar und mündete in eine umfassende Fallanalyse pro Einrichtung (7.). Dabei wurden die Ratings der Filmsequenzen und die Beobachtungen in der Realsituation gegenübergestellt, um so methodische Herausforderungen für die Qualitätsbewertung einer Kita zu identifizieren. Im Juli 2018 fand darüber hinaus ein Expertenworkshop am DJI statt, bei dem die durch das Projekt beauftragten externen Raterinnen und Rater die Werte der Qualitätsmessungen zweier beispielhafter Videosequenzen gemeinsam diskutierten. Die Diskussionsergebnisse flossen ebenfalls in die Auswertung ein.

Ergebnisse der qualitativen Studie

Die folgenden Abschnitte fassen die zentralen Ergebnisse der qualitativen Studie in vier zentralen Thesen zusammen. Diese basieren auf sechs qualitativen Fallrekonstruktionen und deren Zusammenführung mit den Beobachtungsprotokollen sowie den Ergebnissen der Dokumentenanalyse. Die Ergebnisse sind zunächst im Kontext der 18 Filmsequenzen aus Kindergartengruppen, die von allen externen Expertinnen bzw. Experten bewertet wurden, zu verstehen. Dabei stellt sich die Frage, ob die Ergebnisse auch auf andere Situationen übertragbar sind. Wie bei allen Ergebnissen qualitativer Forschung muss dies vor dem Hintergrund des jeweiligen Forschungskontextes reflektiert werden.

Bei der vorliegenden Studie scheinen sich bezüglich der Qualitätsbewertung in Kindergartengruppen immer wieder ähnliche Herausforderungen zu ergeben. Diese Vermutung kann durch zwei Annahmen begründet werden: Zum einen sind die Vorgaben in den Instrumenten standardisiert, also unabhängig vom Kontext vergleichbar anzuwenden. Entsprechend scheinen die methodischen Herausforderungen bei der Beobachtung ähnliche zu sein, unabhängig davon, in welcher Kindergartengruppe das Instrument angewendet wird. Zum anderen sind die Interaktionen zwischen Fachkräften und Kindern im Kindergartenalter Kernpunkt der Bewertung der Filmsequenzen. Zentrale Merkmale dieser pädagogischen Interaktionen, wie beispielsweise ihre Einzigartigkeit oder Komplexität, sind vermutlich auch in allen Kindergartengruppen ähnlich zu beobachten (Steinke 1999; Lincoln/Guba in Döring/Bortz 2016).

These 1: Ergebnisse von Qualitätsbewertungen sind vor dem Hintergrund des dahinterliegenden Qualitätsverständnisses und der Schwerpunkte des Instruments zu interpretieren. Nur dann gewinnen sie an Aussagekraft.

Es scheint naheliegend, dass Qualität aus unterschiedlichen Perspektiven unterschiedlich bewertet wird. Dieser Aspekt hat aber weitreichende Konsequenzen für die Weiterentwicklung der Qualität im Arbeitsfeld. Denn die Diskussion um eine beständige Weiterentwicklung der Qualität frühpädagogischer Praxis ist wesentlicher Bestandteil des Fachdiskurses innerhalb der Pädagogik der frühen Kindheit und Motor für vielfältige Entwicklungsprozesse in Praxis, Politik, Forschung und Theorie (Kalicki 2015). So ist es eine entscheidende Frage, wer mit welchem Qualitätsbegriff agiert. Bezogen auf standardisierte Beobachtungsergebnisse in der Kita muss entsprechend geklärt werden, welches Verständnis

von frühkindlicher Bildung, Betreuung und Erziehung den Beobachtungsinstrumenten eigentlich zugrunde liegt.

Ergebnisse der Dokumentenanalyse zeigen, dass in den standardisierten Beobachtungsinstrumenten unterschiedliche inhaltliche Schwerpunkte gesetzt werden und der Begriff „Qualität“ demzufolge unterschiedlich gefüllt wird. So kristallisierten sich in den Handbüchern und einzelnen Items die folgenden Qualitätsverständnisse heraus:

- Das Instrument CLASS-PreK fokussiert vor allem auf Bildungsergebnisse als Indikator für pädagogische Qualität. Auch wenn sich keine explizite Definition des Begriffes Qualität im Handbuch findet, so kann dennoch aus den einzelnen Items geschlossen werden, dass die Interaktionen, deren Feinfühligkeit und Anregungsgehalt sowie die Berücksichtigung kindlicher Interessen und Autonomie vorrangig dem Ziel der kindlichen Schulfähigkeit dienen sollen. Ein solcher Ansatz spiegelt das US-amerikanische System der frühkindlichen Bildung, Betreuung und Erziehung wider, das in vielerlei Hinsicht den deutschen Leitorientierungen, wie etwa der Alltagsintegration von Bildungsprozessen, widerspricht. Bisherige Ergebnisse anderer Forschungsprojekte zeigen entsprechend, dass die CLASS-PreK zur Untersuchung des Zusammenhangs zwischen Interaktionskompetenz und Schulfähigkeit der Kinder eine gute Grundlage bietet (vgl. von Suchodoletz/Gunzenhauser/Larsen 2015; Wertfein/Mayer 2018).
- Die Instrumente LIS und ASOS bauen hingegen auf dem Verständnis auf, wonach entscheidende Voraussetzung für Qualität in einer Kindergruppe eine Prozessqualität ist, die sich im Wohlbefinden und in der Engagiertheit eines jeden einzelnen Kindes zeigt. „Quality in care and education begins with the daily question about how each of the children is doing in the setting“ (Laevers u.a. 2012, S. 17). Verändert sich die Lernumgebung der Kinder so, dass sich alle darin gut entwickeln können und wohl fühlen, ist das ein entscheidender Schritt für die Weiterentwicklung der Qualität.
- Vergleichbar mit der CLASS-PreK findet man auch im Instrument GInA keine eigenständige Definition des Begriffes Qualität. Dies sicher auch, weil die Autorinnen bzw. Autoren selbst deutlich machen, dass die GInA-Skalen keine globale pädagogische Qualität messen. Das Instrument fokussiert bei der Bewertung des pädagogischen Handelns – wie auch das Instrument CLASS-PreK – auf die Interaktionen zwischen Fachkraft und Kind(ern). Dabei bekommen diese aber einen anderen Stellenwert als in dem US-amerikanischen Instrument. Interaktionen sind die Grundlage für qualitätsvolle Beziehungen. Voraussetzung hierfür ist, dass die Interaktionen Sicherheit, Orientierung, Anregung und Anerkennung bieten, was durch eine emotionale Zuwendung der Fachkraft als auch durch Beteiligung und Wertschätzung der Kinder realisiert werden kann.

- Das Qualitätsverständnis der KES-RZ, das im Handbuch ausführlich dargelegt wird, baut auf dem der ECERS-R auf, entspringt also ebenso einer US-amerikanischen Vorstellung von Frühpädagogik. Die Skala wurde – anders als es bei dem Instrument CLASS-PreK der Fall ist – an die pädagogische Praxis in Deutschland sprachlich als auch an manchen Stellen inhaltlich angepasst. Am Instrument wird deutlich, dass sich pädagogische Qualität dadurch auszeichnet, dass die Kindertageseinrichtung den Bedürfnissen der Kinder nach Lernen und Bildung gerecht wird, dabei auch deren Wohlbefinden im Mittelpunkt steht sowie die Familien unterstützt werden. Neben der Struktur- und Orientierungsqualität kennzeichnet im Sinne der KES-RZ pädagogische Qualität insbesondere die Prozessqualität. Diese wird breit aufgefasst, so dass auch die Voraussetzungen für Prozessqualität und Interaktion, wie konzeptionelle Aspekte und räumlich-materiellen Ausstattung und deren Nutzung, Grundlage für die Bewertung sind. Der Schwerpunkt der Items liegt dabei auf Aspekten der Ausstattung und Gestaltung der Räume.

Die Instrumente entspringen unterschiedlichen pädagogischen Denktraditionen. Welche Aspekte des pädagogischen Handelns anhand des Instruments bewertet werden sollen und welche Aussagekraft das Bewertungsergebnis hat, kann man nur dann nachvollziehen und einordnen, wenn man das dahinterliegende Qualitätsverständnis berücksichtigt. Unterschiedliche Bewertungen können in unterschiedlichen Qualitätsverständnissen gründen.

These 2: Standardisierte Bewertungen pädagogischen Handelns in der Kita sind überwiegend eine Momentaufnahme. Sie bilden einen Ausschnitt des pädagogischen Alltags ab.

Die Bewertung von Qualität durch standardisierte Beobachtungen in der Kita ist in großem Maße davon abhängig, welche Situationen die Beobachtenden zur Einschätzung bekommen. Um sich von einem solchen Zufallsprinzip zu distanzieren, verweisen die Autorinnen und Autoren der Instrumente CLASS-PreK und KES-RZ auf deren Gültigkeit. So beziehen sie sich u.a. darauf, dass sich die Bewertungen nicht ändern, wenn man an einem anderen Tag die Kita besucht und somit stabil bleiben. Dies wurde in mehreren Studien bestätigt. Und dennoch bleibt es eine methodische Herausforderung, durch die stark zeitlich begrenzte Beobachtungsdauer keine entscheidenden Momente zu „verpassen“, die die Bewertung verändert hätten. Ein umfassendes Bild von der Qualität des pädagogischen Handelns zu bekommen ist insbesondere dann schwierig, wenn eine alltagsintegrierte, situationsorientierte Pädagogik in der Kita umgesetzt wird. Denn dort variiert der Tagesablauf sehr und verändert sich spontan, je nach den Interessen, Ideen und Bedürfnissen der Kinder. Anders, Roßbach und Tietze (2016) stellen, bezogen auf der Bewertung sprachförderlicher Interaktionen in der Kita im Rahmen der Evaluation des Bundesprogramms „Schwerpunkt-Kitas Sprache und Integration“, zur Diskussion, ob eine 3- bis

4-stündige Beobachtung ausreicht, um ein repräsentatives Bild einer Kindergartengruppe abbilden zu können.

Einen solchen Gedanken greifen die Instrumente GInA und LIS bzw. ASOS auf. So wird im Handbuch des Instrumentes GInA ausdrücklich darauf verwiesen, dass nur kurze Ausschnitte des Kita-Alltags betrachtet werden und noch nicht klar ist, wie viele von diesen kurzen Filmsequenzen benötigt werden, um eine Aussage über die Interaktionsqualität in der gesamten Kindergartengruppe treffen zu können. Die Autorinnen und Autoren der Instrumente LIS und ASOS hingegen gehen davon aus, dass man zu verschiedenen Zeitpunkten beobachten muss, um ein ganzheitliches Bild von Kindern und Fachkräften zu erhalten.

Die Beobachtungsprotokolle aus den vier Kindergartengruppen zeigen, dass allein durch eine drei- oder viertägige teilnehmende Beobachtung eine Fülle an Informationen über die Situation der Einrichtung, über die Fachkräfte als auch über die Interaktionen zwischen den Fachkräften und den Kindern gesammelt werden kann. Die Fallanalyse einer Kindergartengruppe kann hier beispielhaft angeführt werden. Sie verdeutlicht, dass schon ein dreitägiger Aufenthalt tiefergehende Einblicke erlaubt, etwa in das Interaktionsverhalten und die Beziehungsgestaltung zwischen den Fachkräften und den Kindern, da man vielfältigere Situationen erlebt. So „verpassten“ die Raterinnen bzw. Rater eine Situation in dieser Kita, bei der die Kinder gemeinsam im Stuhlkreis Lieder einübten, welche vermutlich eine Bewertung zur Qualität der Interaktionen beeinflusst hätte.

Alle Gruppen versammeln sich zum gemeinsamen Singen. Dabei werden die Kinder von den Fachkräften immer wieder zum stillen Sitzen und Zuhören angehalten, bei Unruhe etwa am Arm gepackt und auf den Boden gesetzt oder mit den Beinen am Aufstehen gehindert. Ein Kind muss sich direkt neben eine Fachkraft setzen, weil es von ihr schon mehrmals zur Ruhe ermahnt wurde. Das Kind beginnt leise zu weinen, ohne dass beobachtet werden kann, dass die Fachkraft darauf reagiert.

Ein weiteres Protokoll in derselben Kita gibt eine Situation wieder, die nach dem Mittagessen beobachtet wurde, als alle Kinder im Bad beim Zähneputzen sind.

Nach ein paar Augenblicken steht Frau M. (Anm.: Fachkraft) im Gruppenraum, in der einen Hand hält sie eine große Aufbewahrungsschachtel für die Zahnbürsten, mit der anderen Hand schiebt sie N. in den Gruppenraum. In lautem Ton sagt sie zu Frau S. (Anm. Fachkraft), die noch im Gruppenraum ist und aufräumt: „Der hat keine Schuhe an! So geht der nicht Zähneputzen!“ Frau S. erklärt, dass seine Hausschuhe nass seien. „Dann ziehst ihm die Straßenschuhe an!“ antwortet Frau

M. „Die sind auch nass“ erwidert Frau S. „Dann holts ihm andere Schuhe. So geht der nicht Zähneputzen.“ Frau M. will aus dem Raum gehen, Norbert will wieder hinterher. „Nein“ sagt Frau M. „die holen dir erst Schuhe.“ Sie setzt ihn auf einen Stuhl, den sie unter einem Tisch hervorzieht und geht schnell aus dem Zimmer. Die Praktikantin, die auch im Raum ist, meint zu Frau S.: „Sollen wir Schuhe von wem anders ausleihen? Der D. ist doch heute nicht da.“ Frau S. schüttelt den Kopf: „Ne, wenn der einnässt, dann sind die Schuhe von wem anders nass, des können wir nicht machen.“ N. sitzt still auf dem Stuhl. (...) Frau S. verlässt den Raum. Die Praktikantin stellt das restliche Geschirr auf den Wagen und fährt diesen in Richtung Flur. Als sie den Wagen an N. vorbeifährt, der immer noch auf dem Stuhl sitzt, sagt sie: „Warte bis Frau S. wiederkommt.“ Sie geht aus dem Zimmer. N. sitzt ganz leise auf seinem Stuhl, jetzt allein im Gruppenraum, dreht sich um, blickt die Wand mit den Bildern, die die Kinder gemalt haben, an. Nach ein paar Augenblicken kommt Frau S. zurück in den Raum. Sie hält eine Zahnbürste in der Hand, kniet sich neben Norbert und gibt sie ihm. „Dann putzt du eben hier Zähne.“ N. steht auf, geht zum Spülbecken der kleinen Küche und beginnt, seine Zähne zu putzen.

Die Bewertungen der Filmsequenzen aus dieser Kindergartengruppe bezogen auf die Beziehungsgestaltung und ein sensibles Interaktionsverhalten liegen unter Anwendung der CLASS-PreK im mittleren bis guten Bereich, bei den Instrumenten ASOS und GInA je nach Sequenz im mittleren bis unteren Bereich. Das KES-RZ Rating bestätigt dieser Kindergartengruppe eine mittlere Qualität (KES-R: 3,79; KES-Z: 4,33). Die Frage, die sich hier stellt, ist, inwieweit die Bewertungen (noch) niedriger ausgefallen wären, wenn die Beobachterinnen bzw. Beobachter über mehrere Tage am Alltag teilgehabt und beispielsweise die oben zitierten, protokollierten Situationen in die Bewertung aufgenommen hätten.

Standardisierte Qualitätsbewertungen können Aussagen zu der jeweils eingeschätzten Situation geben, scheinen aber vorrangig Momentaufnahmen zu sein, die das abbilden, was an diesem Tag und in dieser Situation bewertbar war. Auch wenn die zeitliche Stabilität als Gütekriterium für manche Instrumente in Studien festgestellt wurde, bleibt zu diskutieren, ob aufbauend auf solchen situativen Bewertungen auf eine Gesamtqualität der Kita geschlossen werden kann. Dies gilt es transparent zu machen und entsprechend im Zuge der Bewertung von Qualität zu reflektieren.

These 3: Standardisierte Bewertungen des pädagogischen Handelns entstehen im Kontext subjektiver Eindrücke und objektiver Ansprüche.

Ziel einer Bewertung der Qualität in Kindertageseinrichtungen sollte eine möglichst objektive Einschätzung sein, damit diese auch vergleichbar wird. Dazu ist es notwendig, dass die Raterinnen und Rater subjektive Interpretationen von dem, was sie beobachten, minimieren und sich an die Vorgaben des Instruments halten. Das ist eine große Herausforderung, denn diese werden als Beobachterinnen und Beobachter des Geschehens zunächst unausweichlich durch ihre subjektiven Empfindungen und Wahrnehmungen beeinflusst. Reichertz bezeichnet Forschende, in der vorliegenden Studie also auch die beteiligten Beurteilenden, als „Instanz der Beobachtungen“ (Reichertz 2016, S. 215), die nicht nur die Erhebung, sondern auch die Auswertung der Daten durch die eigene Involviertheit verändern. Er geht so weit, dass man diese Subjektivität zwar minimieren, etwa durch Reflexion, aber nie vollständig eliminieren könne. Die Frage ist also, wie objektive Bewertungen überhaupt möglich werden.

Ergebnisse der Dokumentenanalyse zeigen, dass sich die Autorinnen und Autoren der Instrumente sehr unterschiedlich mit der Herausforderung, objektive Bewertungen in den Qualitätseinschätzungen zu gewährleisten und die Bedeutung subjektiver Eindrücke der Beobachterinnen und Beobachter zu minimieren, auseinandersetzen.

- In den Instrumenten ASOS und LIS wird die Frage nach der Subjektivität kaum thematisiert, zugleich aber auch kein Anspruch auf größtmögliche Objektivität erhoben. Sie schlagen vor, wiederholte Beobachtungen über einen längeren Zeitraum durchzuführen. Ergänzend hierzu sollen Kontextanalysen durchgeführt werden, um mögliche Einflussfaktoren auf das Rating (wie z. B. Raum-, Gruppen- oder Materialsituation) zu ermitteln.
- Durch Schulungen und regelmäßige Reliabilitätsprüfungen sollen mittels des Instruments CLASS-PreK möglichst objektive Einschätzungen gewährleistet werden. Akzeptable Werte ihrer Interrater-Reliabilität, also dem Maß der Überschneidung von Ratingergebnissen durch mind. zwei CLASS-PreK-Raterinnen bzw. -Rater, verweisen auf ein gewisses Maß an Objektivität bei der Bewertung. Zu berücksichtigen ist dabei, dass die einzelnen Subskalen der CLASS-PreK aus drei bis vier Merkmalen bzw. Items bestehen, welche jeweils in die Beobachtung miteinbezogen und anschließend in einen Gesamtwert überführt werden. Für die Errechnung einer Interrater-Reliabilität sind dabei, so weist es zumindest das auf dem Markt zugängliche Manual aus, nur diese Gesamtwerte für die einzelnen Subskalen wichtig, die vorgesehenen Notizen in den Beobachtungsbögen spielen bei solchen statistischen Berechnungen keine Rolle. So wird allein aufgrund einer hohen Interrater-Reliabilität nicht deutlich, ob die Raterinnen bzw. Rater die einzelnen Items unterschiedlich bewertet, aber bei der Einschätzung der Subskala womöglich denselben

Gesamtwert festgelegt haben. Das nachzuvollziehen ist auch deshalb schwierig, weil die Items zum Teil unterschiedliche Aspekte enthalten (Mayer/Beckh 2018). In der Subskala „Language Modeling“ werden zum Beispiel Items angeführt, die sich auf die Peer-to-Peer-Interaktionen beziehen, auf die Art, wie Fachkräfte Fragen stellen, wie sie auf das, was die Kinder sagen, reagieren oder auf ihren Sprachgebrauch. Da es keine einzelnen Zahlenwerte für die Items gibt, wird die Bewertung einzelner Facetten nicht deutlich.

- Für die Anwendung des Instruments GInA zu Forschungs- und Evaluationszwecken wird ebenso eine Schulung empfohlen. Die Übereinstimmung der Beurteilenden wird, basierend auf statistischen Überprüfungen von elf Sequenzen, als zufriedenstellend eingeschätzt. Nur einzelne Items lagen unter einem noch zufriedenstellenden Wert, sind also nur schwer objektiv einzuschätzen. Mit einem expliziten Austausch zwischen den Beurteilenden konnten diese Werte aber nochmals erhöht werden. Ähnlich wie bei der CLASS-PreK bietet das Instrument für jede Bewertungsstufe aber mehrere ausschlaggebende Aspekte an, von denen nur eines zutreffen muss. So muss etwa in der Subskala 1 „Zuwendung zeigen“ für die Bewertung mit 7 (höchste Ausprägung) nur eines der Items zutreffen: „Es sind deutliche Signale von Zuwendung (...) über die gesamte Sequenz zu beobachten“ *oder* „Die Fachkraft folgt allen Kindern in ihren Aktivitäten/Äußerungen feinfühlig und gibt passgenaue Impulse der Zuwendung“ *oder* „Die Fachkraft wirkt kongruent in ihren verbalen/nonverbalen Signalen; sie beachtet dabei die Heterogenität der Kinder (...) und geht feinfühlig darauf ein“. Welches Item letztendlich für die Raterinnen bzw. Rater maßgeblich war, kann aufgrund des Zahlenwerts allein nicht nachvollzogen werden.
- In der KES-RZ wird, ähnlich wie in der CLASS Pre-K, die Notwendigkeit von Schulungen für die Anwendung des Instrumentes thematisiert. Auch in der KES-RZ wird auf eine in Studien geprüfte hohe Interrater-Reliabilität verwiesen. Dennoch bleibt die Frage, inwieweit eine einmalige Schulung ausreicht, um möglichst intersubjektiv übereinstimmende Bewertungen zu erhalten. Darüber hinaus bleibt auch hier ein gewisser Interpretationsspielraum, wie es zu Einschätzungen mancher Items kommt.

Die Frage nach der Vermeidung subjektiver Interpretationen bei standardisierten Beobachtungen scheint insbesondere bei der Bewertung von sensitiven und responsiven Interaktionen eine Rolle zu spielen, wie die Zusammenfassung der sechs Fallrekonstruktionen zeigt. Bei der Bewertung von Aspekten, die sich auf ein sensitives und responsives Interaktionsverhalten beziehen, sind die Abstände der Ratingergebnisse der drei Instrumente sehr unterschiedlich. In den Handbüchern und Items wird das, was als sensitives Interaktionsverhalten verstanden wird, durchaus eindeutig, wenn auch unterschiedlich, definiert. So hätte man erwarten können, dass die Ratingergebnisse diese unterschiedlichen Schwerpunkte der Beobachtung konsequent abbilden. Aufbauend auf den Fall-

rekonstruktionen zeigt sich hier offensichtlich eine methodische Herausforderung. Bewertungen, die Wärme, Feinfühligkeit oder Wertschätzung erfassen sollen, sind möglicherweise subjektiven Interpretationen unterworfen, die sich allein durch ein standardisiertes Vorgehen nicht vollständig eliminieren lassen.

Betrachtet man die Fallanalyse einer Kindergartengruppe, so findet sich ein mögliches Beispiel für solche subjektiven Interpretationen:

Die videografierte Sequenz zeigt ein pädagogisches Angebot, bei dem eine Fachkraft mit einigen Kindern am Tisch etwas bastelt. Die übrigen Kinder beschäftigen sich selbstständig im gesamten Raum, spielen auf dem großen Teppich, im Hochbett oder essen ihre Brotzeit. Am Ende der Filmsequenz kommt ein Kind zur Fachkraft, das mit drei anderen Kindern auf dem Hochbett gespielt hat. Es klagt, dass ein anderes Kind ihm auf den Kopf gehauen hätte. Die Fachkraft ruft alle Kinder, die oben gespielt haben, zu sich. Das Kind wiederholt, dass ihm auf den Kopf geschlagen worden wäre. Als alle beteiligten Kinder vor ihr stehen, spricht die Fachkraft die Konsequenz aus, dass vorerst keiner mehr auf dem Hochbett spielen dürfe, wenn es da Ärger gebe.

Im Expertengespräch mit den externen Raterinnen und Ratern wurde diese Szene kontrovers diskutiert. So wurde auf der einen Seite wahrgenommen, dass die Fachkraft auf das Kind, dem auf den Kopf geschlagen wurde, nicht angemessen reagiert habe – obwohl das Kind Schmerzen und einen Bedarf an Zuwendung hatte und in einer Stresssituation war. Die Fachkraft aber sprach lediglich eine Kollektivstrafe aus. Von anderen Teilnehmenden wurde die Situation so interpretiert, dass das Kind zufrieden mit der Reaktion der Fachkraft war, da sie klar formuliert, dass anderen Schmerzen zufügen nicht akzeptabel ist. Obwohl die Fachkraft in der Situation mehr Empathie hätte zeigen können, sei das Kind aber deswegen nicht frustriert oder unglücklich gewesen. Die Vermutung liegt nahe, dass sich die unterschiedlichen Einschätzungen der Situation auch in der standardisierten Bewertung durch die jeweiligen Expertinnen und Experten niederschlagen. So zeigten sich hinsichtlich der Bewertung der Sensitivität der Fachkraft leichte Unterschiede in den Ergebnissen: Während mittels CLASS-PreK in der Subskala „Teacher Sensitivity“ die bestmögliche Bewertung von 7 vergeben wird, fallen die Werte der Instrumente ASOS in der Skala „Sensitivity“ (6) und GInA in den Subskalen „Zuwendung zeigen“, „Aufmerksam zuhören“, „Verstehen ausdrücken“, „Aufgeschlossen sein“ (Durchschnittliche Bewertung von 5) etwas schlechter aus.

Dieses Beispiel zeigt, dass das beobachtete Handeln der Beteiligten unterschiedlich interpretiert werden kann. Im Expertengespräch wurde deutlich, dass eine standardisierte Einschätzung, welche Bedürfnisse ein Kind hat und ob die Fachkraft diesen auch gerecht wird, durch eigene subjektive Empfindungen als Beobachterin bzw. Beobachter deutlich erschwert wird. Was als offen, zugewandt oder interessiert interpretiert wird, hängt auch von (beruflichen) Erfahrungen,

Einstellungen und den Biographien der Raterinnen bzw. Rater ab. Diese scheinen gerade für solche „weichen“ Beobachtungsdaten relevant (vgl. dazu Reichertz 2015).

These 4: Qualitätsbewertungen einer Kindergartengruppe lassen keine eindeutigen Rückschlüsse auf das Wohlbefinden einzelner Kinder in der Kita zu.

Das kindliche Wohl in der Einrichtung zu sichern und zu fördern ist Auftrag der Kindertageseinrichtung. Dies lässt sich aus zwei Perspektiven heraus begründen: Zum einen wird das Wohl des Kindes in Artikel 3 der in Deutschland ratifizierten UN-Kinderrechtskonvention (BMFSFJ 2018) als das vorrangigste Bewertungskriterium für alle Maßnahmen im öffentlichen und privaten Bereich benannt. Kindertageseinrichtungen als Teil des öffentlichen Bereichs können sich diesem Auftrag also nicht entziehen (Ben-Arieh 2010; Bertram 2016). Zum anderen haben Kindertageseinrichtungen nach § 22 des Achten Sozialgesetzbuch (SGB VIII) den Auftrag, die Entwicklungs- und Bildungsprozesse des Kindes zu unterstützen. Voraussetzung für solche kindlichen Entwicklungs- und Lernprozesse ist, dass sich das Kind in der Einrichtung wohlfühlt und die Erfüllung seiner grundlegenden Bedürfnisse im Mittelpunkt steht (Mashford-Scott/Church/Taylor 2012). Das Sichern und Fördern kindlichen Wohlbefindens ist demnach Kernelement pädagogischer Arbeit und sollte konsequenterweise auch als wichtiger Indikator pädagogischer Qualität betrachtet werden.

Die Bewertungen der Filmsequenzen unter Anwendung der LIS zeigen, dass es teilweise deutliche Unterschiede zwischen dem in der konkreten Situation erlebten Wohlbefinden der Kinder gibt. Dies führt zu dem Rückschluss, dass nicht alle Kinder in demselben Maße von der erlebten pädagogischen Qualität profitieren und Fachkräfte Kinder unterschiedlich gut erreichen. Das haben auch andere Studie bereits bestätigt (vgl. dazu Smidt 2012).

Tabelle 6: Bewertung Well-Being durch LIS in Fallbeispiel

Well-Being				
	Erste Beobachtung	Zweite Beobachtung	Dritte Beobachtung	Vierte Beobachtung
Kind 1	4,0	3,5	4,0	4,0
Kind 2	5,0	4,5	4,5	4,5
Kind 3	3,5			
Kind 4	3,5			
Kind 5	4,0			

Die Tabelle gibt diese Bewertungen beispielhaft für ein videografiertes pädagogisches Angebot wieder, bei dem die Kinder mit der Fachkraft am Tisch basteln. Die Tabelle zeigt die unterschiedlichen Werte für die Kinder, die zwischen 3,5 und 5 variieren (bei einer Skala von 1 bis 5). Auch verändern sich die Werte für das Wohlbefinden teilweise über den Zeitverlauf zu den unterschiedlichen Beobachtungszeitpunkten, zwischen denen immer ein paar Minuten liegen. Zwei Kinder wurden mehrfach beobachtet.

Das Instrument LIS ist das einzige in der Studie angewendete Instrument, das eine zielkindorientierte Beobachtung des Wohlbefindens vorsieht. Eine Teilauswertung der Dokumentenanalyse zeigte, dass in den Instrumenten GInA, CLASS-PreK als auch ASOS keine Einschätzung des individuellen kindlichen Wohlbefindens in der Kindergruppe vorgesehen ist. Im Fokus der Beobachtungen stehen die Fachkräfte und Rahmenbedingungen. Die Rating-Ergebnisse der Instrumente lassen demnach Aussagen darüber zu, ob Kinder Bezugspersonen vorfinden, die Anteil nehmen, auf ihre Bedürfnisse eingehen und die Erfüllung dieser ermöglichen und somit kindliches Wohlbefinden fördern (Deci/Ryan 1993). Eindeutige Aussagen dazu, inwiefern diese Umgebung tatsächlich eine positive Wirkung auf das Wohlbefinden einzelner Kinder zeigt, können hingegen nicht gemacht werden. Hohe Qualitätsbewertungen unter Anwendung dieser Instrumente sind also nicht zwingend Ausdruck dessen, dass sich alle Kinder in der Kindergruppe bzw. der beobachteten Situation wohlfühlen. Ebenso wenig sagen niedrige Qualitätsbewertungen zwangsläufig aus, dass sich alle Kinder entsprechend unwohl fühlen (Eberlein/Schelle 2018).

Auch die Fallanalyse gibt Hinweise, dass ein solcher Rückschluss nicht automatisch möglich ist. Im oben genannten Beispiel, bei dem die Fachkraft mit den Kindern am Tisch bastelt, wird das Handeln der Fachkraft durchgängig im mittleren bis hohen Qualitätsbereich bewertet. Insbesondere unter Anwendung des Instruments ASOS werden für diese Sequenz die Werte 6 (Sensitivity), 6 (Stimulation) und 7 (Autonomy) vergeben, was auf eine hohe Interaktionsqualität hindeutet. Setzt man diese Bewertungen in Bezug zu den unterschiedlichen

Werten für das Well-Being der Kinder kann man auch hier schlussfolgern, dass eine hohe Interaktionsqualität nicht ohne weiteres in gleichem Maße das Wohlbefinden aller Kinder positiv beeinflusst. Beobachtungsergebnisse anhand der untersuchten Instrumente müssen diesbezüglich entsprechend reflektiert werden (Eberlein/Schelle 2018).

4.

Rückschlüsse für die Weiterentwicklung

Ziel der qualitativen Studie war es zu analysieren, wie Qualität in der Kita durch standardisierte Beobachtungsinstrumente erfasst wird und welche methodischen Herausforderungen sich dabei zeigen. Von den Ergebnissen der Studie können Weiterentwicklungsbedarfe für die Bewertung des pädagogischen Handelns abgeleitet werden.

4.1 Empfehlungen zur Weiterentwicklung

1. Bei der Bewertung des pädagogischen Handelns in der Kita sollten inhaltliche Schwerpunkte und methodische Herausforderungen transparenter werden.

Die dargestellten Ergebnisse zeigen, dass im Diskurs um eine gesteigerte Qualität genauer differenziert werden sollte, welche Akteure mit welchem Qualitätsverständnis agieren. Die Kindertageseinrichtungen werden konstant mit neuen Anforderungen konfrontiert. Daraus ergibt sich eine große Bandbreite an Qualitätsfaktoren, die sich beispielsweise von flexiblen Öffnungszeiten, einem besseren Fachkraft-Kind-Schlüssel, der Qualifikation der Fachkräfte und ganz aktuell über die Gebührenfreiheit in Kitas erstreckt und sich entsprechend hauptsächlich auf strukturelle Faktoren der Qualität beziehen. Diese Bandbreite spielt je nach Adressatinnen bzw. Adressaten, seien es die Eltern, Fachkräfte, Leitungen, Träger oder die Kinder selbst, eine größere oder kleinere Rolle. Flexible Öffnungszeiten sind beispielsweise ein besonders wichtiges Thema für Eltern, während für die Kinder selbst andere Aspekte im Vordergrund stehen könnten.

Diese Vielfalt an Qualitätsverständnissen findet sich auch bei denen im Projekt analysierten Messinstrumenten wieder. So konnte gezeigt werden, dass diese durch ihre inhaltliche Ausrichtung und unterschiedlichen methodischen Herangehensweisen auch zugleich eine Vielfalt an verschiedenen Facetten von Qualität abbilden. Diese Vielfalt sollte bewusst transparent gemacht und auch methodische Schwierigkeiten offensiv benannt werden. So wird deutlich, welchen Ausschnitt der pädagogischen Realität mit welchen Leitorientierungen und methodischen – womöglich auch nicht zu lösenden – Schwierigkeiten durch die Instrumente erfasst werden. Für die Aussagekraft einer Qualitätsbewertung sind diese Informationen unabdingbar.

Mit Blick auf wissenschaftliche Studien oder das Monitoring von Qualität in Kindertageseinrichtungen, wie beispielsweise das Monitoring zur Umsetzung des Gesetzes zur Verbesserung der Qualität in der Kindertagesbetreuung („Gute-Kita“-Gesetz) gewinnen diese Informationen nochmals besondere Bedeutung. Daran schließt sich die Frage an, wie und ggf. unter Anwendung welcher Instrumente eine mögliche Steigerung der Qualität überprüft werden kann.

2. Kindliches Wohlbefinden sollte Bestandteil der Qualitätsbewertung sein.

Ergebnisse der Dokumentenanalyse verdeutlichen, dass das kindliche Wohlbefinden bei den analysierten Beobachtungsinstrumenten – mit Ausnahme des Instrumentes LIS, welches dieses als zentrale Variable erachtet – nur indirekt Teil einer Bewertung ist. Hier zeigt sich ein Weiterentwicklungsbedarf bei der Bewertung der pädagogischen Arbeit in frühpädagogischen Institutionen. Es gilt zu prüfen, wie Wohlbefinden in der Kita differenziert beobachtbar und erfassbar wird, und darüber hinaus einen methodischen Perspektivwechsel zu vollziehen: Der Blick wird bei der Bewertung von Qualität nicht mehr (ausschließlich) auf die Gruppenebene gelenkt, also auf die Lernumgebung des Kindes, sondern auf das Kind selbst und seine individuell erlebte Qualität.

Das von Laevers, Moons und Declercq (2012) entwickelte „Process-Oriented Monitoring System for the Early Years (POMS)“, das auch die LIS enthält, kann als Ausdruck eines solchen methodischen Perspektivwechsels verstanden werden. Zielkindfokussierte Beobachtungen des individuellen „well-beings“ und „involvement“ der Kinder sowie begleitende Beobachtungen des Interaktionsverhaltens der Fachkraft und der Lernumgebung sind bei diesem Verfahren Grundlage für eine Weiterentwicklung der pädagogischen Qualität. Solche Beobachtungen des kindlichen Wohlbefindens machen es möglich, den individuell erlebten Qualitäten des einzelnen Kindes in der Qualitätsbewertung einer Einrichtung mehr Gewicht zu verleihen und so die Bewertungen des Lernumfeldes entscheidend zu ergänzen. Dadurch gewinnen Aussagen über Qualität in der Kita mehr Gewicht (Smidt/Roßbach 2015; Melhuish 2001; Dunn 1993; Hallam u.a. 2009).

Darüber hinaus wird man damit dem Umstand gerecht, dass Kinder als Voraussetzung für Selbstbildungsprozesse und zur Steigerung ihres Wohlbefindens unterschiedliche Anregungs- und Beziehungsangebote benötigen. Für die Realisierung und Weiterentwicklung einer individuellen und kompensatorischen Unterstützung und Begleitung kindlicher Entwicklungs- und Bildungsprozesse in Kindertageseinrichtungen scheint dies unerlässlich.

4.2 Ausblick

Auch in den nächsten Jahren werden die Fragen nach Qualität in der Kindertageseinrichtung die Entwicklung des gesamten Arbeitsfeldes der frühkindlichen Bildung und Betreuung entscheidend prägen: Durch den fortwährenden Ausbau der Kindertagesbetreuung haben immer mehr Kinder länger und früher am System der Kindertagesbetreuung teil, sodass der Bereich entsprechend an gesellschaftlicher Bedeutung gewinnt. Damit steigt auch die öffentliche Verantwortung, in diesen Institutionen Lernumgebungen für Kinder zu schaffen, die sich durch ihre hohe Qualität positiv auf deren individuelle Entwicklung und Bildungsprozesse auswirken. Das Gesetz der Bundesregierung zur Verbesserung der Qualität in der Kindertagesbetreuung („Gute-Kita“-Gesetz) kann als erste bedeutende Konsequenz einer solchen gesamtgesellschaftlichen Verantwortung verstanden werden. Abzuwarten bleibt, wie mögliche Erfolge der Umsetzung dieses Gesetzentwurfes und damit die Weiterentwicklung der Qualität vor Ort abgebildet und gesichert werden können.

Angesichts dieser Entwicklungen stellt sich umso drängender auch die Frage an die Forschung, welche Faktoren für die pädagogische Qualität entscheidend sind, wie diese transparent abgebildet und damit sie auch gezielt weiterentwickelt werden können. Denn standardisierte Qualitätsbewertungen sollten vorrangig der Weiterentwicklung der Qualität im Arbeitsfeld dienen. Das Projekt „Methodenstudie: Qualität in der Kindertageseinrichtung (MS-Kita)“ konnte erste methodische Herausforderungen herausarbeiten. Eine vertiefte Auseinandersetzung damit, wie diese Herausforderungen bewältigt und insbesondere transparenter gemacht werden können, wäre an dieser Stelle gewinnbringend und für die Weiterentwicklung der Qualität im frühpädagogischen Bereich fruchtbar.

Literatur

- Anders, Yvonne/Roßbach, Hans-Günther (2014): Empirische Bildungsforschung zu Auswirkungen frühkindlicher, institutioneller Bildung: Internationale und nationale Ergebnisse. In: Braches-Chyrek, Rita/Röhner, Charlotte/Sünker, Heinz/Hopf, Michaela (Hrsg.): Handbuch Frühe Kindheit. Opladen/Berlin/Toronto, S.335–347.
- Anders, Yvonne/Roßbach, Hans-Günther/Tietze, W. (2016): Methodological challenges of evaluating the effects of an early language education programme in Germany. In: International Journal of Child Care and Education Policy, 10. Jg., Nr. 9.
- Ben-Arieh, Asher (2010): From Child Welfare to Children Well-Being. The Child Indicators Perspective. In: Kamerman, Sheila B./Ben-Arieh, Asher/Phipps, Shelley (Hrsg.): From Child Welfare to Child Well-Being. An International Perspective on Knowledge in the Service of Policy Making. 1. Auflage. Dordrecht, S. 9–24.
- Bertram, Hans (2016): Kindliches Wohlbefinden. Von Kinderarmut und Fürsorge zur kindlichen Teilhabe. In: Diskurs Kindheits- und Jugendforschung, 11. Jg., Nr. 3, S. 269–286.
- Bertelsmann Stiftung (2018): ElternZOOM 2018. Schwerpunkt: Elternbeteiligung an der KiTa-Finanzierung.
- Bundesministerium für Familien, Senioren, Frauen und Jugend (2018): Übereinkommen über die Rechte des Kindes. VN-Kinderrechtskonvention im Wortlaut mit Materialien.
- Deci, Edward L./Ryan, Richard M. (1993): Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. In: Zeitschrift für Pädagogik, 39. Jg., Nr. 2, S. 223–238.
- Dunn, Loraine (1993): Proximal and distal features of day care quality and children's development. In: Early Childhood Research Quarterly, 8. Jg., Nr. 2, S. 167–192.
- Eberlein, Noemi/Schelle, Regine (2018): Aspekte kindlichen Wohlbefindens als Indikator für pädagogische Qualität? Eine Dokumentenanalyse von Qualitätsmessinstrumenten für den frühpädagogischen Bereich. In: Diskurs Kindheits- und Jugendforschung, 13. Jg., Nr. 4, S. 387–402.
- Hallam, Rena/Fouts, Hillary/Bargreen, Kaitlin/ Caudle, Lori (2009): Quality from a Toddler's Perspective: A Bottom-Up Examination of Classroom Experiences. Early Childhood Research & Practice, 11 Jg., Nr. 2. <http://ecrp.uiuc.edu/v11n2/hallam.html> (26.02.2018).
- Harms, Thelma/Clifford, Richard M./Cryer, Debby (2015): Early childhood environment rating scale. 3. Aufl. New York.
- Kalicki, Bernhard (2015): Pädagogische Qualität und Qualitätssteuerung: Konzepte und Strategien. In: Kalicki, B./ Wolff-Martin, C. (Hrsg.): Qualität in aller Munde. Freiburg, S. 12–22.
- Kuckartz, Udo (2016): Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. – Weinheim/Basel.
- Laevers, Ferre/Moons, Julia/Declercq, Bart (2012): A Process-Oriented Monitoring System for the Early Years (POMS). Leuven.
- Lamnek, Siegfried (2005): Qualitative Sozialforschung. Weinheim/Basel, 4. Auflage
- Mashford-Scott, Angela/Church, Amelia/Tayler, Collette (2012): Seeking Children's Perspectives on their Wellbeing in Early Childhood Settings. In: International Journal of Early Childhood, 44 Jg., Nr. 3, S. 231–247.
- Mayer, Daniela/Beckh, Kathrin (2018): Erfassung pädagogischer Qualität in Kindertageseinrichtungen. In: Frühe Bildung. 7 Jg., Nr. 2, S. 67–76.

- Melhuish, Edward C. (2001): The quest for quality in early day care and preschool experience continues. In: *International Journal of Behavioral Development*, 25 Jg., Nr. 1, S. 1–6.
- Pianta, Robert C./La Paro, Karen M./Hamre, Bridget K. (2011): *Classroom Assessment Scoring System (CLASS) manual PRE-K*. 7. Aufl. Baltimore/London/Sydney.
- Reichertz, Jo (2015): Die Bedeutung der Subjektivität in der Forschung. In: *Forum: Qualitative Sozialforschung*, 16 Jg., Nr. 3, Art. 33. <http://www.qualitative-research.net/index.php/fqs/article/viewFile/2461/3889> (08.11.2018).
- Reichertz, Jo (2016): *Qualitative und interpretative Sozialforschung. Eine Einladung*. Wiesbaden.
- Schelle, Regine/ Eberlein, Noemi/ Durand Judith (i.E.): Aspekte einer „Sensitiven Responsivität“ als Qualitätsmerkmal von Interaktionen – Ergebnisse einer Dokumentenanalyse standardisierter Beobachtungsinstrumente für den frühpädagogischen Bereich. In: *Forschung in der Frühpädagogik Band 12*.
- Smidt, Wilfried (2012): *Zielkindbezogene pädagogische Qualität im Kindergarten. Eine empirisch-quantitative Studie*. Münster.
- Smidt, Wilfried/Roßbach, Hans-Günther (2015): Educational process quality in preschools at the individual child level: findings from a German Study. In: *Early Child Development and Care*, 186 Jg., Nr. 1, S. 78–95.
- Steinke, Ines (1999): *Kriterien qualitativer Forschung*. Weinheim/München.
- Suchodoletz, Antje von/Gunzenhauser, Catherine/Larsen, Ross A. A. (2015): Die Beobachtung von Interaktionen im Kindergartenalltag. Das Individualized Assessment Scoring System (in-CLASS). In: *Frühe Bildung*. 4. Jg., Nr. 4, S. 211–217.
- Tietze, Wolfgang/Roßbach, Hans-Günther/Nattefort, Rebecca/Grenner, Katja (2017): *Kindergarten Skala (KES-RZ)*. Revidierte Fassung mit Zusatzmerkmalen. 4. Aufl. Kiliansroda.
- Weltzien, Dörte/Fröhlich-Gildhoff, Klaus/Strohmer, Janina/Rönnau-Böse, Maïke/Wünsche, Michael/Bücklein, Christina/Hoffer, Rieke/Tinius, Claudia (2017): *Gestaltung von Interaktionen – Ein videogestütztes Evaluationsinstrument*. Manual. Weinheim/Basel.
- Wertfein, Monika/Mayer, Daniela (2018): Pädagogische Qualität in der Tagesbetreuung für Kinder bis drei Jahre. In: Schmidt, Thilo/Smidt, Wilfried (Hrsg.): *Handbuch Empirische Forschung in der Pädagogik der frühen Kindheit*. Münster/New York, S.389–406.

Gefördert vom



Bundesministerium
für Familie, Senioren, Frauen
und Jugend

Deutsches Jugendinstitut e. V.
Nockherstraße 2
D-81541 München

Postfach 90 03 52
D-81503 München

Telefon +49 89 62306-0
Fax +49 89 62306-162

www.dji.de